universität
wien

Dissertation

Titel der Dissertation

„Carnap's early semantics"

Verfasser

MMag. Georg Schiemer

angestrebter akademischer Grad

Doktor der Philosophie (Dr. phil.)

Wien, Juni 2010

Studienkennzahl lt. Studienblatt: 092 296

Dissertationsgebiet lt. Studienblatt: Philosophie

Betreuer: Ao. Univ. Prof. Dr. Richard Heinrich

Contents

Introduction

In recent years an intensified interest in the technical and historical details of Carnap's philosophy of mathematics and logic could be witnessed. In my dissertation I take up this debate and focus on his early formative contributions to a theory of formal semantics. According to a dominating view regarding Carnap's work, a 'semantic turn' - eventually culminating in his three volume book project *Series in Semantics* (Carnap (1942), Carnap (1943), Carnap (1947)) - took place only after his confrontation with Tarski's program of "scientific semantics" and the acceptance of serious limitations of the syntactical method in *Logical Syntax of Language* (1934). This has recently been put into proper perspective in a number of contributions (Awodey and Carus (2001), Awodey and Reck (2001), Bonk and Mosterin (2000), Reck (2007), and Goldfarb (2005)). It is noted that there had been a short phase of theoretical engagement with semantic issues in Carnap's work well before his work on logical syntax. These considerations were closely linked to a larger project on the methodology and formal reconstruction of modern axiomatics. Carnap's contributions to the topic include two published articles, (Carnap (1930) and Carnap and Bachmann (1936)), as well as the recently published typescript *Untersuchungen zur allgemeinen Axiomatik* of 1928 (Carnap 2000). These works include early formal definitions of the notions of 'model', 'model extension', and 'logical consequence' as well as a number of theorems concerning the metatheory of axiomatic theories. Whereas Carnap's metatheoretic results have been examined in the literature (see above), no closer attention has so far been dedicated to Carnap's conception of the basic semantic notions in his theory of axiomatics. In my dissertation, I provide a detailed conceptual analysis of the technical details of Carnap's early semantics and contextualize his results in their historic and intellectual environment. By drawing to unpublished material from Carnap's *Nachlass*, in particular his notes on the second projected part of *Untersuchungen*, I also discuss certain interpretative issues related to his heterodox conception of formal models and the tacit semantics for his type-theoretic logic.

The dissertation comprises three chapters that are intended for separate publication. Chapter 1 critically surveys Carnap's attempt to provide a logical reconstruction of formal axiomatics in a type-theoretic setting. The chapter has two aims. First, to analyze the kind of mathematical structuralism and the specific semantics implicit in Carnap's account. Second, to show that Carnap's attempt to capture one of the central semantic innovations of formal axiomatics, i.e.

the definition of classes of structures, with his notion of “*Explizitbegriffe*” is severely limited due to these semantic assumptions.

Chapter 2 examines Carnap’s theory of extremal axioms with a view to the auxiliary concepts of models and model extensions introduced in Carnap (2000). The motivating question here is: “Does Carnap’s early conception of models allow domain variation?” My discussion focuses on these aspects and responds to a commentary in Hintikka (1991) which states that, for several reasons, Carnap’s version of extremal axioms is seriously flawed. I present an alternative interpretation of Carnap’s theory and show that Hintikka’s claims cannot be upheld once a more balanced account of Carnap’s early conception of models is provided. I describe in detail two conventions used by Carnap to allow domain variation for his formal models. Finally I show that the notion of model extension underlying his theory of extremal axioms, if properly understood, clearly captures the informal use of model extension in the axiomatic tradition.

Chapter 3 also discusses Carnap’s theory of extremal axioms, but this time from a more historical perspective. The aim in this chapter is to provide a historically sensitive reconstruction of the evolution of Carnap’s theory between 1928 and 1936. The main focus is thereby set on the unpublished documents of Part Two of *Untersuchungen* from his *Nachlass*. Carnap’s theory is surveyed with an eye to clarifying two points: the first concerns the main mathematical influences, most importantly by the set theorist Abraham Fraenkel, on Carnap’s specific formalization of minimal axioms. The second point regards Carnap’s specific understanding of the connection between extremal axioms and other metatheoretic notions of completeness, most notably the categoricity of an axiom system.

Each chapter of the dissertation is concerned with a specific aspect or a specific interpretive issue of Carnap’s early formal semantics. As such, each is self-contained and readable as an independent paper. There is, however, a broader perspective shared in all chapters that can also be considered a more general motivation of this dissertation. It concerns a certain reservation against the universal validity of a still dominant interpretive scheme in the history of logic, viz. van Heijenoort’s (and Hintikka’s) classic distinction between the ‘*universalist*’ and the ‘*model-theoretic*’ tradition (e.g. van Heijenoort (1967), Hintikka (1988)). Briefly, their claim is that there exist two essentially incompatible traditions in the history of logic and that each of the major proponents of modern logic – from Frege, Russell, Carnap to Quine – can be neatly placed either within the one or the other camp. As I show in the dissertation, Carnap’s work in *Untersuchungen* from the late 1920s clearly does not fit within this dualist picture. To be clear, there are on the one hand (as will be shown in Chapter 1) strong '*universalist*' leanings in Carnap's work, particularly due to the logicist background in his treatment of axiomatics. Nonetheless, this does not prevent him on the other hand from developing an account of formal models for axiomatic theories that anticipates constitutive features of modern formal semantics (see Chapter 2). Thus, at least in this case, the van Heijenoort-Hintikka interpretive framework, if conceived in a strong dichotomic sense, does not help to clarify but rather conceals the real conceptual transformations in the evolution of modern logic present in his work on axiomatics. As I hope to show in the following, only a more balanced account of Carnap's early semantics *between* the lines of the two grand logical traditions will allow specifying his particular contribution to the "semantic revolution" in logic in the first third of the last century.

Chapter 1: Carnap's Explizitbegriffe: axiomatics, structuralism, and logicism

1.1 Introduction

This chapter discusses Carnap's attempt in the late 1920s to provide a formal reconstruction of modern axiomatics.[1] One interpretive theme commonly addressed in the recent scholarly literature concerns Carnap's underlying logicism in his philosophy of mathematics in the 1920s, more specifically, his genuine attempt to "reconcile" the logicist approach of reducing mathematics to logic with the mathematical method of formal axiomatics.[2] It was by Reck and others that the central idea in Carnap's manuscript *Untersuchungen* was to balance a *Fregean* (or *Russellian*) foundational stance with the modern model-theoretic viewpoint introduced in Hilbert's *Grundlagen der Geometrie* (Hilbert 1899).[3] It was also shown that Carnap's attempt to provide a logicist reconstruction of axiomatics is limited in several ways. No closer attention, however, has so far been dedicated to some of the details of his attempted reconciliation.

Our principle aim in this chapter is to undertake a closer analysis of Carnap's theory of axiomatics from *Untersuchungen* and the related works, specifically of the impact of an implicit logicist assumption underlying his early semantics for axiom systems. In this respect the concept of an "Explizitbegriff" ("explicit concept") will stand in the center of attention. Carnap's notion of "Explizitbegriffe" is particularly insightful in the evaluation of his reconciliatory approach between formal axiomatics and logicism. On the one hand it closely mirrors Frege's own reconstruction of Hilbert's method in Hilbert (1899). On the other hand, the concept is clearly introduced to capture formally one a central semantic innovation for formal axiomatics, namely the idea that an axiom system defines its class of structures. Moreover, the concept implies a structuralist account of mathematical theories that has leanings to both traditions.

The subsequent discussion of these points is guided by several interpretive questions: First, in what sense is Carnap's version of a mathematical structuralism expressed here effectively

[1] Carnap's early contributions to axiomatics comprise two published articles, Carnap (1930) and Carnap and Bachmann (1936), as well as Carnap (1929) and the posthumously published typescript *Untersuchungen zur allgemeinen Axiomatik* (Carnap 2000) written around 1928.
[2] Compare in particular Awodey and Carus (2001), Reck (2004), Reck (2007), and Bonk and Mosterin (2000).
[3] For instance, Awodey and Carus (2007) characterize *Untersuchungen* as a "large-scale project to reconcile axiomatic definitions with logicism, and transform implicit into explicit definitions." (ibid, 29)

comparable to the structuralism implicit in the modern axiomatic tradition, in particular in (Hilbert 1899)? Second, what are the semantic assumptions attached to this structuralist approach? Here, in particular, a closer look at the specifics of Carnap's attempt to provide a formal semantics for axiomatic theories, viz. a theory of formal models, is insightful. Finally, given his logicist background, does his reconstruction meet the methodological and semantic desiderata of the modern axiomatic tradition?

The chapter will be structured in this way: Following a brief presentation of the crucial semantic innovations in formal axiomatics (Section 1.2.1) and of Frege's early attempt of a logicist presentation of formal axiomatics (Section 1.2.2), Carnap's account of axiomatics will be outlined in Section 1.3. Section 1.3.1 introduces his notion of "Explizitbegriffe" of an axiom system. In Section 1.3.2 the kind of mathematical structuralism involved in this theory will be discussed. Carnap's early notion of a formal model as defined in *Untersuchungen* is presented in Section 1.3.3. Given this rather expository part, we will show in Section 1.4.1 that Carnap's attempt to reconstruct the semantics of axiomatic theories is based on a specific logicistic premise, namely a logical definability condition for formal models. Carnap's definability assumption is compared with similar accounts suggested by Tarski (Section 1.4.2) and Bachmann (1.4.3). We will argue in Section 1.5 that this premise has strong limiting effects for Carnap's overall attempt to express the model-theoretic semantics of formal axiomatics. This will be shown for cases where Carnap employs quantification over the models of a theory is necessary (Section 1.5.1). It also effectively restrains the tacit semantics of his underlying type-theoretic background language (Section 1.5.2). In Section 1.6, it will finally be shown that, as a consequence of this, Carnap's "Explizitbegriffe" fail to capture adequately the semantic innovations introduced by Hilbert and others in modern axiomatics.

1.2 "Methodological turn(s)" of formal axiomatics

It is not intended here to provide an all-encompassing picture of the conceptual transformations implied by the rise of formal axiomatics at the turn of the last century. Such a task would have to discuss the epistemological and methodological innovations in the works of Dedekind, Peano and the Italian 'Peanists', Pasch, and, most importantly Hilbert in contrast to the traditional axiomatic method.[4] Crucial points in this transformation include the change of the "epistemic status" of axioms, the new methodology for axiom choice and justification, the logical rigorization of proofs, as well as the strict "separation of the mathematical and logical [spheres] from the spatial-intuitive [sphere]" (Bernays 1922, 192), e.g. the complete abstraction from intuition in mathematical concept formation.

1.2.1 Three semantic innovations

What is relevant for our discussion is the fact that these "new methodological turn(s)" in formal axiomatics also led to a second revolution, namely a fundamental transformation of the semantics of a mathematical theory. With the abstraction from the intuitive sources of a theory, modern model-theoretic semantics could emerge. This is probably most explicit in Hilbert's *Grundlagen der Geometrie* (Hilbert 1899).[5] For the present purposes, it will be instructive to pin down three distinct but closely interrelated semantic innovations.

A first important difference to the classical axiomatics lies in the *disinterpreted* character of the primitive terms of a mathematical theory. Given this schematic treatment, the primitive concepts (like, in Hilbert's work, 'point', 'line', 'between' etc.) can then be reinterpreted in different ways.

A second innovation resulting directly from this new, formal treatment of the mathematical terminology is the implementation of a theory of models, in Hilbert's words "systems of things" that to allow the reinterpretation of implicitly defined mathematical primitive terms. In Hilbert's *Grundlagen*, the use of models (and the idea of model variation) plays a methodologically fruitful role in his theory, more precisely in the metatheory for his axiom system: different analytic models ranging over a set of algebraic numbers and over the complete ordered field of real numbers are devised in order to prove the consistency of the axiom system and different independence results for his axioms.[6]

---

[4] See de Jong and Betti (2010) for a detailed account of the "classical model of science" preceding the modern axiomatic method. For a discussion of the historical evolution of modern axiomatics see e.g. Grattan-Guinness (2000).

[5] For a more general account of the emergence of modern semantics in the axiomatic tradition see Hintikka (forthcoming). For a recent survey on the semantic innovations in Hilbert's book and, more generally, on Hilbert's new conceptualization of the relations of intuitive, analytic and axiomatic geometry compare Hallett (2008).

[6] See Hilbert (1899, §2). Compare Hallett (2008) on the connection between the two innovations: "The basic technique which Hilbert adopted for this investigation is that of modelling, more strictly, of translating the theory to be investigated into another mathematical theory. For this, it is essential (…) that the primitive concepts are not tied to their usual fixed meanings; they must be free for *reinterpretation*." (ibid, 211)

A third innovation less explicit in Hilbert (1899) shows the algebraic background of his approach. It consists in the fact that an axiom system defines its class of models.[7] This idea is best visible in Hilbert's presentation of his *axiom of completeness* as a continuity axiom in later editions of *Grundlagen*. The axiom implicitly involves the quantification over the possible models of his theory.[8] In modern terms, this more or less corresponds to the model-theoretic notion of the "class of structures" defined by an axiom system. Briefly, the class defined is the one consisting of all models that satisfy the axiom system.[9]

One can schematically present the three semantic innovations of formal axiomatics in this way:

(1) The <u>schematic conception of the primitive terms</u> of a theory, and thus their reinterpretability.

(2) The <u>use of formal models</u> for the interpretation of a theory (as well as the related notions of truth or satisfaction in a model). This includes the systematic use of <u>model construction</u> and <u>variation</u> for establishing metatheoretic results.

(3) The notion of a <u>definition of a class of models</u> of an axiomatic theory.

The reason for pinpointing (1)-(3) is to refer to them in the following as adequacy conditions for a logical reconstruction of the modern axiomatic method. Thus, any attempt to provide a formal presentation of axiomatics can only be considered successful if it manages to capture in an adequate manner each of these three semantic desiderata.

As mentioned in Introduction, the specific case to be tested in terms of (1)-(3) will be Carnap's explication of (*Hilbertian*) axiomatics. Before turning to his approach, however, a closer look at his immediate predecessor Frege and the latter's attempt to come to terms with Hilbert's new conception of axiomatics is instructive.

[7] Compare Hintikka on this point: "The basic idea of the [GS: axiomatic] method is the capture of a class of structures as the models of an axiomatic system." (Hintikka forthcoming) He further underlines the general relevance of this semantic approach in modern axiomatic thinking: "The very purpose of the axiomatic method is to study some class of structures by constructing an axiom system whose models they are." (ibid.)

[8] See Hilbert (1900, 25). Compare Section 3.2 for a fuller discussion of Hilbert's axiom.

[9] Compare Hodges (1997, 30-37) for a technical presentation of the concept.

1.2.2 Frege on Hilbert (1899): a shared assumption?

The appropriate starting point for our discussion of the adequacy of Carnap's logicist reconstruction of formal axiomatics is the notorious controversy between Hilbert and Frege at the turn of the last century.[10] Their widely discussed correspondence is illustrative since it provides a historical document that highlights some of the main conceptual differences between two opposing logical camps: Frege, on the one side, as the founder of logicism and proponent of a "universalist" conception of logic; Hilbert as an early proponent of the "model-theoretic tradition" on the other side.[11] I will not go here into the details of Frege's objections to Hilbert's conception of axioms as implicit definitions, nor of Frege's own conception of axioms and primitive terms in geometry and of his heterodox views on axiomatic metatheory (in particular concerning Hilbert's indirect consistency proofs). It suffices to note that due these differences, their conceptions were sometimes characterized as simply irreconcilable.

Nevertheless, as was recently stressed by various commentators, the conceptual differences between them are far smaller than often stated. In particular, in view of the semantic innovations of the modern axiomatic approach, it was pointed out that Frege's approach to axiomatics is not, in principle, incompatible with the model-theoretic approach.[12]

Concerning condition (1), one can find in Frege a convention to by-pass the schematic conception of logic necessary for Hilbert's treatment of axiomatics primitives as formal terms. This is simply to treat primitives as free (higher-order) variables of the logical language in use, thus allowing a similar kind of generality constitutive for Hilbert's approach. In Carnap's notes on Frege's lecture "Logic in Mathematics" from 1914 Frege's intuition underlying this presentation is made clear. In a remark on Hilbert's axioms Frege states that:

> Later Hilbert not only uses the word "between" with a different meaning, he also often uses "point," "line," "plane," (…) differently from Euclid. What is unclear, then, is this: he never says so explicitly and he never makes clear how else he understands them. *Often he uses expressions as indicating indefinitely, just as we use letters*." (Reck and Awodey 2004, 166, my emphasis)

[10] See Frege (1980) for the original correspondence between the two mathematicians. Compare Blanchette (2009) and the references given there for recent commentary on the debate.

[11] As is well known, this distinction between two logical traditions goes back to van Heijenoort (1967). For more recent discussions see Goldfarb (1979) and Hintikka (1988).

[12] Compare, e.g., Blanchette (2009) and Tappenden (1997).

"Letters" are understood here as free variables of the underlying formal language. It should be noted that this way of symbolizing the primitives of an axiomatic theory was common practice at that time. A similar treatment of *schematic* terms as variables can be found in the works of Russell, the American postulate theorists, and - as we will see - also Carnap. In fact, the same convention was also suggested by Hilbert himself three decades after his controversy with Frege. In Hilbert and Bernays (1934), a partial formalization of the axiom system of Hilbert (1899) is developed. Here, they state: "Speaking properly, the primitive relations have for formal axiomatics the role of *variable* predicates." (ibid, 7)[13]

Frege says less on condition (2), i.e. the systematic use of models for studying the (metatheoretic) properties of an axiomatic theory. In fact, it seems as though the very idea of a model as introduced in Hilbert's *Grundlagen* was foreign to Frege's understanding of axiomatics. For instance, in a letter dated January 6, 1900, he discussed a possible modification of Hilbert's notion of implicit definitions. In comparing Hilbert's version to that of Stolz, he critically adds:

> And to repair the damage that I believe to find in your definitions one will have to proceed in a similar way; except that it will be much more difficult because what is in question is not just a single system but three systems (of points, lines and planes) with their manifold relations. Incidentally, what are you calling a system here? I believe it is the same thing which is elsewhere called a set or a class and which is best called the extension of a concept. (Frege 1980, 47)

This clearly suggests that Hilbert's informal notion of a model as a "system of things" that satisfies the axioms was (to put it mildly) not fully acknowledged by Frege.

Despite this obvious disinterest in Hilbert's strategic use of models, Frege *does* make an interesting contribution to a formal presentation of (3). To understand his approach, it should be noted first that the modern version of (3) in terms of a "class of models" can be given an alternative formulation that more adequately catches the historical accounts of the idea. In this version, an axiom system defines not a class of models but an abstract structure (or a higher-order property) that is shared by (or applies to) all possible models of the theory. Thus, in the case of Hilbert's theory of Euclidian plane geometry, (3) can be understood as the specification of a general (relational) structure instantiated by all possible Euclidian planes. More generally, one can rephrase (3) by the equivalent version:

---

[13] "In Wahrheit aber haben für die formale Axiomatik die Grundbeziehungen die Rolle von *variablen* Prädikaten." (Hilbert and Bernays 1934, 7) Compare Bonk and Mosterin (2000, 16).

(3*) An axiom system <u>defines an abstract structure</u> shared by all models satisfying the axioms.[14]

It was argued recently that there is no deep conceptual difference between Frege and Hilbert on condition (3) since the former in fact proposes a logical reconstruction of (3*).[15] The crucial convention used by him is the ascent to higher-order concepts. Frege holds that while the single axioms, as assertive statements, cannot be understood as (implicit) definitions in Hilbert's sense, the axiom system nevertheless provides an explicit definition of a "second-level concept" that holds for all first-order concepts defined in the theory. In the letter cited above, Frege explicitly discusses this point:

The characteristic marks you give in your axioms are apparently all higher than first-level (…). It seems to me that you really want to define second-level concepts but do not clearly distinguish them from first-level ones. (ibid, 46)

In a letter to Liebmann, his view of Hilbert's (1899) alleged confusion of the hierarchy of concepts is stated in closer detail:

It seems to me that Prof. Hilbert first had the idea in mind of defining second-level concepts; but he does not distinguish them from first-level concepts. And this explains what Hilbert's explanations always leave unclear: how the same concept seems to be defined twice. It is just not the same concept. At first it is a second-level concept, afterwards a first-level concept falling into it. It is a mistake to mix up the two and to use the same word (e.g. 'point') in both connections. (ibid, 93-94)

---

[14] (3*) is closely captured by Corcoran's insightful remark on the historical notion of a structure (of a domain) as opposed to its modern understanding of structure *as* a model: "From a philosophical and historical point of view it is unfortunate to that the term 'mathematical structure' [GS: as the class of mathematical objects and functions and sets over them] is coming to be used as a synonym for ‚mathematical system'. In the earlier usage (…) two mathematical systems having totally distinct elements can have the same structure. This in this sense a structure is not a mathematical system, rather a structure is a "property" that can be shared by individual mathematical systems. At any rate a structure is a higher order entity. (…) For mathematical purposes it would be possible to 'identify' a structure with the class of mathematical systems having that structure, but such 'identification' may tend to distort one's conceptual grasp of the ideas involved." (Corcoran 1980, 188) For a detailed history of the notion of 'structure' in modern mathematics and logic compare also Mancosu (2006, 210).

[15] See e.g. Blanchette (2009) and Tappenden (1997). Blanchette paraphrases (3*) in terms of a higher-order relation: "A set *AX* of sentences containing *n* geometric terms defines an *n*-place relation R_{AX} holding of those *n*-tuples which, when taken respectively as the interpretations of *AX*'s geometric terms, render the members of *AX* true." (Blanchette 2009 (SEP))

The idea outlined here is that axiomatic primitives like '*point*' effectively refer to two concepts. In Frege's first sense, '*point*' expresses a first-order concept that applies to a fixed class of elements. In addition, it also refers to a second-order concept, i.e. a concept "in" which all first-level concepts '*point*' (with a fixed extension) that meet the conditions specified by the axioms fall. The second-level concepts, say '$point^2$', does not denote a specific system of points, but a class of systems of points, i.e. it holds for (extensionally) different concepts '$point^1$'.[16] Thus, Hilbert's axiom system is reconstructed as providing an explicit definition of the second-order concept of "Euclidian Space" under which all geometric systems satisfying the axioms fall.[17]

Now, despite Frege's clear reservations against the Hilbert's use of "systems of objects" with their "manifold relations", it seems only a small step from this idea of higher-order concepts defined by the axioms to a properly semantic notion like (3*). As is suggested by Blanchette and others, there is thus no deep disagreement between Frege and Hilbert on this point. The actual differences between their accounts result only from two different ways to express essentially the same semantic idea.[18] In particular, in Frege's case, the main tool to express the second- or higher-order concepts explicitly defined by an axiom system is that of quantification in a higher-order pure logic (henceforth HOL). Quantification over individuals in sentences like "For all *x*, *x* is point and …" specifies a specific class of elements designated by 'point^1'. Second-order quantification in sentences like "For all classes of points, …." then expresses the second-level concept which holds for all concepts '$point^1$'. Now, Frege is nowhere explicit on how the higher-order concept defined by an axiom system is to be formalized. Nonetheless, the idea can easily be recast in HOL. A possible tentative formulation of Frege's idea is this: an axiom system *AS* specifies a concept $\mathbf{P^2}$ such that all concepts satisfying *AS* fall under this concept, i.e. formally, $\mathbf{P^2}(X^1) \leftrightarrow (\forall X^1)(\boldsymbol{AS}(X^1))$.
I will return to this specific interpretation of Frege's view in Section 5. For the moment, it suffices to say that *prima facie*, there seems to be no deep conceptual difference between the informal model-theoretic version of (3) in Hilbert's *Grundlagen* and the one outlined by

[16] Essentially the same idea can be found in the latter's notes on Frege's 1914 lecture (Awodey and Reck (2004, 163-166).
[17] Compare Reck (2004, 159-160) on this point.
[18] This affinity was already pointed out in Stein (1988). Stein discusses a variant of (3*) in Hilbert (1899): "(…) hence the axioms taken together constitute a *definition*. Of what? Of a *species of structure*." Then, he remarks on Frege's "undervalued insight" concerning nature of this definition: "(…) a Fregean "second-level concept" simply *is* the concept of a species of structure." (ibid, 254)

Frege. More generally, one could be inclined to argue that - despite the underlying, deeper divergences between the mathematicians on the nature of axioms, the role of intuition in geometry, the correct conception of logic, etc. - Frege's account does in fact present a valid reconstruction of Hilbert's semantics for his theory of geometry. At least two of the crucial semantic desiderata of formal axiomatics, namely (1) the reinterpretability of mathematical terms as well as (3*) the classification of an abstract structure by the axiom system are captured in the former's reconstruction.

With this background we will turn to Carnap's formal presentation of axiomatics. Our aim will be to see whether this claim concerning the adequacy of a Frege-style logical reconstruction can be upheld once the technical details left implicit in Frege's account are fully specified.

1.3 Carnap's axiomatics as "applied logistic"

We know from Carnap's lecture notes of Frege's 1914 lecture that he was well acquainted with Frege's understanding of (*Hilbertian*) axiomatics. Moreover, the continuity with Frege's conception is also more than evident in Carnap's own work on the axiomatic method, from the informal account in "Eigentliche und Uneigentliche Begriffe" (Carnap 1927) to more formal presentations in *Abriss* (Carnap 1929) and *Untersuchungen* (Carnap 2000).[19] In particular, it is not difficult to identify several points of contact with Frege's reconstruction of the semantic innovations made by Hilbert. Generally, it can be safely said that the main rationale for Carnap's work was to provide a formal explication of Frege's intuitions and, in a way, to reconcile the axiomatic method with the latter's logicist approach to mathematics.

1.3.1 "Explizitbegriffe" for axiom systems

Carnap's understanding of modern axiomatics is informally outlined in the Section 2 of *Untersuchungen* (See Carnap, 2000, 87-90).[20] The basic idea is that a formal axiom system is symbolized in a "basic discipline" ("Grunddisziplin"), viz. a fully interpreted type-theoretic language (henceforth STT for simply theory of types). The "primitive symbols" ("*Grundzeichen*") do not have a fixed interpretation but can be freely interpreted. In the

[19] See in particular Awodey and Carus (2001) and Reck (2004) for a closer discussion.
[20] Compare again Awodey and Carus (2001) for a similar discussion of Carnap (2000).

following passage the difference to "contentual axiomatics" ("inhaltliche Axiomatik") is explicitly addressed:

> One can also understand the primitive concepts as unspecified elements and relations of an unspecified domain where it is only stipulated how they relate to each other as specified in the axioms. If in different domains elements and relations are found that satisfy these formal specifications, then the axiom system can be applied to each one of these domains; in each one of these domains also the consequences of the axiom system hold given the respective interpretation. (ibid, 88)[21]

Carnap continues to remark that this version is the "method commonly used by the mathematician," most notably Hilbert's conception in Hilbert (1899).

As mentioned above, his logical presentation of formal axiom systems in STT is closely inspired by Frege's approach. Just as was indicated by Frege, the primitive symbols (expressing the primitive concepts) of a system are symbolized by free predicate variables (of a specified type). Axioms are conceived as propositional functions that range over the primitive symbols. The axiom system is also presented as a propositional function that expresses the conjunction of its axioms and contains a "model variable" ("Modellvariable") of a specified type, e.g.:[22]

$$\mathrm{f}(R_{\mathrm{n}}, S_{\mathrm{n}}, T_{\mathrm{n}}) \sim \mathrm{f}(M_{\mathrm{n+1}})$$

Given this formalization, we can see how Carnap intends to meet the semantic desiderate described above. It is clear from the passage just cited that condition (1), i.e. the reinterpretability of the primitive terms is clearly met here. Formally, this is done (as by Frege before) by the symbolization of the primitive terms in terms of free variables of STT.

We will postpone the discussion of Carnap's attempt to tackle condition (2) to Section 2.3. There, it will be shown that he, unlike Frege, had a distinct idea of how Hilbert's "systems of things" can be recast in formal terms.

---

[21] "Oder aber man faßt die Grundbegriffe auf als unbestimmte Gegenstände und Beziehungen eines unbestimmten Gebietes, von denen nur festgelegt wird, daß sie sich so zueinander verhalten, wie es in den Axiomen bestimmt wird. Finden sich auf verschiedenen Gebieten Gegenstände und Beziehungen, die diese formalen Bestimmungen erfüllen, so kann das Axiomensystem auf jedes dieser Gebiete bezogen werden; auf jedem dieser Gebiete gelten dann auch in entsprechender Deutung die Folgerungen des Axiomensystems." (ibid, 88)

[22] This conception of axiomatic systems as propositional functions was first introduced in published form in Carnap (1929, 70-72). It is important to note here, however, that this convention to treat axiomatic theory as propositions functions was quite common at the time Carnap was working on the topic. Carnap is referring to Huntington (1913) for the first explication of this position. Compare Jané (2006, 18-27) on a similar conception in the works by the 'Peanists' Padua and Pieri. See Mancosu (2006, 212-216) for a broader historical survey of this "widespread conception". Finally, see Section 3.4.1 for a similar account in Fraenkel (1928).

Concerning condition (3) (or (3*)), Carnap's attempt was to provide a logical reconstruction of Frege's original idea that an axiom system explicitly defines a higher-order concept. His formal version of this is most explicitly discussed in his logic manual *Abriss der Logistik* (Carnap 1929). In the second part of the book, titled "Applied Logicist," Carnap describes – in line with the *Untersuchungen* – two ways of understanding an axiom system: (1) as an interpreted theory and (2) as a formal system in Hilbert's sense. Accordingly, two formal presentations in STT can be provided for the theory in question: the primitive terms in (1) are to be expressed via non-logical constants, in the case (2), as we have seen, via free variables (Carnap 1929, 71).[23] Given this, Carnap is quick to point out that an axiom system in question not only provides implicit definitions of its primitive terms, but also an explicit definition of something comparable to Frege's second-level concepts, namely an "Explizitbegriff". I quote the central passage on this point in Carnap (1929) in full length since it will provide the basis for our further discussion:

In every AS (…) one or more primitive concepts are implicitly defined as "improper concepts". In addition, a specific, namely a logical concept is defined explicitly by any AS of n primitive concepts: the "explicit concept" ("Explizitbegriff") for the AS; this concept is a class for $n = 1$, otherwise an n-ary relation. If $x, y, \ldots \alpha, \beta, \ldots P, Q, \ldots$ are the primitive variables of the AS and if we name the conjunction of axioms (that is a propositional function) AS($x, y, \ldots \alpha, \beta, \ldots P, Q, \ldots$), then the definition of the explicit concept for this AS is:

$$\hat{x}, \hat{y}, \ldots, \hat{\alpha}, \hat{\beta}, \ldots, \hat{P}, \hat{Q}, \ldots \{AS(x, y, \ldots, \alpha, \beta, \ldots, P, Q, \ldots)\}$$

The explicit concept of Peano's AS of the numbers e.g. is the class of number sequences that satisfy the AS; this is the logical concept *prog* (class of the progressions (...)). (ibid, 71-72)[24]

[23] Compare the following passage: "Jedes AS (…) ist zunächste ein System von Zeichenkombinationen, für dessen Deutung dir folgenden Möglichkeiten vorliegen: 1. Die Zeichen der Grundbegriffe („Grundzeichen") bezeichnen bestimmte nichtlogische Begriffe; es sind also *nichtlogische Konstanten*; die Axiome und die deduzierten Lehrsätze sind Aussagen über diese Begriffe. 2. Die Grundzeichen haben an sich noch keine Bedeutung, sondern bekommen sie erst durch das AS: sie sollen das bezeichnen, was die durch das AS angebebenen Eigenschaften und Beziehungen hat („implizite Definition"); exakter ausgedrückt: die Grundzeichen sind *Variable*, die Axiome und Lehrsätze (…) Aussagefunktionen." (Carnap 1929, 71).

[24] "In jedem *AS* werden (...) ein oder mehrere Grundbegriffe als „uneigentliche Begriffe" implizit definiert. Es wird aber außerdem durch jedes *AS*, wenn es n Grundbegriffe hat, ein bestimmter, und zwar ein logischer Begriff, explizit definiert, der "Explizitbegriff" des *AS*, dieser Begriff ist für $n = 1$ eine Klasse, sonst eine n-stellige Relation. Sind etwa $x, y, \ldots \alpha, \beta, \ldots P, Q, \ldots$ die Grundvariablen des AS, und bezeichenen wir die Konjunktion der Axiome (also eine Aussagefunktion) mit AS($x, y, \ldots \alpha, \beta, \ldots P, Q, \ldots$), so lautet die Definition des Explizitbegriffs dieses AS: $\hat{x}, \hat{y}, \ldots, \hat{\alpha}, \hat{\beta}, \ldots, \hat{P}, \hat{Q}, \ldots \{AS(x, y, \ldots, \alpha, \beta, \ldots, P, Q, \ldots)\}$. Der Explizitbegriff des Peanoschen *AS* der Zahlen ist z.B. die Klasse der Zahlfolgen, die das *AS* erfüllen, das ist der logische Begriff *prog* (Klasse der Progressionen (…))." (ibid, 71-72)

It is important to have a precise understanding of what is stated here. According to Carnap, setting up an axiom system implies the definition of a (logical) “Explizitbegriff” as a class or relation whose members are the interpretations of the theory. For instance, an axiom system with one primitive binary relation R (as in his example of basic “Peano arithmetic”) explicitly defines the class of higher-level relations R_1, R_2, etc., i.e. the “class of progressions” that satisfy the axiom system.[25] For an axiom system with n primitive terms (e.g. Hilbert’s system in Hilbert (1899), the respective “Explizitbegriff” denotes a higher-level n-ary relation (in modern terms, a class of n-tuples).

Now, Carnap’s notion of “Explizitbegriffe” clearly mirrors Frege’s informal talk of higher-level concepts. It has to be understood as a Frege-style attempt to formally model or capture the semantic innovation (3), i.e. the concept that an axiom system specifies its own model class. Note that Carnap seems closer to the modern model-theoretic understating of (3) than Frege originally was. The former explicitly speaks of a “class” determined by the ‘Explizitbegriff’ consisting of those interpretations that “satisfy the AS”. Moreover, he uses set comprehension to present his concept formally.

1.3.2 A structuralist account

Carnap’s account of ‘Explizitbegriffe’ in *Abriss* can also be considered as a formal representation of the structuralist variant (3*). In his example of Peano arithmetic, he mentions the “class of number sequences”. In the same sentence he holds that the axiom systems also defines “the logical concept *prog*” (ibid, 72). Thus, the ‘Explizitbegriff’ is viewed here as a (logically defined) concept or property that applies to all interpretations of the theory. This seems to mirror closely the “structuralist methodology” (Reck and Price 2000) underlying the modern axiomatic method. Briefly, structuralism holds that mathematics is the study of structures or of structural properties of the objects that leaves unspecified the particular nature of objects.[26] This idea is at least implicit in Hilbert axiomatic approach. Here, a theory does not study the specific intuitive nature of particular geometrical objects or collections of objects but their ‘logical relations’, i.e. an abstract structure. In a second, independent step the models as the instances in such a relational structure are studied.[27]

[25] For a detailed presentation of Carnap’s theory of basic arithmetic see 2.5.1 as well as 3.4.3.
[26] See Shapiro (2000, 257-260).
[27] See again Bernays (1922, 192-193). Compare also Hintikka on the implicit structuralism in modern axiomatics: ”The structuralist orientation of modern mathematics naturally leads to the use of axiomatization. To understand a kind of structure, for instance the structure of a group, is to have an overview over all its

Carnap's own structuralist account of axiomatics becomes visible at several places in his work from the late 1920s. In Carnap (1929), in discussing different examples of 'Explizitbegriffe' of mathematical axiom systems, he explicitly mentions of the structures of a theory. Beside Peano arithmetic, he mentions different systems of projective geometry and topology, stating that:

(...) the explicit concept of a geometrical AS, e.g. of an AS of projective geometry (...) presents the logical concept of the respective type of space (e.g. the concept "projective space"). In this way, geometry can (like arithmetic (...)) also be presented as a branch of logistic instead of an applied case of logistic to a non-logical field. (ibid, 72)[28]

Here, as in the case of Peano arithmetic, the respective axiom systems in geometry define an abstract "logical concept" of a particular (Euclidian, projective or topological) space that is instantiated by different (Euclidian, projective or topological) models. Moreover, the specification of this concept can be effected in purely logical terms, thus geometry itself can be reconstructed as a "branch of logistic". (I will return to this latter point in Section 1.4.) This structuralist account of geometry is further specified in an interesting exchange between Scholz and Carnap in 1932 on the latter's conception of 'Explizitbegriffe' for axiomatic geometry in *Abriss*. Scholz, in a letter dated April 19, 1932, remarks on the "interpretation of Euclidian space":

I have thought about a lot (...) whether it makes sense to call every sextuple of argument values that satisfy the Hilbert axioms of an Euclidian space (...). If this makes sense, then the Euclidian space (of course in the sense of the class of Euclidian spaces) would be the explicit concept ("*Explizit-Begriff*") defined by the Hilbert system. Do you otherwise know an interpretation that gives us the One Euclidian space? (RC 102-72-07)[29]

---

instantiations. In an axiomatic system, this is accomplished by capturing all those structures as the models of the system." And: "For this method is the natural one for a structuralist. The axioms determine a class of structures as models of the axiom system." (Hintikka (forthcoming))

[28] "(…) der Explizitbegriff eines geometrischen AS, z.B. eines AS der projektiven Geometrie (…) stellt *den logischen Begriff der betreffenden Raumart* dar (z.B. Begriff "projektiver Raum"). Auf diese Weise kann auch die Geometrie, anstatt als Anwendungsfall der Logistik auf ein nichtlogisches Gebiet, dargestellt werden als Zweig der Logistik selbst (wie die Arithmetik (…))." (ibid, 72)

[29] "Ich habe (...) viel darüber nachgedacht, ob es sinnvoll ist, jedes Sextupel von Argumentwerten, die die Hilbert-Axiome befriedigen, einen Euklidischen Raum zu nennen (…). Wenn die Sache sinnvoll ist, so wäre der Euklidische Raum (natürlich im Sinn der Klasse der Euklidischen Räume) der durch das Hilbert-System definierte Explizit-Begriff. Oder kennen Sie eine Interpretation, die uns den Einen Euklidischen Raum liefert?" (RC 102-72-07)

Carnap's reply in a letter (dated May 3, 1932) is particularly insightful:

To me, it does not seem to be a question of truth but of convenience what one names "Euclidian space". In case one prefers to speak not of various Euclidian spaces but of *the* Euclidian space, it seems convenient to name as "Euclidian space" the structure shared by the Hilbertian sextuples. (RC 102-72-04)[30]

This clearly shows Carnap's structuralist conception of geometrical theories. (Euclidian) space is defined here as the abstract structure shared by all models of Hilbert's axiomatization.[31] Finally, Carnap's structuralism is also evident in his more general account in *Untersuchungen*. Here, the formal reconstruction of axiom systems and the model theory for them that makes essential use of the notion of structures. Briefly, a theory can express a number of 'structures' where each structure defines an isomorphism class of models (Carnap 2000, 121-123). If the axiom system defines only one structure (as in the case of Peano arithmetic and Hilbert's Euclidian geometry), it is categorical ("*monomorph*"). It is non-categorical ("*polymorph*") in case it defines more than one structure (ibid, 128-129).[32]

Given these remarks it is obvious that Carnap shares with the axiomatic tradition a structuralist understanding of mathematics. Nonetheless, our initial concern was to see how exactly his account of axiomatics can be positioned between Hilbert's model-theoretic viewpoint and a Frege's 'universalist' approach. With this in mind, a more refined view of what Carnap's version effectively amounts to will prove to be necessary.
For this task a comparison with two variants of mathematical structuralism in the contemporary discussion will be insightful. In the current debate, two (among other) different conceptions are frequently discussed, namely (a) "relative structuralism" and (b) "unversalist" or "eliminative structuralism".[33] Both accounts hold that the genuine subject field of a formal mathematical theory are abstract structures and not (as for instance in synthetic geometry,

[30] "Es scheint mir keine Wahrheitsfrage, sondern nur eine Zweckmässigkeitsfrage zu sein, was man als "euklidischen Raum" bezeichnen will. Möchte man lieber nicht von mehreren euklidischen Räumen, sondern von dem euklidischen Raum sprechen, so dürfte es zeckmässig sein, als „euklidischen Raum" die den Hilbertschen Sextupeln (…) gemeinsame Struktur zu bezeichnen. (…)." (RC 102-72-04)
[31] For a related account of formal geometry compare Carnap's dissertation (Carnap 1922), in particular of chapter 1 on "formal space".
[32] An example of a noncategorical theory discussed by Carnap is the axiom system of topology presented in Abriss §33 (Carnap 1929, 75-78). See 2.2 for a more detailed discussion of Carnap's conception of axiomatics.
[33] For an overview of the current debate see Shapiro (2000). For a detailed comparison of (a) and (b) see Reck and Price (2000, 348-354 and 354-361 respectively). For a further discussion of (a) compare Parsons (1990).

facts about the intuitive nature of the field of investigation). The central difference between them concerns alternative assumptions on the additional “semantic implications” of the structuralist doctrine (see Reck and Price 2000, 347). In (b), the abstract structures studied by a theory can be instantiated by models or set-theoretical structures. Hence, talk about abstract structures can thus be replaced by talk about a specific interpretation of a theory once the structure has been interpreted model-theoretically. The denotation of the primitives and the truth of a theory are specified relative to a given model. Moreover, the consequence relation between a sentence and the theory can be expressed via the notion of model-theoretic consequence (see ibid, 349-350). Thus, a relative structuralist position can be characterized as the combination of two assumptions: mathematical structuralism and standard model theory.

In contrast to this, position (b) does not turn to model theory for a semantic theory in support. Instead of interpreting the primitives of a theory and its domain of discourse in a particular model, the eliminative structuralist uses higher-order logic to circumvent talk about relative truth and specific model construction. Here, consequence is not defined model-theoretically but via the introduction of a universally closed conditional of the form

$$(\forall R)(\forall S)(\forall T)[AS(R, S, T) \rightarrow \varphi(R, S, T)]$$

where R, S, T are higher-order variables for the primitive terms of an axiomatic theory AS and $\varphi(R, S, T)$ is a specific sentence of the theory. The closed formula then essentially states that for any tuple of assignments to the primitives, φ follows logically from AS. In short, the sentence expresses a true sentence in HOL. Note that this too marks a specific semantics assumption, just like the adoption of model-theory in the case of (a). The respective difference in the semantic assumptions between the two versions is specified in Reck and Price (2000): the crucial characteristic of (a) is that - instead of a model-theoretic interpretation of a structure – use of higher-order quantification is made that allows generalization over all possible interpretations of the variables R, S, T of an axiom system. For the case of Peano arithmetic this has the following effect:

> (…) note that we again “abstract away” – now by *generalizing* – from what is peculiar about any particular model of PA_2. (…) In fact, (…) any reference to specific models of PA_2, or to particular objects and functions in them, has disappeared completely (even in a relative or model-theoretic sense). Instead, what we assert with an arithmetical statement p is now something about *all* objects, *all* one-place functions, and *all* one-place predicates and sets; since the main logical operators in q [GS: a universally closed sentence like the one above] are unrestricted universal quantifiers. (ibid, 356)

In comparison to (a), this difference is further specified in this way:

In relativist structuralism the basic idea is to pick a particular model of the relevant theory and to explain the reference of mathematical terms with respect to it; in universalist structuralism the appeal to particular objects and the use of referring terms is simply quantified away. (ibid, 361)

Given this, does Carnap's structuralist account of axiomatics (and in particular his reconstruction of 'Explizitbegriffe') side with one of these two variants and its respective "semantic implication"? As we will see, it is difficult to provide a clear-cut answer to this question. Consider again Carnap's two main intellectual influences on this theory of axiomatics: Hilbert's position in *Grundlagen* and Frege's logical reconstruction. It is clear that the semantic innovations in Hilbert (1899), in particular his introduction and systematic use of formal models (viz. (2)) make Hilbert an early proponent of relative structuralism. In contrast, Frege's suggested formulation of innovation (3*) via the use of higher-order concepts in 1900 and again in 1914 can be viewed as a kind of eliminative structuralism.[34] The difficulty in positioning Carnap's account within one of the two camps of mathematical structuralism results from the fact that it deliberately was formulated as a reconciliation between Hilbert's and Frege's viewpoints. We emphasized the strong continuity between Carnap's notion of 'Explizitbegriffe' and Frege's reconstruction of (3*). In fact, one can easily reformulate Carnap's use of set comprehension in the explication of this notion

$$\hat{x}, \hat{y}, ..., \hat{\alpha}, \hat{\beta}, ..., \hat{P}, \hat{Q}, ...\{AS(x, y, ..., \alpha, \beta, ..., P, Q, ...)\}$$

by a universally quantified sentence in his "applied logistic" of the form

$$(\forall x)(\forall y)\ ..(\forall \alpha)(\forall \beta)..(\forall P)(\forall Q)..(\mathrm{AS}(x, y, .. \alpha, \beta, .. P, Q, ..) \rightarrow \mathrm{Ex}_{\mathrm{AS}}(x, y, ..\ \alpha, \beta, .. P, Q, ..))$$

Here, $\mathrm{Ex}_{\mathrm{AS}}$ stands for the 'Explizitbegriff' of the axiom system in question. This clearly looks like an instance of the more general scheme of eliminative structuralism discussed above. In fact, given his examples of 'Explizitbegriffe' in Carnap (1929), one could view the quantifiers in the sentence as introduced in order to "abstract away" from the reference to particular models of the class of interpretations specified by the axiom system. (I will return to this interpretation in Section 5.)

Can one infer from this that Carnap was an eliminative structuralist? To a certain point, as far as the *Fregean* background of his reconstruction goes, we would argue that one can. Nonetheless, it would be a mistake to hold that he was an eliminative structuralist throughout.

[34] More generally, it is sometimes argued that eliminative structuralism and in particular the convention to treat mathematical primitives not as schematic terms but as variables, has close ties to the logicist accounts. Compare Parsons (1990, 311-315).

In particular, closer attention to his explication of several semantic notions in the context of his work on axiomatics shows that the situation is more intricate in his case. Where Carnap differs from Frege is in the fact that he, unlike the latter, makes a serious attempt to come to terms logically with Hilbert's semantic innovation (2), viz. the introduction and use of models of an axiom system.

1.3.3 Formal models in *Untersuchungen*

No closer account of what Carnap understands by an interpretation of an axiomatic theory is provided in Carnap (1929). Fortunately, in his *Untersuchungen* written around the same time as *Abriss*, a detailed formulation of the notion of a "formal model" of an axiom system can be found.[35] In §2.3 of the manuscript, the following definition is given:

If we write for f(R, S, T), in short f(M), and if R_1, S_1, T_1 are specific, e.g. arithmetic relations that are admissible values of the variables R, S, T, then we can also introduce an abbreviatory notation for the ordered system of relations R_1, S_1, T_1, say M_1. The *model* M_1 is then *a value of the model variable M*. Each of the relation variables R, S, T of the axiom system has a specific type (…); the sentential function for f(R, S, T) only yields a true or false sentence if the values of these types are inserted. If the three relations R_1, S_1, T_1 are admissible values of the variables R, S, T, i.e. if they can adopt the type of the three variables R, S, T (by declaring a specific domain of discourse in them to be the domain of individuals), we name the system of relations (R_1, S_1, T_1), or short M_1 an "*admissible model*" of fM, irrespective of whether M_1 satisfies fM or not; fM_1 is then in any case meaningful, i.e. either a true or a false sentence. An admissible model M_1 of fM is only then also a model of fM if fM_1 is not only meaningful, but also true. (Carnap 2000, 95, notation slightly changed)[36]

[35] Compare also Chapter 2 for a detailed survey of Carnap's conception of models.

[36] "Schreiben wir für f(R, S, T), kurz f(M), und sind R_1, S_1, T_1 bestimmte, etwa arithmetische Relationen, die zulässige Werte der Variablen R, S, T sind, so können wir auch für das geordnete Relationensystem R_1, S_1, T_1 ein abkürzende Bezeichnung einführen, etwa M_1. Das Modell M_1 ist dann ein Wert der Modellvariablen M. Jede der Relationsvariablen R, S, T des Axiomensystems hat einen bestimmten Typus (…); nur dann, wenn Werte dieser Typen eingesetzt werden, ergibt die Aussagefunktion für f(R, S, T) eine wahre oder falsche Aussage. Wenn die drei Relationen R_1, S_1, T_1 zulässige Werte der Variablen R, S, T sind, d.h. wenn sie (dadurch, dass ein bestimmter Gegenstandsbereich in Ihnen zum Individuenbereich erklärt wird) den Typus der drei Variablen R, S, T annehmen können, so nennen wir das Relationensystem (R_1, S_1, T_1), oder abgekürtzt M_1, ein „zulässiges Modell" von fM, gleichgültig, ob M_1 fM befriedigt oder nicht; fM_1 ist dann jedenfalls sinnvoll, nämlich entweder eine wahre oder eine falsche Aussage. Ein zulässiges Modell M_1 von fM ist nur dann auch Modell von fM_1, wenn fM_1 nicht nur sinnvoll, sondern wahr ist." (Carnap 2000, 95, notation slightly changed)

We have seen above that axiom systems were conceived by Carnap as "theory-schemata" logically presentable as systems of sentential functions that range over several predicate (or relation) variables. Systems with more than one predicate variable of the form f(R, S, T) are abbreviated by a "model variable" M of a specified type. The resulting axiom system f(M) allows different empirical and formal interpretations. The term "formal model" is reserved for the second type of interpretations. The variable M thus ranges over the set of adequate models ("zulässige Modelle") of f (ibid, 95).

Carnap's account has to be considered as an early attempt to provide a logically precise definition of models, viz. of condition (2) in the axiomatic context. Whereas different informal version of the notion had been used in the 1920s and before, it is mainly in Carnap's *Untersuchungen* as well as in Tarski's work at the same period that first attempts of a formal explication were given.[37] This said it is obvious that Carnap's definition of a model is - in several ways - heterodox compared to modern account of models. In recent work Carnap's *Untersuchungen*, it was characterized as "something like model theory" (Awodey and Carus (2001, 145) or as an "early heuristic form of the modern concept" of models (Bonk and Mosterin (2000, 38). However, so far no further qualification of Carnap's account has been given. Nonetheless, I believe that for reasons shown below, it is worth the effort to look more closely at where precisely Carnap's conception of models differs from the modern account. Some aspects are clearly anachronistic: first, note that in contrast to standard model theory models are devised for theories formulated not in first-order but in higher-order logic, particularly in type-theoretical setting.[38,39] Moreover, Carnap is operating in *Untersuchungen* with a pure logical language with an empty signature, whereas, in the modern account, models are usually constructed for theories in languages with a non-empty signature. This difference in the kind of the formal language is directly reflected in the Carnap's conceptual architecture of models. Briefly, a model (for a formal language) is conceived now as a tuple $\langle D, I\rangle$ where D designates the (non-empty) universe and I the interpretation function that

[37] See Badesa (2004) for an account of early model theory in the work of Löwenheim and Skolem. Compare Mancosu (2006) on Tarski's early notion of a model. See Sieg and Schlimm (2005) for a detailed discussion of the notion of a system in Dedekind's work on arithmetic, in particular Dedekind (1888). See Jané (2006) for a more general discussion of the semantic notions 'system of things', 'application', and 'representation' (among others) in the work of the Italian 'Peanists' as well as the American postulate theorists Veblen, Young and Huntington.

[38] In fact, Carnap gives an explicit specification of the type-theoretic setting in which his theory of models is conceptualized. One of his contributions to a systematic theory of models is the above distinction between the '*structural adequacy*' as a type-theoretic condition for 'admissible models' and the '*truth*' in a model. This two-step approach to models clearly mirrors Carnap's strong background in *Principia Mathematica*, particularly the theory of relations presented there. For more on this see Chapter 2.

[39] Compare, however, Shapiro (1991), where a model theory of SOL is outlined.

assigns to the non-logical symbols of the language elements of D (sets of elements of D, relations on D, and functions from D^n to D, respectively).[40] In Carnap's notion, there exists no interpretation function I, but only valuations to the variables R, S, T, viz. the primitive terms of the theory. A model M_1 is thus simply conceived as an n-tuple of relations (or more precisely, as we will see, as a sequence of relational constants) of a specified type that are assigned to the 'primitive symbols' and that "satisfy" the axiom system f.[41]

Finally, returning to our discussion above, it follows from Carnap's heterodox understanding of formal models and its subsequent use in his discussion of axiomatic theories and their metatheory that his account cannot be classified as an eliminative structuralism *tout court*. Just like the relative structuralist Hilbert, Carnap uses the notions of "truth", "satisfaction" of an axiom system *relative* to a given model and does not "abstract away" from particular interpretations of the theory. One can thus conclude that he held a hybrid structuralist position that combines both elements from (a) with elements from (b). We saw above that he seems to favour the semantics involved in unversalist structuralism in discussing the 'Explizitbegriffe' of an axiom system. This does not retain him from formulating an early theory of models and of using particular models in his study of axiomatic systems. We will return to the implications of Carnap's hybrid structuralist account in Section 1.5.

Given Carnap's presentation of formal models (and thus his version of (2)), we are now halfway to turn to the main question of this chapter, namely 'Does Carnap's formal reconstruction adequately capture the semantic innovation (3) of modern axiomatics?' Before addressing the question, however, a specific logicist assumptions implicit in his early semantics has to be specified. As we will see in Section 1.6, it is here, that Carnap's theory of axiomatics runs into problems.

1.4 A logicist reconstruction of axiomatics

We mentioned in the Introduction that Carnap's theory can be considered as a 'logicistic' reconstruction of Hilbertian axiomatics. This has to be refined. More than Frege, Carnap was aware of key methodological innovations (1) and (2) of modern axiomatics. His use of 'applied logicist' in the formalization of schematic primitives and models of a theory was clearly supposed to meet these desiderata. Moreover, Carnap also provided a reconstruction of

---

[40] See, e.g., Enderton (2001, 80-82)

[41] For a more detailed survey on these differences to the modern notion, particularly with respect to the concept of domain variation for models see Section 2.4.

the model-theoretic idea expressed in (3). Given this, on could infer that his theory clearly meets the adequacy conditions of formal axiomatics.

Such a diagnosis would miss to take into account, however, that Carnap's heterodox version of (2) was still closely tied to the traditional logicist program in the foundations of mathematics.[42] Interestingly, in Carnap's *Untersuchungen* and later on, this point is nowhere explicitly addressed. In fact, at one point in Carnap (2000), when introducing his "basic discipline" for his study of axiomatics Russell's project of the logical reduction of mathematics is mentioned in a sentence only to stress that his work on axiomatics is "independent of this conception" (ibid, 62). (Compare also Section 1.4.5 below.) This suggests that there was no closer conceptual connection between Carnap's work on axiomatics and his more general logicist leanings in the philosophy of mathematics at that time. However, as we will see, the situation is much more intricate than this.

As is well known, the logicist reduction of mathematics to logic depends crucially on the formal notion of interpretation. A mathematical theory T_1 (expressed in language L_1) is interpreted in a second logical theory T_2 (expressed in the logical language L_2 in the following way: (i) the mathematical terms of T_1 are defined in L_2, i.e. for every primitive or defined concept P of T_1 in language L_1 a formula φ is given in L_2 that defines the concept; (ii) the range of the quantifiers in L_2 is restricted to the range of the quantifiers of L_1 via quantifier relativization. Given (i) and (ii) all axioms and theorems of T_1 can be expressed as sentences of L_2 and can be deduced from the axioms of T_2. In the case of the classical logicist reduction, T_2 is logical axiomatic system like Russell's theory of types and L_2 a pure (type-theoretical) logical language.[43]

How does this classical logicist program relate to Carnap's logical reconstruction of formal axiomatics?[44] Evidently, in both cases the notion of an interpretation in terms of a translation plays a central role. In Carnap's axiomatics, it is, as we have seen, the translation of a formal, uninterpreted axiomatic theory into a fully interpreted logical system, his "basic discipline". The first step in this translation was outlined above: primitive terms are expressed as variables, the axioms and sentences of the theory as propositional functions of STT. In addition, there is a second step in the translation that has not been considered so far. What has to be added to the above picture is that also the semantics, in particular his formal models are

[42] For Frege's and Russell's logicist programs see Burgess (2005). For a broader account of Carnap's own evolving conceptions of logicism over his intellectual career see Bohnert (1975).
[43] For a more detailed definition see e.g. Burgess (2005, 50-51).
[44] Compare in particular Reck (2004, 172-175) on this point.

subject to interpretation in pure logic. Thus, a full logical reconstruction of an axiom system also concerns the interpretations of the system.

Carnap does not mention this additional condition concerning the interpretations of an axiom system, neither in *Untersuchungen* nor in *Abriss*. The most explicit account of it can be found in his contribution to the *Königsberg* discussion on the foundations of mathematics in 1930. Here, in aiming to provide a bridging account of the three foundational schools, he makes the following well-known remark on the "logical analysis of the formalistic system":

> 1. For every mathematical sign one or more *interpretations* are found, and in fact purely logical interpretations.
> 2. If the axiom system is consistent, then upon replacing each mathematical sign by its logical interpretation (or one of its various interpretations), every mathematical formula becomes a *tautology*.
> 3. If the axiom system is complete (…), then the interpretation is unique; every sign has exactly one interpretation, and with that the formalist construction is transformed into a logicist one. (Hahn and al. (1931, 143-144) quoted from: Awodey and Carus (2001, 153))

Carnap's understanding of a 'logicist' reconstruction of the formalistic method, i.e. formal axiomatics as the interpretation in pure logic is outlined most distinctively here. Point 1 highlights the above assumption, viz. the logical interpretability of the mathematical primitives of the axiom system. If the system is understood as formal then this condition effectively amounts to the logical interpretability of the formal models of an axiom system. Given this, we can refine our above question in the following way: what does Carnap understand by "purely logical interpretations" of an axiom system? And: applied to his theory of models, how does he conceive the logical definability of a model?

1.4.1 The definability of models

One of the earliest commentators of Carnap's *Untersuchungen*, the mathematician Hasso Härlen, already took notice of such a logicist assumption concerning the models of an axiom system. In commenting on a draft of the manuscript in a letter dated January 1, 1928, he remarked that:

The existence of cases of application [is here] equivalent to the existences of models. This claim is hypothetical; I would like to name it the hypothesis of logicism: that no structure can be described axiomatically if it is not already describable in logic. (RC 081-01-34)[45]

And then, in emphasizing the difference between Carnap's account and Hilbert's use of analytic models in Hilbert (1899) he adds:

Hilbert mapped geometry to a mathematical model, not to a formal one in your sense. Basically, this amounts to the same, but Hilbert's mapping differs not principally from those of Lobatschefsky, Klein, and Beltrami. (RC 081-01-34)[46]

The conceptual difference mentioned here between the "mathematical models" used by Hilbert and Carnap's "formal models" is important.[47] For Härlen, using the latter as presentations of the former implies an additional assumption, a "hypothesis of logicism," namely that every mathematical axiomatic structure can be "captured logically".

Now, it is important to note that the definability of models of an axiom system can be understood in two ways. In the first and weaker sense, the definability assumption simply requires that for every consistent mathematical axiom system, (at least) one purely logical model can be constructed. In the second, stronger sense, the definability assumption requires that all models of a formalized theory can be presented in logical form, i.e. as logical models. Both versions of the definability assumption have been discussed in the recent literature on Carnap's axiomatics, sometimes without sufficiently stressing their difference.[48] For instance, Reck (2007), in discussing *Untersuchungen*, explicitly argues for the stronger assumption:

---

[45] "Vorhandensein von Anwendungsfällen [is here] gleichbedeutend mit Vorhandensein von Modellen. Diese Behauptung ist hypothetisch; ich möchte sie die Hypothese des Logizismus nennen: dass keine Struktur axiomatisch erfassbar ist, die nicht schon logisch erfassbar ist." (RC 081-01-34)

[46] "Hilbert hat die Geometrie auf ein mathematisches Modell abgebildet, nicht auf ein formales in ihrem Sinn. Im Endeffekt ist das ja das gleiche, aber dem Sinn nach ist Hilberts Abbildg. nicht prinzipiell von denen Lobatschefskys, Kleins, Beltramis verschieden." (RC 081-01-34)

[47] Härlen's claim concerning the continuity between the 'intuitive models' introduced by Beltrami and Klein for non-Euclidian geometries and Hilbert's analytic models in Hilbert (1899) cannot be discussed here. See, however, Webb (1995).

[48] In recent commentary on the metatheoretic results in Carnap's 1928 manuscript, most notably Awodey and Carus (2001) and Reck (2007), this logicist definability assumption underlying his model theory is also mentioned, specifically in the discussion of Carnap's '*Gabelbarkeitssatz*'. The theorem states the equivalence of the three notions of completeness discussed by Carnap, i.e. categoricity ('*monomorphism*'), syntactic completeness ('*deducibility*') and semantic completeness ('*non-forkability*') (Carnap 2000, 133-139). Carnap's theorem is generally flawed for several reasons (see Awodey and Carus (2001) and Goldfarb (2004)). Nonetheless, it was argued that – given Carnap's definability assumption for models – the equivalence between semantic completeness and categoricity does in fact hold. Compare Awodey and Carus (2001) and Reck (2007).

"Carnap assumed that every model of a higher-order theory is definable." (ibid, 195) In discussing Carnap proof of the 'Gabelbarkeitssatz', he notes in Reck (2004) that:

> Carnap implicitly assumes that, for any model *M*, "being isomorphic to *M*" is *expressible in the simple theory of types*. This assumes that *any model M is definable in simple type theory*, which is not true as became clear after Carnap's work. (ibid, 170, my emphasis)

In contrast, in discussing the same proof, Awodey and Reck (2002) describe Carnap's tacit assumption in this way: "In particular, he in effect assumed that any consistent theory has a model that is definable within simple type theory, which is false." (ibid, 35)

We will return to the limiting effects of these definability constraints on models in Sections 1.5 and 1.6. For the moment, our interest lies in the question which of the two versions Carnap actually assumed. Not surprisingly, it is difficult to give a clear-cut response to this question based on Carnap's writings in the 1920s, given the fact that the logicist assumptions were (mostly) left implicit in his project. Nonetheless, one can get a clearer picture of his account when comparing it with two similar accounts of 'logicized' axiomatics of two of Carnap's contemporaries, Alfred Tarski and Friedrich Bachmann.

1.4.2 Tarski's "effective interpretation"

Carnap's "Gabelbarkeitssatz" from 1928 is usually discussed in reference with Tarski and Lindenbaum (1935) where the theorem was stated correctly for the first time. Here, in contrast to *Untersuchungen*, the additional assumption concerning the logical definability of models was made explicit.[49] The main result of the paper is the formulation of a theorem that connects the logical definability of concepts to an invariance condition under permutations of the logical "universal domain" V of a type-theoretic language. The theorem states that all notions definable in STT are invariant under all permutations of V and vice versa. The central point is that this invariance condition holds for "every relation between objects (individuals, classes, relations, etc.) which can be expressed by purely logical means", viz. in terms of STT (ibid, 385).[50]

---

[49] See fn.48. Compare in particular Bonk and Mosterin (2000, 41-43), Coffa (1991, 282), and Awodey and Carus (2001, 160). For an independent treatment of the paper compare Goméz-Torrente (2002).

[50] It should be noted here that a kind of invariance based approach to logically definable notions close to Tarski's account is already discussed in Carnap's conception of axiomatics. Compare a a passage on "structural properties" of relations in Carnap (2000) where this is made explicit: "The structural properties are in a way the

Applied to axiomatic theories this invariance theorem translates into a second theorem concerning the logical definability of mathematical relations. Tarski and Lindenbaum paraphrase it in this way:

Intuitively speaking, every relation between objects (…) which can be expressed in terms of logic and geometry is invariant with respect to every one-one mapping of space onto itself in which the relation *a* is preserved, i.e. with respect to every similarity transformation (and not only with respect to every isometrical transformation)” (ibid, 388)

Note that with the move to specific mathematic ‘deductive theories’ there is also a shift here from definability within a pure logical language (like STT) to definability within a formal language with a non-empty signature, i.e. with non-logical terminology (say STT*). For their example of an axiom system of Euclidian geometry, they explicitly stress the limits of logical definability compared to geometric definability (in this extended logical language): briefly, there are geometrical relations that are definable in STT*, but not in STT (ibid, 389).[51]

The consequence of this for Tarski’s treatment of his version of the ‘Gabelbarkeitssatz’ is this: the equivalence of semantic completeness (in Tarski’s terms “non-ramificability”) and categoricity is conceived as a direct “application of [Theorem] 1,” i.e. the first invariance theorem stated above. For this reason, an additional condition concerning the logical definability of the primitive terms of the axiomatic theory is explicitly mentioned (ibid, 390). Their crucial Theorem 10 states: “Every non-ramifiable axiom system which is effectively interpretable in logic is categorical.” (ibid, 391) The “effective interpretability in logic’ of a theory mentioned here is specified in this way:

[Theorem 10] (…) proves to hold under a supplementary assumption. We define an axiom system ‘$\alpha(a, b, c, \ldots)$’ to be *effectively interpretable in logic* if there is in logic a sentential function ‘$\tau(x, y, z, \ldots)$’ such that the following three logical formulas are logically provable:

$$(x, y, z, \ldots) : \tau(x, y, z, \ldots) . \supset . \alpha(x, y, z, \ldots);$$

$$(\exists x, y, z, \ldots) . \tau(x, y, z, \ldots);$$

$$(x', x'', y', y'', z', z''): \tau(x', y', z', \ldots). \tau(x'', y'', z', \ldots).$$

properties invariant under isomorphic transformations. They are of particular importance for axiomatics.” (Carnap 2000, 74)

[51] Tarski and Lindenbaum note for the case of geometry: “The distinction between logic and geometry comes to light, however, in the discussion of three termed relations; for, as is well known, there are infinitely many three- (and many-termed) relations which are definable in Euclidian geometry.” (ibid, 389)

$\supset . \ x' = x'' . y' = y'' . z' = z''$ ….. (ibid, 390-391)

Several points should be mentioned here. Note that Tarski's convention to formalize axiom systems in STT closely mirrors Carnap's own approach outlined above. A sentential function is logically expressed via the scheme '$\alpha(a, b, c, \ldots; x, y, z, \ldots)$'where a, b, c, …. stand for "extralogical" constants and x, y, z, …. for free variables of a given type. An interpreted axiom system can be expressed by the closed formula '$\alpha(a, b, c, \ldots)$' where a, b, c, … express the primitive terms of the theory. A formal, disinterpreted system is then symbolized as '$\tau(x, y, z, \ldots)$' i.e. a sentential (or propositional) function in pure STT.

This said, Tarski's additional condition concerning the "effective interpretability" of axiomatic theories is best understood as a translation of an interpreted theory of the form '$\sigma(a, b, c, \ldots)$' expressed in STT* (i.e. here STT with the signature $\{a, b, c, \ldots\}$) into a theory of the form '$\tau(x, y, z, \ldots)$' expressed in pure STT. This method bears strong similarities with the method of logicist reduction outlined above. In Tarski's case, the interpretation can be effected in two ways: in the above passage the three conditions fix (i) that there is an universally quantified implication from the sentential function τ (expressing the logical axiom system) to the sentential function α (expressing the disinterpreted axiom system); (ii) that there exists an interpretation of for τ and (iii) that τ holds for a unique sequence of values for x, y, z, …. The second way outlined in Tarski and Lindenbaum (1935) to characterize the "effective interpretation" has an interesting implication on level of models of an axiomatic theory.[52] There is an insightful reformulation of the conditions (i) to (iii) of the above passage:

(…) the axiom system '$\alpha(a, b, c, \ldots)$' is effectively interpretable in logic if and only if there are logical constants 'a'', 'b'', 'c'', … (undefined or defined) such that the sentence '$\alpha(a', b', c', \ldots)$' obtained by substituting these logical constants for the primitive terms in the axiom system discussed is logically provable." (ibid, 391)

In this kind of translation, the primitives of an interpreted theory expressed by the non-logical constants a, b, c, … are substituted by logical constants a', b', c', ..., i.e. primitive or defined terms of STT. The axiom system is thus effectively interpretable in STT if the transformed

[52] For a closer comparison of Tarski's and Carnap's conception of formal models see 2.4.2.

system ‘$\alpha(a', b', c', ...,)$’ can be deduced in Tarski’s logical system (i.e. a theory of types plus the axiom of infinity).[53]

Tarski conception of models in the early 1930s is similar to that of Carnap. Models are in both cases understood as tuples of relations that are substituted for the variables in the formalized axiom system ‘α(x, y, z, …)’. Given this, the sequence of logical relations $\langle a', b', c', ...\rangle$ has to be understood as a logical model of the theory. Thus, “effective interpretability” means the construction of a particular model for an axiom system that consists of logical constants. This is in effect the translation of a theory in a mathematical language into a theory in pure HOL where variables expressing primitive notions are substituted by logical notions. Conceived in this way, the “effective interpretability” of an axiom system then simply states that there is (at least) one model of the theory that is definable in pure STT.

1.4.3 Bachmann’s “logical constitution of a model”

The notion of an “effective interpretation” outlined in Tarski and Lindenbaum (1935) is not the only specification of a logical definability assumption comparable to the “hypothesis of logicism” in Carnap’s formal models. A similar account can be found in a relatively unknown work of the mathematician Bachmann, a later collaborator of Carnap (Bachmann 1934).[54] Bachmann’s subject of his inaugural dissertation, the formal presentation of different axiom systems of arithmetic and the study of their logical relations shares various points of contact with both accounts. As in Tarski’s notion here too the aim is explicitly to provide a logicist presentation of axiomatic theories.[55] In the introduction, Bachmann presents this task in this way:

[53] Tarski and Lindenbaum mention a possible limit to the effective interpretability of a theory in a given type level of the background language. They add: “It sometimes happens that an effective interpretation of an axiom system in logic can be found only at a higher type (this applies, for example to the axiom system of the arithmetic of real numbers).” (ibid, 391) Compare 2.4.3 for a similar account of *type ascent* in Carnap’s logical presentation of models for axiom systems.

[54] Compare the discussion of Bachmann (1934) in 2.4.2.

[55] In his introduction, he explicitly speaks of an “allgemeine Untersuchungen zur Axiomatik der Arithmetik” (ibid, Introduction). There is no indication, despite the strong similarity to Carnap’s *Untersuchungen* that Bachmann knew Carnap’s 1928 typescript when working on his dissertation. Nonetheless, in Bachmann’s bibliography, explicit reference is made to Carnap (1929). Moreover, there exists an extensive discussion of the treatment of extremal axioms in Carnap (1930, 46-47). This topic was the subject of the later collaboration between the two. Compare also Chapter 2.

The subject of this work is the problem of the deducibility of arithmetic from logic or – to be more precise – the problem of the deducibility of an interpretation of arithmetic from logic. (ibid, Introduction)[56]

Bachmann mentions two "subtasks" for the "solution of the problem":

1) The logical characterization of the models of arithmetic (the presentation of an axiom system of arithmetic)
2) The logical constitution of a model of arithmetic (the presentation of a system of constants and the proof that this system of constants holds in an axiom system of arithmetic. (ibid, Introduction)[57]

This is clearly not the typical approach of reducing arithmetic to logic according to classical logicism. Nevertheless, Bachmann's close ties to this program are evident, however, in the specifics of these two tasks. He discusses (1) for different axiomatizations of arithmetic, including Peano arithmetic, Dedekind's version of arithmetic via chains, and an axiom system of elementary arithmetic (ibid, 27-52).[58] As in Tarski's and Carnap's versions, axiom systems are treated here as "sentential forms" ("*Aussageformen*") and the un-interpreted primitive terms as variables.[59] Models are also conceived in essentially the same way as in Carnap (2000). In Bachmann's terms, a model for the system of elementary arithmetic f(x R α) is specified in this way: "Every triple of values that if substituted for x R α transforms the axioms in true sentences is called a model of arithmetic." (ibid, 1)
Given these striking similarities, two subtle differences to Carnap's account are worth to be mentioned. The first concerns Bachmann's discussion of his first subtask of the "logical constitution of a model" for axiomatic arithmetic, i.e. the "presentation of a system of constants" as a model for the system. Whereas Carnap's account is, as we will see, restricted to a presentation of different sequences of "constants of the basic system" without any further specification of these constants, Bachmann is clear about the admissible range of constants to be allowed here. He states that the constants constituting a model have to name concepts that

[56] "Der Gegenstand dieser Arbeit ist das Problem der Ableitbarkeit der Arithmetik aus der Logik oder – wie ich mich genauer ausdrücken möchte – das Problem der Ableitbarkeit einer Interpretation der Arithmetik aus der Logik." (ibid, Introduction)
[57] "1) einer logischen Charakterisierung der Modelle der Arithmetik (Angabe eines Axiomensystems der Arithmetik); 2) der logischen Konstituierung eines Modells der Arithmetik (Angabe eines Konstantensystems und Beweis, dass dieses Konstantensystem einem Axiomsystem der Arithmetik genügt)." (ibid, Introduction)
[58] Compare Chapter 3 for a closer discussion of these systems.
[59] For a given theory of arithmetic, he states that: "Our axiom system consists of four sentential functions in which apart from the constants of logic and bound variables also three free variables x R α appear." (ibid, 1)

are definable in sense of Frege's reductive definitions. This is spelled out explicitly as a condition for his task of showing the "deducibility of an interpretation" of arithmetic. What is missing - and probably deliberatively left out in Carnap's treatment due to his shift of attention from logicist reduction to axiomatics – is the specific method of the "presentation" ("*Aufweis*") of "a logically constituted Modell" ("ein logisch konstitutiertes Modell") (ibid, 9). In Bachmann (1934), for the case of elementary arithmetic, such a task consists in the presentation of a sequence ⟨0, *Succ*, *NN*⟩, where each concept is given an explicit reductive definition. Here, unlike in Carnap, the assumption concerning the logical definability of models is made explicit. So, the first task of step (2) of a logicized reconstruction of axiomatic arithmetic is to present at least one purely logical model just as in Tarski and Lindenbaum (1935).

A second difference to Carnap's account concerns Bachmann's second task of step (2), the "proof that this system of constants holds in an axiom system of arithmetic" (ibid, Introduction). Following the presentation of a logical model (as a sequence of logically defined constants) for the axiom system in question, Bachmann holds that, in a second step, one has to "then show of its definientia that they form a model of arithmetic" (ibid, 9). This is the proof that the logically defined notions '0', '*Succ*', and '*NN*' "satisfy" the conditions specified in the axiom system (see ibid, 18-26). Note that in Carnap's account, the relevant notions used for this like 'truth in a model' or 'satisfaction' are left informal, at least in 1928.[60] Nevertheless, it is clear that they are - at least in principle - considered to be semantic in nature. In Bachmann's case, the proof that the models "hold" in an axiom system is understood in purely syntactic terms. We have seen that for him models are a set of logically defined relational expressions. An interpretation then holds in an axiom system if the sentence resulting form the substitution of the variables in f(x R α) by such constants is provable in the logical background system. This "deducibility of an interpretation of arithmetic" from logic is not understood model-theoretically. The proof that a model of the form ⟨0, *Succ*, *NN*⟩ "satisfies" the axiom system is given "by exclusive use of the means of proof of logic" (ibid, 8). Note that also in Tarski and Lindenbaum (1935), the adequacy of a effective interpretation of an axiom system is not expressed via a semantic notions of truth but syntactically, via the "logical provability" in STT (see Coffa 1991, 282-283) This clearly underlines his logicist motivation of both accounts.

[60] Compare 2.4.1 for a more detailed discussion.

Given Bachmann’s two-step approach, it is important to note that the “logical constitution” of a model of arithmetic means for him, just as in Tarski’s case, the construction of a single logical model. Here again, definability seems to be understood in the weak form.

Nevertheless, there exists a crucial difference between Tarski and Bachmann concerning the general motivation of the formal reconstruction of axiomatics. In Tarski and Lindenbaum (1935), the “effective interpretation” is conceived as one among alternative ways to present logically mathematical theories and their models. Given the convention that axiom systems are expressed via propositional functions with several variables, models can here either be presented as sequences of genuinely mathematical and thus non-logical relations of a language STT*. They can also be presented as sequences of logical expressions, “defined or undefined” in pure STT.[61]

In contrast, in Bachmann (1934) the *exclusive* aim of the study is to devise a logic reconstruction of axiomatics close to classical *Fregean* reduction. It is in this understanding of logical formalization that the stronger definability assumption concerning models can be identified. Consider again Bachmann’s second task for his intended “deducibility of an interpretation of arithmetic from logic:” to present a single “logically constituted model“ that holds in the theory. Given his logicist account, any model of elementary arithmetic that can be subject to this second task has to be logical in the sense specified by him. In contrast to Tarski’s account, there are thus no alternative ways to formalize axiom systems and their models. The resulting stronger version of the definability assumption is explicitly mentioned in Bachmann (1934). In his discussion of the “proof that the system of constants holds for an axiom system,” he gives an explanation why the models must consist only of logical constants in this context:

> For the second task to succeed it is necessary that the three constants are constructed from the primitive signs of logic and bound variables, in short: that they are logical constants. A sentence with nonlogical constants is only deducible from logic if it is transformed into a universally valid sentential form in case the constants are substituted by variables. Since the sentential form A-AS I(x, *R*, α) is not universally valid, this approach is not possible. (ibid, 9)[62]

---

[61] See Goméz-Torrente (2009) for a broader survey of Tarski’s “logical pluralism”, i.e. the different uses of logical language for the formalization of mathematical theories in the 1930s.

[62] “Damit die Lösung der zweiten Aufgabe gelingt, ist es notwendig, dass die drei Konstanten aus den Grundzeichen der Logik und gebundenen Variablen aufgebaut, kurz: dass sie logische Konstanten sind. Eine Aussage über ausserlogische Konstanten ist nämlich nur dann aus der Logik ableitbar, wenn sie bei Ersetzung der Konstanten durch Variable in eine allgemeingültige Aussageform übergeht. Da nun die Aussageform A-AS I(x, *R*, α) nicht allgemeingültig ist, ist dieser Weg nicht gangbar.“ (ibid, 9)

This passage is central for Bachmann's understanding of the logicist reconstruction of axiomatics. The theory A-AS I(x, *R*, α) is not universally valid since one can, as Bachmann points out, construct "non-models" that fail to satisfy the theory (ibid, 5-6). It follows from this that any model constructed for the theory has to consist exclusively of logical constants in order to allow this proof of the validity the interpreted system. Thus, in Bachmann's specific presentation of axiomatics, any model for an axiom system has to be logical for exactly this reason. All models relevant for consideration are thus "logically constituted" in this sense.

1.4.4 Carnap's substitutional conception of models

How do Tarski's "effective interpretation" of an axiomatic theory and Bachmann's "logical constitution of a model" relate to Carnap's tacit definability assumption? Recall again the latter's remarks on the "logical analysis of the formalistic system" in Königsberg. Here his intention was obviously to provide a *logicized* version of formal axiomatics close in spirit to both Tarski's and Bachmann's accounts. Moreover, his talk of "one or more (…) in fact purely logical interpretations" of the mathematical primitives suggests that he too proposed the weak version of logical definability in 1930.[63]

However, we saw from Bachmann (1934) that this type of reconstruction if considered the exclusive formal method to treat axiomatic theories effectively implies the stronger version, namely that all relevant models are logically definable. Turning to *Untersuchungen*, we noted that Carnap's primary aim here was not to provide logicist reconstruction in Bachmann's sense. Carnap stresses at the outset of his manuscript that the principle objective is to provide formal explications, i.e. "*precise definitions for the concepts*" used in axiomatics in a formal framework and not a *Fregean* reduction (Carnap (2000, 59)). In this sense, Carnap's account in 1928 is closer in motivation to Tarski's discussion of deductive theories in 1935. Nonetheless, Carnap's theory differs from the latter's more pluralist account with respect to the specific formal framework. In particular, Tarski conceived of formal languages with non-logical terminology. One of his conventions of a formal reconstruction of the semantics of a theory was the introduction of (a sequence of) non-logical mathematical constants $\langle a, b, c, \ldots \rangle$ that are assigned to the variables of an axiom system '$\alpha(x, y, z, \ldots)$'. This is ruled out in

[63] Note that a similar account of a "logical interpretation" (in contrast to a "descriptive interpretation") of an axiomatic theory in a type-theoretic theory can still be found in Carnap (1939). The idea expressed here is essentially identical to Tarski's method of an "effective interpretation" (compare ibid, 22).

Carnap's framework with the deliberate limitation to a pure type-theoretic logic. In consequence, all formal models for an axiom system have to be expressed in pure STT. So, whereas his general approach in 1928 is closer to Tarski's formal reconstruction, he effectively assumed logicist premises similar to those highlighted in Bachmann (1934). The main tacit import of this is that he in fact assumed the stronger version of the logical definability of models. The fundamental difference to Tarski's account of formal reconstruction of axiomatics is that the latter does *not* impose – as does Carnap – that all possible interpretations of an axiom system are logical. For the latter, the "effective interpretability" of an axiom system is a special case of model construction. For Carnap, a similar interpretability is conceived as a general condition for providing an interpretation. Here, the underlying assumption is that all models are in fact expressible in his "basic discipline".[64]

This tacit condition of the *interpretability* in STT is evident in Carnap's substitutional conception of models. We have seen that models were treated as *n*-tuples of relations, more specifically as relational constants in *Untersuchungen*, i.e. as relational terms with a fixed meaning of the interpreted background language. The assignments to the "primitive variables" are not understood extensionally in terms of objects, sets, or relations but substitutionally, in terms of constants of STT. Provided that STT is understood here as a pure logical language without non-logical terminology, these constant expressions are all logical (as in Bachmann's theory).

There is strong evidence that Carnap understood his formal models in this way, in 1928 and also in later works. Compare a passage in *Untersuchungen* where he further specifies his account of models:

---

[64] It is important to clearly distinguish between the two assumptions – weak and strong definability – here. The versions are sometimes mingled in the recent debate. Compare a passage in Reck (2007) on the relation between Carnap's tacit assumption and Tarski's "effective interpretation": "Carnap assumed that every model of a higher-order theory is definable. This is made explicit in the correct, but restricted, version of the result published, a few years later, in Lindenbaum and Tarski (1935)." (ibid, 195) What Tarski effectively imposes here, is, as we have seen, the weaker constraint on model definability. Note, however, that the stronger version clearly implies the weaker one. In a perceptive footnote in Awodey and Reck (2001), this point is stressed: "Note that this restriction, that the axiom system be "effectively interpretable" (…), is in effect imposed by Carnap in the Axiomatics as well, where all sentences are interpretable in the basic system." (ibid, 169)

(…) in short, we speak of “models” of an axiom system und thereby mean logical constants, i.e. “*systems of concepts of the basic discipline* (these are mostly systems of numbers). (ibid, 94)[65]

This is further highlighted in his distinction between “formal models” and “realizations”, i.e. empirical interpretations of an axiom system:

As the values of the primitive relations of an axiom system both logical as well as non-logical constants can occur; the axiom system can either be applied to concepts of the basic discipline or also to real concepts (“*Realbegriffe”*) (concepts of a nonlogical, empirical state of affairs). (ibid, 93)[66]

In this conception, mathematical or formal models are exclusively composed of the logical terms of the “basic discipline”. The same account can also be identified in a later paper, Carnap and Bachmann (1936). In specifying the notion of an isomorphism correlation between “*n*-place models (i.e., sequences with *n* members)” the authors hold here that such a correlation is defined “(…) over the field of this relation, i.e., over the totality of constants of the basic, assumed language which can appear as elements of models (...)“ (Carnap and Bachmann (1981, 74). Here again, the underlying language is a pure and higher-order. This substitutional conception of models is still present in the first English edition of *Logical Syntax of Language* from 1937. A small section not yet included in the original German version is devoted to the “axiomatic method“ (ibid, §71e). Here, one can find a definition of models and model domains that is comparable to the one in Carnap and Bachmann (1936) and Carnap (2000):

In the first method, the domain of the interpretations of a certain primitive symbol is the domain of the substitution-values of the variable. If, as is usual, it is a case of primitive variable within a system of types, then the same relations of types must hold between the symbols of the model as hold between the corresponding primitive variables. (Carnap 2002, §71e)

According to this method, a model for the AS is to be understood as a series of *k* substitution-values $U_1, \ldots U_k$ for the primitive variables. (ibid, §71e)

---

[65] “(…) wir sprechen kurz von “Modellen” eines Axiomensystems und meinen damit logische Konstanten, also „*Systeme von Begriffen der Grunddisziplin*“ (und zwar sind dies meist Systeme von Zahlen).“ (ibid, 94)
[66] “Als Werte einer Grundrelation eines Axiomensystems können sowohl logische als auch nicht-logische Konstanten auftreten; das Axiomensystem kann angewendet werden auf Begriffe der Grunddisziplin und auch auf Realbegriffe (Begriffe eines nichtlogischen, empirischen Sachverhalts.)“ (ibid, 93)

Each of these passages cited here describes the same picture: formal models as conceived as sequences of constants of a fully interpreted type-theoretic language with an empty signature. These constants are assigned to the "primitive variables" of a theory also expressed in that language. Since there is no non-logical terminology in languages like STT, all relevant constants used for this kind of model construction have to be logical or logically definable in STT.

1.4.5 The "scope of logic" in *Untersuchungen*

Carnap's substitutional conception of models imposes strong conditions on the scope of logical expressions of his "basic discipline". The expressibility condition effectively presupposes that there is a range of logical constants in the language sufficient to express all classes, relations, etc. that are referred to in models. Several questions arise from this: First, what exactly does Carnap understand by logical constants in *Untersuchungen*? Second, what is the "scope of logic" ("*Gebiet der Logik*") in his "basic discipline"? Finally, what does logical definability mean in context of his theory? Before turning to a discussion of the limiting effects of this logicist assumption for his overall theory of axiomatics, these points have to be clarified.

Typically, constant expressions can be added in two ways to a formal language. They can be stipulated at the outset with the specification of the primitive symbols of the language. They can also be introduced via definitions from the given vocabulary. For example, if a certain formula of the language specifies a certain class of (or relation on) the language's interpretation, a new constant can be added to the language to name this set (or relation).
A similar approach can be found in Carnap's specification of his interpreted "basic discipline" in *Untersuchungen*. The main difference to a modern conception lies in the fact that here the constant expressions are all logical. Thus, the language does not contain non-logical expressions (i.e. "external signs" ("*Fremdzeichen*") denoting possible "external concepts" of an axiom system) (see Carnap (2000, 89-90).[67] In setting up his background logical theory, Carnap gives the following classification of the possible logical terms:
(1) proper *logical* constants ($\forall$, $\exists$, $\rightarrow$, $\neg$, …);

[67] "The present investigation discusses, as is usual in mathematics, only axiom systems without external concepts." (ibid, 90)

(2) constants of absolute arithmetic (*successor, number*, 0, 1, 2, …);

(3) constants of absolute set theory (*class*, $\in$);

(4) constants of relation theory (Carnap 2000, 59-70).

The relevant distinction made here is between the logical constants in (1) and those logical constants in a broader sense (2) to (4).[68] Combined, they constitute the syntactic resources of his ”basic discipline” in which the models for a theory are constructed. Two points should be noted here. First, the notion of logical concepts is obviously conceived here in a much more general way compared to the modern account. STT does not only include the proper logical terminology of (1) but in addition terms from arithmetic, set theory, and most importantly, relation theory. Compare Carnap's following remark on this generalized account of logic:

> The schools diverging in the question of providing a foundation to mathematics will come to agree that in the basic discipline the usual arithmetic, set-theoretical and also logical concepts have to be present; we want to designate the concepts of these fields (…) with the unifying expression "*logical concepts*". (ibid, 62)[69]

Thus, any formalization of an axiomatic theory in terms of the theory of types does in fact draw to resources not included in a modern account of HOL. These are in particular terms of mathematics and constant expressions of relation theory.

Given Carnap's generalized "scope of logic", the question arises how the expressions in classes (2) to (4) are related to the proper logic in (1).[70] Carnap in *Untersuchungen* is more or less silent on how the terms of (2) to (4) are supposed to enter into the "basic discipline". He mentions that these concepts can be logically constructed from proper logical concepts by reductive definitions as in the work of Frege and Russell (see ibid, 61-62). Nonetheless, he leaves undecided whether they should be conceived with such a "logistic meaning". In fact, in a side remark, Carnap is quite explicit that he is not interested in the logicist project in *Untersuchungen*:

---

[68] Compare Carnap: "By "logical constants" we understand all signs of the basic discipline, i.e. the signs of logic in the closer sense (...) as well as the arithmetical signs (e.g.: 1, 2, + etc.)." (ibid, 89) At other places of *Untersuchungen*, also the signs of set theory and relation theory are mentioned.

[69] "Die in der Frage der Begründung der Mathematik auseinandergehenden Richtungen werden darüber einig sein, dass in der Grunddiziplin die üblichen arithmetischen, die üblichen mengentheoretischen und außerdem noch logische Begriffe vorkommen müssen; die Begriffe dieser Gebiete wollen wir (…) mit der zusammenfassenden Bezeichnung "*logische Begriffe*" nennen." (ibid, 62)

[70] Note again that all notions of the "basic discipline" are (in contrast to their disinterpreted 'axiomatic' correlates) understood as "absolute", i.e. as fully interpreted ("*inhaltlich*") (ibid, 60).

For the following investigation of axiomatic problems it is irrelevant whether the set- theoretical and arithmetical concepts are presented as independent primitive concepts or if they are deduced from those [concepts] of logic. The results are independent of the mentioned (Russellian) conception. (ibid, 62)[71]

Thus, he deliberately leaves open the question whether the primitive mathematical notions like '0', '1', '2', '+', '<', '>', '∈', etc. should be conceived as undefined primitive logical constants or as logical constants defined from the core logical vocabulary in (1). This marks a crucial difference to the decidedly logicistic account in Bachmann (1934). For Bachmann, the relevant logical constants used in a model of axiomatic arithmetic are particularly those mathematical terms that can be given a logicist definition in pure logic (compare Bachmann (1934, 8-9). In *Untersuchungen*, these constituents of the formal models are simply treated as logical primitives. It is deliberately left open whether they are defined or undefined notions.
However, this alleged indifference towards this question in Carnap (2000) is doubtful at closer inspection. The generalized account of logical concepts is almost per definition conceptually tied to a logicist and thus reductive approach. The whole idea of classifying genuinely mathematical terms like '*number*', '*successor*', '*point*', '*betweenness*' etc. *as logical notions* as Carnap is grounded on the idea that they can be defined in proper logic. That the genuine mathematical concepts in the scope of logic are to be definable in terms of pure logic is explicitly stated in a later work by Carnap (Carnap 1948). Here the generalized scope of logic from 1928 is still upheld. He gives the following explication of the expanded class of logical signs:

Further, all those signs are regarded as logical which are definable by those mentioned [GS: those of the above class (1)]; (…), all signs of the system of [*Principia Mathematica*] by Whitehead and Russell and of nearly all other systems of symbolic logic, all signs of mathematics (including arithmetic, analysis of real numbers, infinitesimal calculus, but not geometry) with the meaning they have when applied in science (…). A defined sign is descriptive if its definiens contains a descriptive sign, otherwise logical. (ibid, 57-58)

[71] "Für die folgenden Untersuchungen axiomatischer Probleme ist es indessen nicht von Belang, ob die mengentheoretischen und die arithmetischen Begriffe als selbstständige Grundbegriffe aufgestellt sind oder aus denen der Logik abgeleitet werden. Die Ergebnisse sind also unabhängig von der genannten (Russellschen) Auffassung." (ibid, 62)

It is precisely this condition of logical concepts that also seems to underlie Carnap's conception of the "basic discipline" in *Untersuchungen*: according to it, a sign is logical if it is primitively (or genuinely) logically or can be defined in terms of genuinely logical expressions.[72] Given this conception, it is highly likely that Carnap - irrespective of the remark cited above - understood the mathematical terms in (2) to (4) of his 'basic discipline' not as primitives but as logically defined terms in the sense of Russell's and Frege's reductive definitions in the part (1) of his logic.[73]

This point concerning the definitional structure of the "basic discipline" can be further strengthened by looking at how the non-elementary, mathematical constants are actually used in the specification of models. We have noted in Section 1.3 that in contrast to Frege, Carnap fully embraced the idea of model variation in the theory of axiomatics. Thus, for a given axiomatic system like Peano arithmetic or Hilbert's Euclidian geometry he accepted the idea that different models can be constructed for each theory. Given his formal presentation of models, this implies that tuples of extensionally different relations (more precisely, relational constants) can be presented in which the system is interpreted. Therefore, there exist – at least in the case of non-categorical theories - not only one sequence of logical concepts but several sequences of extensionally different concepts that satisfies the theory. Carnap's "basic discipline" has to include different logical signs of, say 'successor' in the case of arithmetic or 'between' in the case of Euclidian geometry that denote extensionally different relations of individual elements. Together, they form the substitution class of the "primitive variables" of an axioms system discussed in Section 1.3.3. Now, it is in principle possible in this "relation-theoretic" presentation of models that these different expressions are introduced as undefined terms into the language. It is far more plausible that they are introduced as constants denoting sets or relations that are explicitly defined in terms of (1) of STT.

[72] A comparable conception of the scope of logic can be identified in the work from the 1930s of Carnap's contemporary Tarski. Compare Goméz-Torrente (1996) on his conception of logical notions: "(…) Tarski reserves his most inclusive use of the word 'logic' for a system of logic based on the theory of types (…), such a "logic" is a system, therefore, in which arithmetical constants can be defined in terms of logical constants, and arithmetic developed as logic." (ibid, 134)

[73] Goméz-Torrente (2002) gives as more general presentation of this practice in Tarski and Lindenbaum (1935) that is also shared by Carnap: "(…) all primitive symbols denoting notions in that language (…) are (logical constants)^{T}; also, if the definition were applicable to defined symbols, all these symbols would be (logical constants)^{T}. Such results agree well with (…) traditional practice (for example the practice of the logicists, but others as well) according to which the constants of the language of the theory of types are logical constants.“ (ibid, 16)

Irrespective of this question, - given Carnap's substitutional conception of models and his generalized scope of logic - we are now in a position to give a more refined definition of his tacit definability assumption (henceforth DA) in *Untersuchungen*:

(DA) All models of a given axiom system are expressible in the 'basic discipline" in the sense that all constitutive parts (sets, relations, and individual constants) of a model are named by constant expressions of STT. These constants are (i) either conventionally fixed to be part of the basic vocabulary or are (ii) introduced via explicative definitions in pure STT.

1.5 Carnap's tacit semantics for STT

Carnap's implicit logicist assumption (DA) has strong, limiting consequences for his overall theory of axiomatics. It effectively restricts the class of possible models of a theory to those expressible in his background logic. This is not necessarily restricting in the case of individual model construction. In this case, (DA) calls for the presentation of a particular *logical* model in the sense of Tarski's "effective interpretation" or Bachmann's "presentation of a logically constituted model."[74] Here, the condition of logical definability causes no deeper problems for Carnap's theory, at least for the axiomatic systems discussed in *Untersuchungen*. For instance, as it was shown by Bachmann (1934) for the case of Peano arithmetic that different models can be presented via explicit 'logicist' definitions of the three primitive terms '0', '*Succ*' and '*NN*'.[75]

Where (DA) imposes constraints on Carnap's theory of models is in cases where he refers to *all models* of a theory. Here, the condition effectively implies the expressibility of *all* models in STT, either in terms of primitive expressions of the language or in terms of defined notions. The real impact of (DA) is thus not on the level of concrete model construction ("*Modellaufweis*") but on the more general level when the class of models of an axiom system is discussed. These instances are formally expressed by Carnap via the quantification over higher-order relations in STT. It is here that talk about the totality of models also implies

---

[74] Note that a similar account of a "logical interpretation" of an axiomatic theory in a type-theoretic theory can still be found in Carnap (1939). The idea expressed here is essentially identical to Tarski's method of an "effective interpretation" (compare ibid, 22).

[75] One possible limitation to the logical presentation of models in a logical system was mentioned in Tarski and Lindenbaum (1935), namely that the definability depends on the specific richness of the assumed background language. They mention the following example: "It sometimes happens that an effective interpretation of an axiom system in logic can be found only at a higher type (this applies, for example to the axiom system of the arithmetic of real numbers)." (ibid, 391) A similar point was also stressed in Carnap and Bachmann (1936). Compare also 2.4.3.

effective constraints on the tacit semantics of his type-theoretic language, particularly of its higher-order segments.

1.5.1 Quantification over models

Quantification over the models of a theory is usually expressed in the metatheory today, e.g. in an interpreted set theory like ZFC. In Carnap's early formal account, in contrast, talk of "all models" was cast in his 'universalist' and also fully interpreted type-theoretic language STT. Recall that models were conceived in Carnap (2000) and later on as *n*-tuples of relational constants of a given type of the form $M_1 = \langle R_1, S_1, T_1, \ldots\rangle$. Given this conception, the generalization over models is then simply expressed as the higher-order quantification over relations (of the type-level of M_1). So, e.g., in the case of Peano arithmetic with an individual term and two first-level primitive terms, the quantifiers binding the "model variable" would range over second-level tertiary relations of STT.

This use of higher-order quantification is frequently used in Carnap's work on axiomatics. Four particular cases should be mentioned here:

(1) The first example concerns his treatment of the notion of logical consequence. In Untersuchungen (and already in Carnap (1927)) a "*Lehrsatz*" is expressed by $(R)(\mathrm{f}R \rightarrow \mathrm{g}R)$, where R is an unspecified predicate (or relation) variable, fR a propositional function expressing the axiom system, and gR a propositional function expressing a sentence "implied" by fR. "*Lehrsätze*" are defined as follows: „gR is called a consequence of fR, if $(R)(\mathrm{f}R \rightarrow \mathrm{g}R)$ is valid" (Carnap 2000, 92) Carnap gives this further specification: "If $\mathrm{f} \rightarrow \mathrm{g}$ is valid, that is, if all models of f are also models of g, we can also say: the extension ("*Umfang*") of f is a "part" of the extension of g" (ibid, 95). Clearly, R is understood here as a model variable that ranges over all possible models of the system f.[76]

(2) A second example where Carnap generalizes over models can be identified in his treatment of "extremal axioms" in the projected second, unpublished part of *Untersuchungen* (and later in Carnap and Bachmann (1936)). "Extremal axioms" are introduced to a theory in order to impose minimal or maximal conditions on the possible interpretations of the theory.

[76] Compare Reck (2007) for an extensive discussion on Carnap's understanding of the consequence-schema $(\forall R)(\mathrm{f}R \rightarrow \mathrm{g}R)$. Reck points out that there is "an ambiguity in his definition of the notion of deducibility, or of logical consequence more generally" that concerns Carnap's assumed conditions under which a "Lehrsatz" is valid (see ibid, 188-189).

His main example of a mathematical maximal axiom is Hilbert's *axiom of completeness* in geometry (see Hilbert 1900).[77] Carnap's formal reconstruction of this maximal axiom has the following form:

$$\mathrm{Max}(F;\, M) =_{\mathrm{df}} \sim(\exists N)\ (M \subset N \,.\, M \neq N \,.\, F(N)).$$

Informally, the axiom states that for a given theory F and a model M of F, there exists no proper extension of M that also satisfies F (Carnap and Bachmann 1981, 77).

(3) A further example of quantification over models, again of a clearly metatheoretic character from a modern perspective, is Carnap's treatment of categoricity (in his terms "*monomorphism*") in *Untersuchungen*. Based on a formal presentation of the notion of "model isomorphism" between models of a given type structure (Carnap 2000, 104-122), a theory f is categorical if $(\exists)\mathrm{f}\ \&\ (P, Q)\ [(\mathrm{f}P\ \&\ \mathrm{f}Q) \rightarrow \mathrm{Ism}_{\mathrm{q}}(P, Q)]$ holds in STT. The formula expresses that the theory f is satisfied and that for any two models P and Q of f there exists an "isomorphism correlation" of type-level q between P and Q (ibid, 128-129).

(4) Finally, we have already seen that an implicit generalization over the possible models of a theory can be identified in Carnap's notion of the "Explizitbegriff" (compare Section 1.3.1). Recall that, for Carnap, an axiom system AS defines in an explicit fashion the class of all models of AS in terms of this higher-order explicit concept EB. We noted that his formal presentation of this via class comprehension, i.e.

$$\mathrm{EB}_{\mathrm{AS}} = \hat{R}, \hat{S}, \hat{T}\{AS(R, S; T)\}$$

can be reformulated as an universally quantified sentence of STT of the form

$$(\forall R)(\forall S)(\forall T)(AS\ (R, S, T) \rightarrow \mathrm{EB}_{AS}(R, S, T))$$

Thus, Carnap's specification of the class of models of a theory in an "Explizitbegriffe" can also be recast in terms of higher-order quantification over relations.

Given these examples, two brief remarks should be made. First, note that all cases can be viewed as additional cues for Carnap's universal structuralism underlying in his theory of axiomatics (in the sense specified in Section 1.3.2). Quantification over models can in a sense be considered as the abstraction from specific instances of his logical model construction described above. Thus, in results that would today run under the label of metatheory - viz. the formulation of categoricity, model maximality, logical consequence, and the definition of

[77] See Sections 2.2 and 3.2 for further details.

classes of structures - Carnap seems to have held a universal structuralist position. The next section will show, however, that this is not the full picture.

Second, in (1) to (4) the specific class of models of an axiom system expressed here via higher-order quantification is directly linked to Carnap's tacit semantics for the type levels, i.e. the ranges of the bound variables of STT. In the context of his formal reconstruction of axiomatics there is thus a direct connection between the formal semantics of a theory, specifically its class of possible models and the interpretation of the higher-order segments of the background language STT. The class of models referred to in each of the above cases is identical to the range of the higher-order quantifier of STT. It follows from this that Carnap's (DA) concerning the formal models has to be directly reflected in his tacit semantics of his formal language. We mentioned above that (DA) effectively restricts the class of models to those expressible in STT. Consequently, it must have a similar effect on the (fixed) ranges of the higher-order segments of Carnap's underlying theory of types.

1.5.2 Higher-order quantification

We have pointed out before that the "basic system" for Carnap's theory of axiomatics was understood as fully interpreted. Compare the following remark on this point in *Untersuchungen*:

> Every treatment and investigation of an axiom system therefore presupposes a logic, specifically a contentual logic, i.e. a system of sentences that are not merely combinations of signs but that have a particular meaning. (Carnap 2000, 60)[78]

Thus, the constant expressions as well as the quantifiers of STT come equipped with a fixed semantic interpretation. It follows also that all formulas in the examples discussed above that generalize over the models of a theory are conceived as interpreted statements of STT.

Unfortunately, Carnap did not get any more specific about his intended interpretation of this "contentful logic" that forms the "basis for axiomatics" (ibid, 60). In particular, no explicit remark can be found in *Untersuchungen* concerning the intended interpretation of the higher-order quantifiers of the "basic discipline". In the following discussion, we will restrict our

[78] "Jede Behandlung und Prüfung eines Axiomensystems setzt also eine Logik voraus, und zwar eine inhaltliche Logik, d.h. ein System von Sätzen, die nicht bloße Zeichenzusammenstellungen sind, sondern eine bestimmte Bedeutung haben." (Carnap 2000, 60) Compare Chapter 2 for a more detailed study of Carnap's conception of type-theoretic languages.

attention for reasons of simplicity to a segment of Carnap's (finite) type-theoretic logic, namely SOL.[79] The pressing interpretive question then is: What is Carnap's assumed semantics for his "applied logicistic", in particular for the second-order domains of relations (of a given arity)?

In order to reconstruct his case, a modern presentation of the spectrum of the possible semantics for SOL will prove to be helpful. Koellner (2010) shows that for a second-order language L, a specific class of L-structures can be defined by a formula $\Phi(x, y)$. An instance of Φ is a structure $\langle M, S\rangle$ where M stands for the first-order domain and S for the second-order domain of the structure respectively (ibid, 1-2). The interesting point is to determine "what constraints" different Φs can place on the possible structures of L. Given this convention, the difference between what is usually classified as "standard" and "nonstandard semantics" relates to two limiting instances of Φ: in the minimal case, Φ restricts the class of models for L to models whose second-order domains S are restricted to the definable subsets of M. In other words, in the resulting "Definable SOL", $S = \mathrm{Def}(M)$, viz. "the set of subsets of M that are definable over M with parameters." (ibid, 31) The other, maximal case is that of "Full SOL." Here, the second-order domains of the models are $S = \wp(M)$, i.e. the second-order domain ranges over the full power set of a given M (see ibid, 33-36).[80]

Generally speaking, type-theoretic languages (and thus SOL) as used in the foundational programs up to 1930s were usually conceived to range over all appropriately segmented objects of a type-theoretic universe.[81] For instance, Shapiro argues for the adequacy of this "standard semantics" in the characterization of the historical use of type-theory in the works of Frege, Russell and others: "(...) standard higher-order logic is a good model of the logicist systems, and it would be a major distortion to think of them as if they were first-order (for example, to impose a non-standard semantics)" (Shapiro 1991, 179). In modern terms, this is

[79] Most of the axioms system Carnap discusses in Carnap (2000) and Carnap (1929) as well as their interpretations can be recast formally in SOL. For example, we saw that quantification over the first-level relations conceived as the interpretations for Peano arithmetic PA^2 can be expressed in SOL. Also, his discussion of "elementary arithmetic" based on the concept of "progressions" as binary first-level relations can be expressed in SOL.

[80] Compare Shapiro (1991, 80-96) for a similar discussion of standard and non-standard semantics for SOL.

[81] See e.g. Ferreirós (2007) and Grattan-Guinness (2000) for two broader historical studies of the evolution of logic during that time.

to say that the bound first-level predicate and function variables of a language were intended to range over the power set of the respective individual domain of the language.[82]

Nevertheless, turning to Carnap's use of "applied logicistic" for formalizing axiomatics, there is direct evidence that he must have conceived the higher-order domains of his interpreted language in a nonstandard sense. In the cases of quantification over models mentioned above, this nonstandard treatment follows directly from his substitutional conception of models outlined in Section 1.4.4. Given the fact that Carnap conceived his models as sequences of relational constants, it is obvious that – in sentences like $(\forall R)(\mathrm{f}R \rightarrow \mathrm{g}R)$ – the range of $(\forall \mathrm{R})$ is clearly restricted to the class of relational expressions of STT. Note that this understanding of higher-order quantification is not limited to Carnap's work on axiomatics. In fact, it can safely be said that prior to the *Logical Syntax of Language* (Carnap 1934), he upheld a substitutional understanding of higher-order domains (as well as of the first order domain) of a formal language.[83] This conception of models and of higher-order quantification necessarily implies a nonstandard semantic close to the above specified sense. The reason for this is that, particularly in mathematical contexts, substitutional quantification is restricted compared to standard extensional quantification. This holds particularly for mathematical domains with indenumerable cardinalities given the usual assumption that a substitution class of expressions of a given type is at most (infinitely) denumerable. Carnap himself was quite explicit on this

---

[82] Against Shapiro's view, these historical logics are sometimes also characterized in terms of a "non-standard semantics." It is argued for instance in Hintikka (1995) that Russell's ramified type theory in *Principia Mathematica* (Whitehead and Russell (1962)) effectively implies as non-standard interpretation. Due to Russell's predicativity constraints on the comprehension principles, the higher-order quantifiers effectively range only over a subset of the power set of the model's universe. The sets of admissible values for the higher-level quantifiers are thus effectively restricted to those elements (concepts, relations, sets,…) that are predicatively definable in type theory. Thus, due to this additional definability condition for higher-order entities, Hintikka follows, Russell and Whitehead "in effect assume a nonstandard interpretation" of the higher-order quantifiers of their logical language (Hintikka 1995, 38)

[83] This becomes obvious in Carnap's correspondence with Gödel from 1932 on several issues concerning drafts of *Logical Syntax*, most importantly, the valid understanding of the notion of 'analyticity' for formulas in a type-theoretic language. Here, in an early version of analyticity for formulas involving a second-order quantifier presented to Gödel, Carnap apparently suggested that a sentence $(X)X(0)$ is analytic iff $F(0)$ is analytic for all predicate constants F of the language. Gödel objections to this use of substitutional quantification eventually led Carnap to a change in mind concerning his conception of higher-order quantification and his subsequent move to an extensional treatment in *Logical Syntax* and later on. Compare Carnap (2002) on this point: "It may happen that, though all these sentences are true, '*M*(*F*)' is nevertheless false – in so far as *M* does not hold for a certain property for which no predicate can be defined in II (…) we will follow Gödel's suggestion and define 'analytic' in such a way the '*M*(*F*)' is only called analytic if *M* holds for every numerical property irrespective of the limited domain of definitions which are possible in II." (Carnap 2002, 106-107) Compare Awodey and Carus (2007, 37-38) and Coffa (1991, 290-293) for detailed discussions of this correspondence and its effects on Carnap (1934).

premise concerning the vocabulary of a formal language at several places in his work.[84] Given these inherent limits of this conception of (first- and higher-order) quantifiers in comparison to the extensional interpretation, it is clear that Carnap's conception of STT was nonstandard in a crucial way.

One might object here that Carnap's early substitutional approach was - in the absence of a sharply drawn syntax/semantics-distinction - in fact unintended or merely the result of a (generally shared) confusion between the expressions of a language and their semantic correlates. In this more charitable reading, he was not deliberately proposing a substitutional theory of models and quantifiers in a strong sense, but rather a quasi-extensional conception cast - for lack of better knowledge - in syntactic terms.[85] In the context of his work on axiomatics, there are in fact a number of passages in *Untersuchungen* that indicate a certain ambiguity in his understanding of the "domains" ("*Gebiete*") of a theory that can be mentioned in support of this view. For instance, at one place when Carnap is discussing Hilbert's consistency proofs for axiomatic geometry he mentions its domains of "individuals of real numbers (or ordered set of those)" (Carnap 2000, 94). More specifically, he states:

> As a model of an axiom system of Euclidian geometry e.g. different triples of real numbers and certain relations between them can be used (…). (ibid, 94)[86, 87]

There is not talk here of linguistic representations as the elements of the respective domain, but simply of relations between (triples of) real numbers, classes, and individuals. Hence, one might be tempted to infer that when Carnap speaks of logical constants (of STT) as the

---

[84] Compare, e.g., Carnap (2002): "(…) obviously the number of aggregate-designations is denumerable in every system. The richness is only assumed by means of axioms, and it is not demonstrable by designations. (names or descriptions)." (ibid, 270)
[85] A similar account can be identified in *Principia Mathematica* where it is not always clear whether Russell and Whitehead are discussing the syntactic expressions or their semantic relata when speaking of propositional functions.
[86] "Als Modell eines Axiomensystems der euklidischen Geometrie können z.B. die Tripel reeller Zahlen und gewisse Beziehungen zwischen Ihnen genommen werden (…)." (ibid, 94)
[87] In a second example he discusses the fact that for every non-formal, e.g. an empirical interpretation *S* of an axiom system, an isomorphic formal model *S*' can be presented. Here again the elements of a system seem to be understood extensionally: "For every system S consisting of such individual elements (including classes or relations of such individual elements or classes (or relations) of classes (or relations) of such individual elements etc.) there is a structure-identical system S' whose individuals are real numbers (or ordered sets of such) (…)." (ibid, 94) "Zu jedem System S, das aus solchen Einzelheiten besteht (einschließlich der Klassen oder Relationen solcher Einzelheiten, oder der Klassen oder Relationen von Klassen oder Relationen solcher Einzelheiten usf.), gibt es daher ein strukturgleiches System S', dessen Individuen reelle Zahlen (oder geordnete Mengen von solchen) sind (…)." (ibid, 94)

constituents of the models or the instances of (higher-order) quantification, what he actually means to speak about are the extensional objects denoted by them.

Even if a certain ambiguity concerning the proper nature of the semantic relata is granted for Carnap's case, one should not overlook the fact that his substitutional account is directly motivated by his logicist background. More specifically, it is a natural consequence of his definability assumption for mathematical concepts. Thus, there is also a deliberate reason for him to restrict higher-order quantification to a nonstandard semantics. This is most explicitly expressed in Carnap (1931) in the context of his discussion of Ramsey's approval of full impredicative comprehension in the theory of types (in contrast to Russell's ramified version in *Principia Mathematica*). In arguing against Ramsey's "concept absolutism" in the quantification over all properties, Carnap states:

> (…) I think we should not let ourselves be seduced by it into accepting Ramsey's basic premise; viz., that the totality of properties already exists before their characterization by definition. (…) I think we ought to hold fast to Frege's dictum that, in mathematics, only that may be taken to exist whose existence has been proved (and he meant proved in finitely many steps). I agree with the intuitionists that the finiteness of every logical-mathematical operation, proof, and definition (…) is required by the very nature of the subject. (Carnap 1931, 102 quoted from: Benaceraff and Putnam 1983, 50)

In a draft of his Königsberg talk on which the article was based in Carnap's *Nachlass*, he gets more explicit about the kind of definability assumed here:

> We in contrast: not theological as Ramsey, but constructivistically: "all properties" means: "all expressions with 1 variable constructed from the basic properties according to given principles of construction." (RC 090-16-09)[88]

Two points should be stressed here. First, Carnap is fully explicit here about his motivation for a deliberate restriction of the higher-order domains: Ramsey's "concept absolutism" has to be rejected from a *Fregean* constructivist point of view. Carnap's explicit constructivism mentioned here essentially coincides with what we described as (DA) for formal models, i.e. the assumption that all mathematical higher-level entities (such as models) have to be

[88] "Wie dagegen: nicht theologisch, wie Ramsely, sondern konstruktivistisch: "alle Eigenschaften" bedeutet: "alle Ausdrücke mit 1 Variablen, die aus Grundeigenschaften nach gegebenen Konstruktionsprinzipien konstruiert sind." (RC 090-16-09)

expressible in the "basic discipline". In Carnap (1931) the condition for the existence of (and thus quantification over) higher-level concepts is that they have to be (logically) constructible or definable.[89]

Second, note that here again a certain ambiguity concerning the effective elements in the range of quantification can be identified: they can either be considered as extensionally understood "properties" or as "expressions" and thus substitutionally. Nonetheless, in either way (and in contrast to Ramsey's suggested full quantification), the higher-order domains are clearly supposed to be restricted to logically definable entities. Hence, Carnap tacit semantics for higher-order quantifiers is significantly nonstandard irrespective of the intrinsic constraints of the substitutional account, due to this logicist constructivism in concept formation and in quantification over properties.

Given this *Fregean* doctrine in Carnap's philosophy of mathematics we can specify his tacit semantics for the higher-order regions of STT in this way: in the case of second-order domains, the assumption implies something comparable to "Definable SOL" specified above. In Carnap's account, the second-order domain S of the intended model $\langle M, S\rangle$ consists of the relations on and sets of M that are definable in the language. Given the fact that the class of defining formulae (and thus the introduced constants denoting them) have to be denumerable, it is clear that this semantics for SOL, i.e. the range over relations, is strongly restricted compared to full second-order quantification over the power set of M. Thus, Carnap's logical constructivist doctrine effectively restricts the usual full higher-order semantics for his formal language STT.[90]

---

[89] In fact, this Fregean constructivism can be considered as a philosophical invariant in Carnap's work from that time. Compare also his remarks in (Carnap 1930/31) on the "logicistic definitions" of Frege and Whitehead/Russell: "The concepts of mathematics can be deduced from the logical concepts (…)" (ibid, 298). A few lines later, he talks of a "formally developed system of basic concepts: starting from some basic concepts all concepts of the mentioned fields [GS: i.e., arithmetic, analysis, and set theory] can be constituted by Stepp-wise explicit definitions." (ibid, 300) For a closer analysis of Carnap's logicism in the 1920s see again Reck (2004, 172-175) as well as Bonk and Mosterin (2000, 28-29).

[90] This fact is eventually noted in *Logical Syntax* where Carnap - under the influence of Gödel's suggestions from 1932 - purposes an extensional treatment of the fully interpreted higher-order quantifiers of his languages LII. Compare the following passage in §55 on "improper operators" of LII: "The universal and existential operators of higher levels – that is to say, with p (or f) – are apparently improper in the majority of languages. (...) Let the ${}^2pr^1$, pr^2, of II_l [GS: the logical sub-language of II] designate (in material speech) a property which belongs to all the number properties which are definable in II_l but, on the other hand, not to all the number properties which are indefinable in II_l (...). Then $(p^1)(pr^2(p^1))$ is contradictory; the class of all closed variants of the operand is, however, analytic: and hence this contradictory sentence can not be a consequence of it. Further, on the same hypothesis $(\exists p^1)(\sim pr^2(p^1))$ is analytic; hence all closed variants of the operand are contradictory; the content of the existential sentence is null and the product of the contents of the variants is the total content; therefore the former is a proper sub-class of the latter." (Carnap 2002, 197)

With this in mind, we are now finally in a position to address the question what implications Carnap's logicist assumption (and the resulting nonstandard semantics for his logical system) have for his attempt to capture formally the model-theoretic condition (3) of modern axiomatics via "Explizitbegriffe".

1.6 A restricted "Explizitbegriff"?

Previous Sections 1.2.2 and 1.3.2 showed that (3) could be recast in structuralist terms as (3*). We mentioned that a proper understanding of (3*) depends on one's additional semantic assumptions concerning the interpretation of the structures referred to here. In a "relative structuralism," this concerns the specifics of one's conception of models. In case of "universal structuralism," the intended interpretation of the higher-order quantifiers in formulas like $(\forall R)(\forall S)(\forall T)\ [AS(R, S, T) \rightarrow \varphi(R, S, T)]$ turns out to be decisive.

As was pointed out, Carnap's own semantic assumptions underlying his structuralist account of axiomatics combine elements from both strands. Here, a heterodox notion of formal models is used vis-à-vis a universal approach in cases where models are quantified out in higher-order quantification. Moreover, we saw that this generalization over the interpretations of a theory is not eliminative in a strong sense of the term since it does not fully abstract from the particular instances. Quantification in Carnap's case remains effectively constrained by the logical definability conditions imposed on the level of individual model construction (i.e. his "*Modellaufweis*"). The implications of this for his reconstruction of (3) (or (3*)) are not difficult to see: the range of the quantifiers in the reformulation of Carnap's "Explizitbegriff"

$$(\forall R)(\forall S)(\forall T)\ (AS(R, S, T) \rightarrow \mathrm{EB}_{\mathrm{AS}}\ (R, S, T))$$

is clearly restricted to those instances of R, S, T that are explicitly definable in his "basic system". Hence, Carnap's talk the classes of models defined by an $\mathrm{EB}_{\mathrm{AS}}$ effectively means the class of logically definable models of the AS.

So much for Carnap's side. What have not been specified so far are the actual semantic assumptions underlying modern axiomatics, e.g. in Hilbert's own work. In particular, for (3), this concerns the intended class of models specified by his axiomatic theories for geometry or analysis. In order to get a definitive answer to whether Carnap's notion of "Explizitbegriffe" provides a valid reconstruction, we have to get a closer picture on the implicit semantics underlying (3) in the modern axiomatic practice.

1.6.1 Frege vs. Hilbert: which shared assumption?

In addressing this, a short detour over the already mentioned Frege-Hilbert controversy will be instructive. In particular, a comparison of two recent commentaries on the historical debate by Tappenden (1997) and Shapiro (2005) will help to clarify the semantic implications of Hilbert's account. Both papers attempt to relativize to a certain degree the received view of the conceptual gap between Frege's *universalist* and Hilbert's *model-theoretic* conceptions. They differ substantially, however, in their understanding of how Frege and Hilbert's respective semantic assumptions can actually be related to each other.

Tappenden (1997) discusses the historical controversy primarily with a view to reconcile Frege's own reconstruction of the axiomatic method with Hilbert's essentially modern account. He starts by stressing the former's clear opposition to the latter's "schematic" understanding of primitive terms and of axioms as "unquantified schemata":

> Frege makes it evident that he is opposed to the idea that one can determine a subject matter by writing down a set of such uninterpreted sentences and indirectly fixing a family of interpretations for them. (ibid, 208)

Nonetheless, Frege, as we noted in Section 1.2, formulated an alternative version of the semantic innovation (3) in the framework of his interpreted higher-order logic. Concerning this reconstruction in terms of higher-order concepts, Tappenden states:

> That is, although he objects to the idea of resting with a set of only partially interpreted schemata and a class of models for them, he has no objection in principle to *the exploration of families of models using second-order quantification*. (ibid, 209, my emphasis)

This point is further strengthened. Tappenden holds that there is in fact no effective logical difference between this "substantive" expression of (3) and Hilbert's own schematic approach in the formulation of axioms. This claim is based on two crucial assumptions: first, that Frege understood higher-order quantification substitutionally. In this sense, the quantifiers range over a substitution class of linguistic expressions and not, as in the extensional treatment, over a universe of higher-order entities (relations, sets, etc.). Second that this substitutional account is semantically equivalent to a schematic conception of logic:

As is well known, the distinction between quantified sentences and schemata essentially disappears when the quantifiers are substitutional; the quantifiers serve just as indicators of substitutional order. (ibid, 210)

With respect to (3) Tappenden's claim can thus be paraphrased this way: If Frege held a substitutional interpretation of higher-order quantifiers, then the class of models quantified over in a sentence like $(\forall R)(AS(R) \rightarrow \varphi(R))$ is in effect identical to the one in Hilbert's (now standard) schematic treatment. Thus, given the additional assumption concerning the implicit semantics of the "substantive" logicist tradition, the two versions of (3) result to be logically equivalent.

I see a problem in this account that does not concern the semantics of Frege's "substantive" logic. (I will simply sidestep the debate here on whether Frege actually held a substitutional account of second-order quantification).[91] The problem, in my mind, rests on a misconception of Hilbert's "schematic" account of axiomatics. It is certainly true that he introduced in Hilbert (1899) the notion of *disinterpreted*, schematically understood mathematical primitives. This is precisely the semantic innovation (1). It would be a mischaracterisation, however, to label his theory as generally *schematic* in character. This can be illustrated by looking at the informal logic used in his expressions of the geometrical axioms. Tappenden's account seems to suggest that Hilbert's axiomatization is schematic in the sense of a first-order theory with axiom schemes. This would be to confuse, however, the *schematic conception of primitives* with a *schematic version of axioms*. While it is true that the bulk of Hilbert axioms for geometry can easily be formalized in FOL, his main continuity axiom, the *axiom of completeness*, can only be expressed in HOL. Generally, the axiom plays a pivotal role in Hilbert's theory of Euclidian geometry. It is introduced to secure the completeness (in terms of model maximality and categoricity) of the axiom system and thus also its adequacy in relation to analytic geometry.[92] The way it is expressed in the *Grundlagen der Geometrie*, it effectively involves quantification over all possible domains (and thus over all models) of the theory, quite similar to the idea of a definition of the class of structures of an axiom system expressed by (3). Given this, it would clearly be misleading to characterize Hilbert's axiomatic approach to geometry as purely schematic.

---

[91] Compare Heck (2007), Linnebo (2004) and the references given there.
[92] See Chapter 3 for further details on Hilbert's axiom and Carnap's reception of it. See also Hallett (2008).

A similar point can be made for the case of arithmetic. Here too, it is sometimes argued that the schematic first-order version of the Peano system does not adequately capture out intuitions about arithmetic.[93] The bone of contention here is obviously the axiom of induction. In the first-order formulation, it is expressed as the axiom scheme

$$(\mathrm{AI}^1) \qquad (\varphi(0) \mathbin{\&} (\forall x)(\varphi(x) \rightarrow \varphi(S(x)))) \rightarrow (\forall x)\varphi(x).$$

The schematic letter φ allows instantiation by all formulas FOL that define a property or set of the intended individual domain. If the axiom is formalized in this sense, then there is in fact no substantial difference between AI^1 and a *Fregean* reconstruction of the axiom in substitutionally interpreted HOL. The possible instances for the schematic letter φ (in the modern schematic version) are effectively the same as those of the bound predicate variable *X* (in the "substantive" quantified version).[94] With the move to second-order Peano arithmetic, φ is also substituted by a bound variable *X*. The second-order version of AI then reads:

$$(\mathrm{AI}^2) \qquad (\forall X)([X(0) \wedge (\forall x)(X(x) \rightarrow X(S(x)))] \rightarrow (\forall y)(X(y))).$$

Unlike in the "substantive version", the range of *X* is not restricted here to definable properties, but includes all properties (or sets) on the given individual domain of the intended model. Thus, in contrast to AI^1, the range of possible arithmetical relations quantified over in AI^2 is not limited by definability in the formal language.[95]

Tappenden's proposed reconciliation between Frege's *substantive-but-substitutional* approach and the modern model-theoretic one holds in the cases of geometry and arithmetic only if the respective formal theories are understood in the first-order version (i.e. without a completeness axiom and a second-order axiom of induction).[96] However, in both Hilbert's and Peano's original presentations of their theories informal but explicit use of higher-order quantification was made: in the first case over all Euclidian "systems of things" in the axiom of completeness, in the second case over all arithmetical properties in AI^2.[97] Thus, both axioms were not conceived as first-order axiom schemata in early modern axiomatics, but as fully interpreted second-order principles of the respective theory. Given this, one could in fact

---

[93] Compare e.g. Kreisel (1972).
[94] For additional, more subtle differences between the two accounts see Chapter 3.
[95] Compare Shapiro (1991, 110-113), for a more detailed account of the differences between the two versions.
[96] Compare again Shapiro (1991) on the difference between the two formalizations: "The second-order axioms of induction, completeness (…) apply to any subset of the domain, whether definable in the language or not." (ibid, 111) More specifically: "Each of the first-order theories, formulated with a scheme, entails that for each model *M*, the relevant principle (induction, completeness, or replacement) applies to every first-order definable subset of, or function on, the domains of *M*. But that is all. They do not, and cannot, state that the principle applies to every subset of the domain." (ibid, 112)
[97] Compare Awodey and Reck (2002) for further details on the historical axiomatizations.

reverse Tappenden's argument in this way: given Frege's "substantive" conception of logic, it is not substitutional (or definable) but full second-order quantification that is needed in order to capture the original intentions behind these axioms in these early axiomatic systems. A formal reconstruction in which the higher-order domains are conceived nonstandard cannot adequately capture the modern axiomatic practice.

In Shapiro (2005), greater focus than on Frege's conception is laid on the actual logical assumptions in Hilbert's account, specifically on the question whether there are "substantive" assumptions in his informal use of logic. Similarly to Tappenden (1997), it is also stressed here that Hilbert is a major proponent of the "algebraic," model-theoretic tradition (in opposition to the "substantive" or, in Shapiro's terms, "assertory" conception of logic of the logicists like Frege) (ibid, 68). However, what Tappenden sees as the overall "schematic" conception of Hilbert's axiomatics, is characterized differently here:

> One can take any algebraic sentence and interpret it directly as a proposition about all systems of a certain sort. Consider, for example, the Euclidian sentence that there is a point that lies between any two distinct points. From the algebraic perspective, this comes to something like this:
> (*) In any (possible) Euclidian system S, for any two distinct objects a, b in S that are 'points-in-S', there is a third object c that is also a point-in-S, and c lies between-in-S a and b. (ibid, 68)

Shapiro points out here an implicit universalist structuralism, more precisely the tacit quantification over all (possible) models of a system in Hilbert's schematic version of expressing his axioms. Concerning the implicit semantics of this structuralist account, he remarks:

> What is the status of statements like (*)? It would seem that for the algebraist [GS: i.e. Hilbert], such sentences must themselves be assertory. (…) Moreover, it would run counter to the spirit of Hilbert's approach to think of the opening quantifier in (*) as itself restricted to a particular system. (ibid, 68)

Shapiro's claim made here is this: The implicit quantifiers in sentences like (*) in Hilbert's theory were conceived to be fully interpreted and to range over all possible models of the axiom system. Thus, as in Frege's treatment of axiomatic systems, there is also a tacitly assumed generalization over the interpretations in Hilbert's theory. His informally used logic in *Grundlagen der Geometrie* is "assertive" (or "substantial") despite the general algebraic nature of the project.

Concerning the notorious controversy with the 'logical universalist' Frege, Shapiro mentions that "such statements [GS: as (*)] are at least implicit in Hilbert's motivating remarks and in the correspondence with Frege." (ibid, 68) Frege's own attempt "to recapture Hilbert's perspective" was thus to construct "statements like (*) in his own logical system. And, of course, for Frege such statements, like all others in mathematics, are assertory." (ibid, 68)

This reading, as the one proposed by Tappenden, both attempt to reconcile the two historical positions by stressing their mutual assumptions. The fundamental difference of Shapiro's to Tappenden's interpretation concerns precisely the nature of this (allegedly) shared assumption: The latter assumes a substitutional interpretation of Frege's logical language. He argues that a "substantive" reconstruction of axiomatics with this restricted semantics meets Hilbert schematic conception. Shapiro, in contrast, emphasizes the fact that not only in Fregean logicism, but also in Hilbert's algebraic context, statements like (*) have an "assertory" sense, viz. they are substantive claims in an informal but interpreted logic.

In Hilbert's case, these logical resources are not visible given the fact that Hilbert (1899) is fully written in colloquial language and thus not a formalized theory. Moreover, we already mentioned that most of his axioms can be expressed formally in FOL. Take for instance the axiom of order underlying Shapiro's example (*). Hilbert and Bernays (1934) give the following formalization of the axioms in FOL:

$$(\forall x)(\forall y)(x \neq y \rightarrow (\exists z)(Zw(z,x,y)))$$

where x, y, z are individual variables that range over points and Zw is a tertiary relation that expresses '*betweenness*' (ibid, 6).

The mathematical motivation of the algebraist to turn to higher-order quantification (over models) does not concern this object-level in which the theory is expressed. It concerns the theory's metatheory. Shapiro's main point is that in Hilbert's axiomatics, despite his schematic conception in the object theory, there has to be an assertive use of logic in his interpreted metatheory. This is necessary in order to develop the well-known (semantic) consistency and independence results in Hilbert (1899) (see ibid, 70).[98] We saw that this full use of (higher-order) logic is also present - on the object-level – in Hilbert's *axiom of completeness*. The axiom is clearly metatheoretic from a modern point of view since it involves the quantification over the models of the base theory.[99] More generally, as Shapiro

[98] For the central importance of a metatheoretic account in Hilbert (1899) see also Hallett (2008) and Hintikka (forthcoming).

[99] Compare 3.2 for a closer discussion of the axiom as well as of Carnap's formal reconstruction of it as a maximal axiom.

makes clear, sentence (*) itself has to be considered as a metatheoretic reformulation of the original axiom in this sense:

(*) For all individual domains P and all assignments to Zw: $(\forall x)(\forall y)(x \neq y \rightarrow (\exists z)(Zw(z,x,y)))$

This is a quantified out version of Hilbert's original axiom that is fully in line with the universal structuralist account described above. Note that the underlying semantics for the meta-mathematical sentence (*) has to be strong enough in order to allow talk of "all ranges of *P*" and "all assignments to *Zw*", hence to quantify over all possible interpretations of the terms as defined by the axiom.[100] The most natural way to secure this is to assume a fully semantics for the language in which (*) is expressed. Thus, irrespective of whether the background language in a given metatheory is formalized or kept informal (as in Hilbert's account), the quantification indicated in (3) as well as in sentences like (*) is most likely supposed to be standard in modern axiomatics. Thus, when Hilbert tacitly assumed the prefix "for all models" for all of his axioms, he most likely *meant* every (mathematically possible) model and every model expressible in a given language.

Here again, the difference between this view and Tappenden's account with respect to (3) should be stressed. Tappenden claims that Frege's reconstruction of (3) via substitutional higher-order logic is equivalent to Hilbert's schematic definition of the class of models. From Shapiro's perspective, the mistake of this view lies in the fact that (3) is certainly not expressed in the object theory, but in the metatheory. As such it is not a schematic but an assertive statement that presupposes sufficiently strong mathematical resources to refer to all models of a theory. Hence, one can view Shapiro's interpretation of the Frege-Hilbert controversy in direct opposition to Tappenden's version: according to the former, Hilbert's metatheoretic and thus assertoric treatment of (*) (and (3)) and Frege's formal reconstruction are in fact comparable under the premise that Frege actually assumed full higher-order quantification in his logical system.

[100] Compare Hintikka on the necessary semantic presuppositions of the axiomatic account: "When a certain type of structure is studied mathematically, this study is not restricted to what can be found out about all such structures or about some subclass of them. A perhaps even more important task in practice is to reach an overview over all such structures." (Hintikka (forthcoming))

1.6.2 Consequences for Carnap's version

Shapiro's understanding of metatheoretic sentences like (*) implicit in Hilbert axiomatics on first sight also seems to apply to Carnap's reconstruction of similar sentences in higher-order quantification. In Carnap too, (*) would clearly be understood as a substantive truth in his "applied logistic". Similarly, this holds for his explication of (3) via "Explizitbegriffe".

Now, one might object to this comparison by stressing an important difference between Carnap's account from 1928 and the modern model-theoretic treatment of this issue. In the modern account sentences like (*) are expressed in an interpreted meta-language. In contrast, in Carnap's case, they are formulated on the object-level, i.e. in a single logical object language. We saw that in his *Untersuchungen*, this 'universal' background system is his type-theoretic "basic discipline", in which both the axiomatic theories and their (semantic) metatheory (including sentences like (*)) are captured. The examples of the metatheoretic sentences or formulas (all clearly semantic in content) discussed in Section 1.5.1 are further cases in point for this.[101]

However, concerning statements about the semantics of an axiom system (like (3)), the missing object-/meta-language distinction is not a consequential difference to the algebraic tradition. First, Hilbert in *Grundlagen der Geometrie* clearly engaged in semantic metatheoretic investigation for his axiom system without possessing anything like a proper meta-perspective, let alone a proper formal meta-language. Here, too, the informal metatheory (in particular his independence and consistency proofs) is expressed in one and the same language as the axiom system, in his case in colloquial *German*. Sentences like (*) described by Shapiro would thus also be expressed in this universal colloquial language.

Second, it is questionable whether a metatheoretic perspective has to be considered as a *conditio sine qua non* for large parts of formal semantics. In Frege's as well as in Carnap's formal reconstructions of axiomatics, there is no stringent reason why (at least some) of their explications of semantics notions cannot be expressed in the proper object language. This is true in particular if the languages in question have strong enough logical resources, e.g.

[101] The missing distinction between meta- and object-theory in Carnap's *Untersuchungen* has been discussed extensively in recent scholarship. Compare Coffa on Carnap's "monoliguistic approach": "Carnap's book was thus inspired by the somewhat epicyclic aim of showing that everything of value in metamathematics can (or should) be expressed within the monolinguistic framework of *Principia Mathematica*." (Coffa 1991, 274) A more differentiated account is given in Reck (2004): "(…) if looked at in more detail his [i.e. Carnap's] approach is only partly metamathematical and model-theoretic, since he does not separate precisely between object-language and meta-language, and since he does not yet have a full understanding of the model-theoretic notion of an interpretation for a formal language at his disposal." (Reck 2004, 170). See Awodey & Carus (2001, 158-161) for a critical discussion of Carnap's 'universalist' approach in his treatment of the "Gabelbarkeitssatz". Compare also Bonk and Mosterin (2000, 35-36).

sufficient (interpreted) higher type-levels to present the semantics of the expressions and formulae of the lower type-levels. Carnap's underlying universal language for the "basic discipline" is finite STT. Here, this condition is clearly met. The higher interpreted type levels can definitely be used to present the semantics, e.g. the models for theories (like Peano arithmetic) expressed in the lower regions of the language. A move to a separate meta-language is therefore not necessary, especially for the axiomatic theories discussed in Carnap (2000).

Where the discussed accounts actually differ from each other is in the implicit background semantics, more specifically in the tacit assumptions concerning the language in which sentences like (*) and also (3) are expressed. In other words, where Frege's and Hilbert's position can conceptually differ from each other is in the semantic understanding of the higher-order quantifiers of the respective (formal or informal) background language.

Applied to Carnap's case, the main point to see is that Hilbert's informal use of quantification and the formal use of quantification in the former's STT are "assertory" or "contentual" in different senses. I fully agree with Shapiro (and contra Tappenden) that there is an assertoric use of quantifiers in 'Hilbertian', viz. modern axiomatic practice, in particular in the presentation of the model-theoretic results like (3). I would add that the semantic or set-theoretical resources necessary for expressing this result are most neatly codified in a SOL (or alternatively in ZFC). Moreover, talk about "all (possible) models" of a theory seems to presuppose a full, standard interpretation of the underlying metalanguage. The quantification involved here ranges over all "possible" models, i.e. all Euclidian structures in which the theory in question holds. Therefore, I take it that the intended metatheory for Hilbert's theory of Euclidian geometry, if formalized, would have to be cast in fully interpreted SOL (or something equivalently strong).[102] It follows from this that any formal reconstruction of (3) has to employ equally strong logical resources if the aim is to capture Hilbert's metamathematical account.

We saw in Section 1.4.5 that for Carnap the possible models of an axiom system are those expressible (or definable) in his formal background language. In particular, the definability assumption led him to treat formal models in a substitutional way, i.e. a tuples of (logical) constants of STT. Moreover, Section 1.5.2 showed that the same logicist definability condition for the introduction of higher-order entities led him to assume a nonstandard

[102] Compare Shapiro (1991) for a detailed discussion of the set-theoretic import of different versions of SOL.

semantics for the higher-order domains of his logical language. In his work on the semantic metatheory of axiomatics, this affects cases that involve quantification over the models of an axioms system: the range of a bound “model-variable” of an axiom system is thus restricted to the logically definable models. It also follows from this that Carnap’s concept of an “Explizitbegriff” effectively determines the class of definable models of an axiom system and not of all mathematically possible models. It therefore remains limited in context to the model-theoretic assumption (3) in modern axiomatics.

1.7 Conclusion

The aim of this chapter was to provide a close survey of Carnap’s early theory of axiomatics from the 1920s. We saw that his formal reconstruction of axiom systems and their model theory was based on a tacit but consequential logicist assumption inherited from Frege and Russell. Carnap’s discussion of formal models is subject to a strong definability assumption in the sense that each model of a theory has to be expressible in this logical background system. One consequence of this was that Carnap effectively upheld a nonstandard semantic conception of higher-order logic, for the case of first-level relation quantifiers, definable SOL. A second consequence was that certain of the crucial semantic innovations of modern axiomatics, most importantly the idea that a theory semantically defines the class if its models could not be adequately captured in his formal reconstruction. These limitations resulting from the definability condition (and thus the nonstandard semantics for his type-theoretic language) not just affect Carnap’s “Explizitbegriffe.” It also has strong limiting effects on his overall project in *Untersuchungen* to provide a formal explication of the (semantic) metatheory of axiomatics. For Carnap’s “Gabelbarkeitssatz,” this was already discussed in Awodey and Carus (2001) and Reck (2007). It is important to stress that the tacit logicist definability assumption for models also affects Carnap’s formal reconstruction of the individual notions of completeness, most importantly that of *categoricity* (“*monomorphism*”). The categoricity of a mathematical theory, e.g. of second-order Peano arithmetic, is nowadays usually proved in a fully interpreted higher-order metatheory (see Shapiro 1991, 82-83). The fact that Carnap assumed a nonstandard interpretation of his universal background logic makes his explication of the notion in *Untersuchungen* problematic at best.

In this chapter we have gained an understanding of the inherent limits of Carnap’s work on axiomatics, in particular his attempt to provide a formal semantics for axiomatic theories. We will see in the next chapter, on the other hand, that his early account of formal models has in several ways anticipated constitutive features of the modern notion.

Chapter 2: Carnap's early semantics: models, extremal axioms, analyticity

2.1 Introduction

One interpretive issue concerning Carnap's early contributions to formal semantics, specifically his theory of models for axiomatic theories formulated in the only recently published manuscript *Untersuchungen zur allgemeinen Axiomatik* in 1928 (Carnap (2000)) has yet not been resolved conclusively. It concerns the question whether Carnap conceived his models to allow for domain variation, i.e. whether he held a variable domain conception (in the following VDC).[103] Whereas the first, published part of Carnap's manuscript does not provide conclusive evidence for deciding on this issue, I will show that the results of the second, projected part, which exists in fragmentary form in Carnap's *Nachlass* (RC 080-01-01 to RC 080-01-33), parts of which were published in Carnap and Bachmann (1936), will give decisive cues for the understanding of his early conception of the domain of a model.

It has been stressed recently, most notably by Hintikka that due to Carnap's 'universalistic' conception of logic – mainly inherited from Russell – the idea of changing a universe of discourse for a language was simply inconceivable for him.[104] Drawing to van Heijenoort's now classic distinction between two opposing conceptions of logic, the *logic as calculus* vs. the *logic as language* (van Heijenoort (1967)), Hintikka argues that the "universality assumption" involved in the second is also a tacit invariant in Carnap's thinking throughout his intellectual career. With regards to semantic issues this background assumption results in a "one domain thesis" (Hintikka 1991) according to which the domain of an interpretation, i.e. the range of the individual quantifiers, is invariantly fixed with the specification of a language:[105]

---

[103] For recent work on Carnap's axiomatic project see Awodey and Carus (2001), Reck (2007), Goldfarb (2005), Bonk and Mosterin (2000), and Awodey and Reck (2002).

[104] See esp. Hintikka (1991) and Hintikka (1992).

[105] "[Carnap] assumes that, in order to interpret a first-order language, we have to specify some one given domain of individuals to which that language pertains. (...) the members of that given domain of individuals can still assume different structures. They can have different properties and different relations to each other." (Hintikka 1992, 175) Hintikka's main focus here is Carnap's specific understanding of possible worlds in his modal semantics in Carnap (1947). Nevertheless, as we will see, he takes the same remarks to hold also for his theory of models for standard logical languages.

I will in the following term this assumption the fixed domain conception (in short FDC).[106] According to Hintikka, FDC concerning logical, formal languages systematically hinders Carnap, being one of the inventors of formal semantics, from a full apprehension of a standard "model-theoretical viewpoint" (ibid.). It is also held that Carnap's tacit FDC can be traced back to his early contributions to the philosophy of logic and mathematics and related topics in formal semantics. In support of this claim, Hintikka (1991) refers to Carnap and Bachmann (1936), a neglected technical article that presents (with slight modifications) results formulated in the second, projected part of *Untersuchungen.*[107] The paper marks a highly innovative contribution to a topic in mathematical axiomatics broadly discussed at the time, namely the use, function and limits of "extremal axioms", i.e. mathematical axioms that impose minimal or maximal conditions on the models of a theory in question. Hintikka makes two interpretive claims about Carnap's treatment of extremal axioms: first that his formal explications misrepresent the actual use of the corresponding axioms in mathematical practice, i.e. that Carnap fundamentally misconceives the usual intentions underlying axioms like Hilbert's completeness axioms for geometry and analysis or Peano's axiom of induction for arithmetic. Second that this misconception on Carnap's side can be explained by the fact that he already presupposed something like FDC in his theory of models at this early phase in his intellectual career.

The principle aim of this chapter is twofold: first, to provide a precise and historically sensitive account of Carnap's early conception of models. For this, I will draw extensively on unpublished material from Carnap's *Nachlass*. Second, it is to argue against both of Hintikka's claims. I will show that the conclusions Hintikka draws for Carnap's tacit semantics result from an inadequate understanding of the latter's heterodox conception of models in *Untersuchungen*. Hence, the former fails to take into account a number of subtle distinctions concerning Carnap's (and Bachmann's) auxiliary notions of 'model structure', 'model domain', and 'model extension' that provide the conceptual framework for Carnap's formal explication of extremal axioms. In this chapter I first sort out different conventions

[106] Hintikka gives the following more general specification of this assumption: "If we cannot in our theorizing assume a different set of language-world relations in general, we cannot assume a different set of name-object relations, either. What this means is that a specification of the domain of individuals is part and parcel of the interpretation of any one language. If the domain of individuals is changed, we are dealing with a different language." (Hintikka 1991, 328) For a similar assessment of Carnap's limitative assumptions underlying his semantics compare Awodey (2007).

[107] Apart from Hintikka (1991), Carnap's theory of extremal axioms has not been subject to extensive scholarly discussion. A short commentary can be found in Bonk and Mosterin (2000) where Carnap's theory is presented and the projected structure of the second part of *Untersuchungen* is outlined. For a short discussion see also Awodey and Reck (2002) and Reck (2007).

introduced by Carnap to allow domain variation for his models in the axiomatic context. Second, it is argued that once a more balanced account of Carnap's underlying semantics of these auxiliary notions is provided, Hintikka's interpretive claims cannot be upheld. To the contrary, a more refined picture of Carnap's account of models and model extensions will suggest that he - despite the heterodox character of his notion –conceived domain variation for his models in a sense structurally comparable to the modern notion.

The chapter will be organized as follows: In Section 2.2 an outline of Carnap's theory of extremal axioms (as presented in Carnap and Bachmann (1936)) is given. Following a discussion of Carnap's notion of a model's structure (Section 2.2.1) and the presentation of a simple example of extremal structures (Section 2.2.2), Hintikka's interpretive claims concerning the theory will be discussed (Section 2.3). The main aim in Section 2.4 will be to develop a detailed analysis of Carnap's heterodox conception of formal models. Certain differences concerning the conceptual framework of Carnap's definition of a formal model between the 1936 paper and *Untersuchungen* will be discussed (Section 2.4.1). This allows outlining what I will term a 'domain-as-fields' conception of models held by Carnap (Section 2.4.2). Moreover, by drawing to unpublished material from Carnap's *Nachlass*, a second convention for domain variation based on a type flexible understanding of his formal background language will be discussed (Section 2.4.3.) With this in hand, an interpretation of Carnap's technical notions of 'submodel' and 'model extension' alternative to Hintikka's account will be presented in Section 2.5. It will be argued - contra Hintikka - that the explications given in 1936 are fully consistent with the informal use of these notions in mathematical practice. Additional evidence for this view is presented by taking into consideration Carnap & Bachmann's treatment of genuinely mathematical examples of model extension from "basic arithmetic" and geometry (section 2.5.1). Finally, an alternative convention for models extensions via language transitions outlined in the final section of Carnap and Bachmann (1936) will be discussed in Section 2.5.2. I will argue that this provides additional support for my reading of Carnap's notions of model and model extension.

2.2 Extremal axioms

In the first part of *Untersuchungen*, Carnap describes of formalized version of axiomatic systems as "theory-schemata" that can be presented logically as systems of sentential functions. The 'basic terms' ('*Grundzeichen*') of a theory are expressed by typed relation or class variables. Systems with more than one primitive term of the form f(P, Q, R, …) are abbreviated by a 'model variable' ('*Modellvariable'*) M of a specified type (see Carnap 2000, 88). The resulting system f(M) allows different empirical and formal interpretations. The term 'formal model' is reserved for the second type of interpretations. The variable M ranges over the set of adequate models ("*zulässige Modelle*") of f. A model M_1 is then defined as an ordered sequence of relations (and thus as a relation itself) of a specified type that present admissible values for the predicate variables (P, Q, R, …) and "satisfy" the axiomatic system f (see ibid, 95).

Given this, the second, projected part of the manuscript deals with different versions of maximal and minimal axioms, i.e. axioms introduced in order to fix the intended interpretation of a mathematical theory (RC 081-01-04 to -33). A mathematical instance of a maximal axiom Carnap discusses extensively is Hilbert's *axiom of completeness* (AC) for geometry first published in the French edition of *Grundlagen der Geometrie* in 1900. As an example of a minimal axiom Carnap refers to Peano's axiom of induction as well as to Fraenkel's *axiom of restriction* devised for axiomatic set theory around 1920.[108] All these cases share a similar limiting structure: they impose that the domain of elements D of the possible models of the axiom system in question cannot be extended/restricted to a larger/smaller domain of individuals D' without violating the relations implicitly defined by the theory. Thus e.g. Hilbert's AC for geometry imposes that the models satisfying AS+AC are the maximal models of the base system AS, i.e. models, whose domain of elements cannot be extended without violating one or more of the relations defined by AS. The underlying idea here is to rule out non-intended models and fix intended models for an axiom system. In Hilbert's case, the aim behind the addition of AC is to rule out models of Archimedean Euclidian Geometry that are submodels of analytic geometry (e.g. certain algebraic models satisfy AS+Archimedean axiom defined for independence proofs in §9 in Hilbert (1900)) and to secure the identity of (the models of) his axiomatic theory with "ordinary" analytic geometry.[109]

As already mentioned we can get a detailed impression of Carnap's attempt to provide a formal explication of these types of axioms around 1928 by looking at his joint paper with the German mathematician Friedrich Bachmann with the unifying title "Über Extremalaxiome" published a couple of years later (Carnap and Bachmann (1936/1981)).[110] Here the topic is

---

[108] For a closer discussion of Carnap's reception of Fraenkel's axiom of restriction see Schiemer (2010).

[109] For a detailed presentation of Hilbert's axioms of completeness see Awodey/Reck (2002) and Ehrlich (1997).

[110] The origin of the 1936 paper can be retraced from Carnap's correspondence with Bachmann documented in Carnap's Nachlass. In a letter dated July 27, 1934, Carnap invites Bachmann to write a joint paper that should

taken up again. At the outset of the paper a provisional explication of extremal axioms is given by elucidating their domain-fixing function:

Axioms of the sort of Hilbert's axiom of completeness, which ascribe to the objects of an axiomatic theory a maximal property in that they assert that there is no more comprehensive system of objects that also satisfies a series of axioms, we call a *maximal axiom*. The same axiomatic role as that of the maximal axioms is played in different axiom systems by *minimal axioms* which ascribe a respective minimal property to the elements of the discipline. Maximal and minimal axioms we call collectively *extremal axioms*. (Carnap & Bachmann 1981, 68-69)[111]

The remainder of the paper is essentially a conceptual clarification and formal explication of their understanding of these informal notions. Before turning to the specifics of their formal versions of extremal axioms, one point concerning Carnap's tacit semantics involved here should be emphasized: the task of giving a formal explication of these axiom types *prima facie* seems to presuppose a semantic apparatus comparable to modern model theory, i.e. an understanding of models and model extensions that allows domain variation. Without such an assumption a discussion of different models (of differing power) of a theory could hardly be meaningful. Note the way in which the intended effects of the maximal axioms are described above: the axioms "*ascribe*" a "maximal property" to the "elements of the discipline", i.e. its "systems of objects". From this intuitive picture alone, it is difficult to see what led Hintikka to claim that Carnap tacitly assumes a FDC in his theory of models. To address this, it will be instructive to take a closer look at Carnap's notion of a model's "structure" that plays a crucial role in his conception of model extensions. As will become clear in the next sections, the central difficulty in assessing Carnap and Bachmann's formal version of extremal axioms lies in the fact that whereas Hilbert (as well as modern accounts) speak of model extensions in terms of extensions of the individual domain of a model, Carnap and Bachmann – unconventionally - speak of extensions of a model's structure.

My main point will be to show that Hintikka's interpretive claims about Carnap's underlying semantics and his misconception of mathematical practice essentially stems from a misreading of Carnap's emphasis of structures over domains of individuals. For clarification,

combine his results from part two of *Untersuchungen* with Bachmann's own work on extremal axioms (RC 028-02-01).

[111] Now, it is well known that extremal axioms like Hilbert's AC are, from a modern point of view, logically problematic for they "conflate formal languages with their model-theoretic semantics" (Shapiro 1991, 185). One of my main points in this paper will be to show precisely in what regards Carnap's account differs from a full-fledged modern conception of model theoretic semantics.

I will briefly outline in the following two sections Carnap's notions of a model's structure and of structure extensions which I take to be at the core of this misinterpretation.

2.2.1 Model structures

A central notion employed by Carnap in his treatment of axiomatics is that of a relational "structure", a concept originally adopted and further developed from the theory of relations in *Principia Mathematica*.[112] The main characteristic of a relation's structure is its abstract character, i.e. the fact that it abstracts from any information about the extension of a given relation possessing this structure. In §22 of Carnap's *Abriss der Logistik* (Carnap 1929) we find the following characterization of the "structural properties of a relation":

> A relation is fully specified by its structure with regards to its formal properties if we understand them to consist of those properties that result from considering exclusively the identity and difference of the relational elements, irrespective of their other nature. (Carnap 1929, 54)[113, 114]

Applied in *Untersuchungen*, Carnap's main intention is to define formal properties of mathematical models axiomatically. Here, the following definition is given:

> The property f*P* of relations is called "structural property" if, in case it holds for a relation *P*, it also holds for every relation isomorphic to *P*." (ibid, 74)[115]

Expressed in Carnap's formal terms, this reads: $(P, Q)[\mathrm{f}P \ \&\ \mathrm{Ism}(Q,P) \rightarrow \mathrm{f}Q]$. A substantial part of Part 1 of *Untersuchungen* is reserved for the exact type-theoretic definition of a "generalized" or "complete isomorphism" $\mathrm{Ism}(P,Q)$ between n-ary relations Q and P where the elements of the different domains are of different types.[116] It is shown that the

[112] See Whitehead and Russell (1962, §21 and §30-38), compare also Russell (1919, 60-61).
[113] "Eine Relation wird durch ihre Struktur vollständig charakterisiert in Bezug auf ihre formalen Eigenschaften, wenn wir hierunter diejenigen ihrer Eigenschaften verstehen, die sich bei bloßer Berücksichtigung der Identität und Verschiedenheit unter den Relationsgliedern ergeben, unter Nichtbeachtung der sonstigen Beschaffenheit der Glieder." (Carnap 1929, 54)
[114] This notion of a relation's structure is closely related to Carnap's concept of "structure descriptions" of relations in Carnap (2003). In §11, he gives the following characterization of a relation's 'formal properties': "By formal properties of a relation, we mean those that can be formulated without reference to the meaning of the relation and the type of objects between which it holds." (ibid, 21)
[115] "Die Eigenschaft f*P* von Relationen heißt eine "strukturelle Eigenschaft", wenn sie, falls sie einer Relation *P* zukommt, auch jeder mit *P* isomorphen Relation zukommt." (ibid, 74)
[116] For Carnap's theory of relations as well as his theory of (model-)isomorphism see Carnap (2000, §1.7) as well as Carnap (1929).

construction of an isomorphism correlation between relations is closely connected to specification of their 'structure'. Two relations that are isomorphic also share "the same structure." This, in turn, can be defined via the isomorphism class of the relation by stating that f*P* holds as a "*structural property*" of a relation if it also belongs to all relations isomorphic to *P* (ibid, 74).[117] Since models are treated as higher-level relations, similar remarks hold also for a model's structure.

This approach can - as Carnap himself specifies - in fact be viewed as an approach to define a structural property of a relation via an invariance condition over "isomorphic transformations" of a certain kind (ibid, 74). However, it is not clear how the transformations referred to here are actually understood, especially when models are considered as the subject.[118] Fortunately, in Carnap and Bachmann (1936), the authors get more specific about the concept of a 'model structure'. Again the definition closely leans on the notion of "complete isomorphism":

> Since the complete isomorphism between *n*-place models (i.e., sequences with *n* members) is a 2^n-equivalence relation, *n*-place relations can be defined over the field of this relation, i.e., over the totality of constants of the basic, assumed language which can appear as elements of models, such that the *n*-place relations have the following properties: for each model there exists *exactly one such relation* which is satisfied by the constituents of the model and is satisfied by the constituents of two different models if and only if the models are completely isomorphic. The relations so determined we will call structures and will say that model M_1 has structure S_1 if '$S_1(M_1)$' is analytic. (Carnap and Bachmann (1981, 74), my emphasis)

This passage is highly instructive in helping to clarify the semantics of Carnap's approach. In the theory of models proposed here, each model comes with a single structure that specifies a relational architecture, i.e. the *arity* and type structure, shared by its isomorphism class. This structure of a model M_1 of type n can itself be expressed as a relation $S(M)$ of type $n+1$.[119] I will return to this in Section 4.

---

[117] A similar remark on the defining features of relations can be found in Carnap's *Abriss*: "A property is characterized as a structural property if it belongs to any relation *R* it also belongs to all relations of the same structure, that is to all relations isomorphic to *R*. Expressed in terms of classes: α presents a structural property if every relation belonging the a relation isomorphic to α also belongs to α (...)." (Carnap 1929, 54)
[118] Compare also Bonk and Mosterin (2000) on this point.
[119] This conception of a model's structure is closely captured by Corcoran's more general remarks on the historical notion of a structure as opposed to its modern understanding *as* a model: „From a philosophical and historical point of view it is unfortunate to that the term 'mathematical structure' is coming to be used as a synonym for 'mathematical system'. In the earlier usage (…) two mathematical systems having totally distinct

Given these preliminaries, we can turn to Carnap and Bachmann's theory of 'extremal structures' of an AS that lay the proper ground for their explication of minimal and maximal axioms. They are also, as I will argue, the root of Hintikka's misinterpretation.

2.2.2 Extremal structures

We have seen that for Carnap in *Untersuchungen* a model's structure characterizes a model in the sense that it fixes its isomorphism class. Accordingly, an axiom system has a 'structure number' referring to the numbers of different structures it defines. Each of these structures fixes one isomorphism class of models. An axiom system with only one structure is categorical ('*monomorphic*'), one with more than one structure non-categorical ('*polymorphic*') (see Carnap 2000, 128).

In Carnap and Bachmann this idea of structures possessed by an axiomatic theory is further developed.[120] Axiom systems are conceived here again as propositional functions can "possess a certain number of structures" that characterize their possible models.[121] A segment T_{SF} of this class of structures can be classified by the relation of a 'proper substructure' holding between set of structures of the system. A structure *S* is a 'proper substructure' of a structure *T*, if *S* and *T* are not identical and there are at least two models *M* and *N*, where *M* is a 'proper part' of *N* and *M* satisfies *S* and *N* satisfies *T*. Put in Carnap and Bachmann's formal terms, *S* is a proper substructure of *T* if the two conditions hold (ibid, 75):

(i) $S \neq T$

(ii) $(\exists M)(\exists N)\ (S(M).T(N).M \subset N)$

It is important to emphasize they way in which Carnap and Bachmann address this issue. They speak of a model's structures and substructures and not, as one would expect, of models and submodels of an axiom system. Also, the 'extremal conditions' imposed by minimal and maximal axioms are in the following cashed out in terms of structures, not in terms of the models themselves. Given the explications above, a 'structural diagram' of an axiom system

---

elements can have the same structure. This in this sense a structure is not a mathematical system rather a structure is a "property" that can be shared by individual mathematical systems. At any rate a structure is a higher order entity. (…) For mathematical purposes it would be possible to 'identify' a structure with the class of mathematical systems having that structure, but such 'identification' may tend to distort one's conceptual grasp of the ideas involved." (Corcoran 1980, 188) For a history of the term 'structure' in modern mathematics and logic see also Mancosu (2006).

[120] In this section I am following closely Section 3 of Carnap and Bachmann (1936).

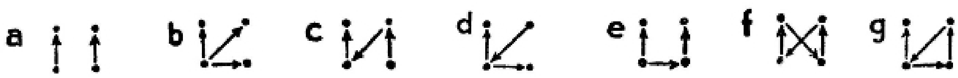
[121] In (RC 081-01-05), Carnap's assumed relation between the models of a theory and the structures described by it is further specified in this way: "Given that a formal AS is satisfied by every model of a structure in case it is satisfied by one model of the structure, we say in this case: the AS is "satisfied" by the structure." (ibid.)

can be defined as the relation T_{SF} of a "*proper substructure restricted to the structures belonging to 'F(M)'* " (ibid, 75).[122] S is termed a '*beginning structure*' of *F* if it belongs to the domain of T_{SF}. *S* is termed an '*end structure*' of *F*, if it belongs to the range of T_{SF}. Finally, *S* is called an '*isolated structure*' if it belongs to *F*, but not to T_{SF} (see ibid, 76).

Note that these definitions are closely related to the theory of relations expressed in Russell and Whitehead's *Principia Mathematica* as well as Carnap's own more introductory presentation in *Abriss der Logistik*.[123] The main innovation in the 1936 paper is its application to formal models with the aim to classify different limiting types of model structures defined by an AS: '*minimal structures*' consist of the beginning- and isolated structures; '*maximal structures*' are the end- and isolated structures of an axiom system (see ibid, 76).

Carnap & Bachmann present a simple example of an axiom system in order to illustrate how these different types of structures and their mutual relations are to be understood. Presumably, the example is also intended to outline how the above given specifications concerning a model's structure are to be related to the type of minimal and maximal conditions that extremal axioms impose on the models of a theory. As I will show in Section 2.5, in this second role it is to a certain extent misguiding.

The axiom system f(*R*) in question consist of four axioms with one basic binary relational variable *R*. Axioms (1) to (3) define the properties of injectivity, asymmetry, and non-reflexivity of *R*. Axiom (4) states that the field of *R* contains exactly four elements.[124] A set of topological "arrow diagrams" illustrate the different kinds of model structures the system captures:[125]



(ibid, 76)

Obviously, each diagram here presents not a specific model but a model's structure that is defined by f(*R*). Each of these individual structures can be interpreted differently, by assigning models with different domains of four elements. What is also immediately evident

[122] An essentially identical approach can be found in the notes for Part 2 of *Untersuchungen* in Carnap's *Nachlass* (RC 081-01-05).

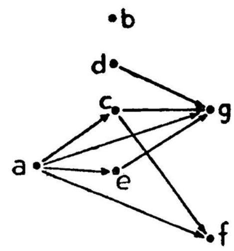
[123] See esp. (ibid, 35-38).

[124] In Part 2 of *Untersuchungen*, cardinality axioms such as this are termed 'absolute existence axioms' that fix the '*Urklassen*' of a theory (RC 081-01-10).

[125] The diagrams are to be understood in the sense that the (cardinality of the) field of a relation *R* (fld *R*) is presented by the points, its domain (dom *R*) by the starting points of the arrows and its range (ran *R*) by their arrowheads. Compare Behmann (1927, 42). See also Carnap (1929, 26-28).

from this illustration is that the different models satisfying these possible structures of the axiom system are not isomorphic to each other, i.e. f(R) does not capture its models up to isomorphism.

The crucial point to see, however, is that Carnap and Bachmann seem to intend the sequence of structures (a) to (g) to illustrate their account of model extensions (and restrictions respectively). A possible model extension for a given model M_1 of the theory (say, of structure (e)) is prima facie conceived as the extension to a model M_2 with one of the succeeding structure (say (g)). Obviously, in this case, no extension of the domain of individuals of M_1 to M_2 is involved. This idea strongly suggests a kind of FDC concerning the models of a theory. This is further strengthened by looking at how the different extremal conditions of axiom systems described above are to be explained via this sequence of structures. Given the topological arrow diagrams above, one can classify the types of relations that are intended to satisfy the extremal conditions of f(R). Thus, Carnap & Bachmann develop a "structure diagram" of the axiom system that allows classifying the maximal and minimal structures of the theory:



(ibid, 76)

The structures (a) and (d) in the graph are beginning structures, g and f are end structures and b is an isolated structure. Therefore (a), (b), and (d) are minimal structures and (b), (f) and (g) are maximal structures respectively (see ibid, 76). Carnap's theory can be further illustrated by building a number of simple models M_a, M_b, … M_g of f(R) that share an individual domain of four natural numbers, say A= {1, 2, 3, 4}, but assign different valuations to the binary relation variable R over A^2 so that $R_a = \langle(1,2); (3,4)\rangle$; $R_b = \langle(1,3); (1,2); (1,4)\rangle$; …; $R_g = \langle(1,3); (4,1); (1,2); (2,4)\rangle$. The way the relations (and therefore the models) can be clustered is by drawing to the notion of a partial relation: e.g. R_a is a partial relation of R_c, R_e, R_f and R_g. Now (g) expresses a maximal structure due to fact that it cannot be extended by adding ordered pairs from A^2 (like (1,4)) without violating one of the axioms (in this case, the axioms 1 and 2). Similar holds for the structures (b) and (f). In contrast, (a) presents a minimal structure because there exists no partial relation of M_a that satisfies the axiom system. The same holds for structures (b) and (d). Finally, b is an isolated structure since there neither exists a partial relation of R_b nor is R_b a partial relation of a more comprehensive relation that satisfies f(R).

2.3 Hintikka's interpretation

Hintikka's critical discussion of Carnap and Bachmann (1936) sets in at this point. As I have mentioned above, it essentially consists of two interpretive claims. They can be paraphrased in this way (see Hintikka 1991, 332):

(1) The specific formalized version of extremal structures in the 1936 paper is nonstandard compared to mathematical practice in the sense that it does not capture the informal conceptions of limiting conditions on a theory's domain expressed in the corresponding axioms by Hilbert, Peano and Fraenkel.
(2) This misconception on Carnap & Bachmann's side is a direct consequence of the tacit, but guiding "one domain assumption" in Carnap's conception of semantics.

Due to his assumed FDC - Hintikka argues – Carnap was forced to conceive the idea of 'extremal assumptions' in a fundamentally different way than they were originally codified in the mathematical completeness and restriction axioms. Hintikka puts the point as follows:

> Because of this one-domain assumption, Carnap and Bachmann had to reinterpret tacitly the entire problem so as to speak of the extremality of a model with respect to properties and relations, not with respect to individuals, as Hilbert among others, had considered the problem. The result was less than magnificent (...). Indeed, the only really interesting thing about Carnap's ill-fated paper seems to be its reliance on the one-domain assumption. (Hintikka 1992, 176)

According to Hintikka, the focal point of this misconception lies in the Carnap's treatment of the extension of a model that is not understood as the extension of the domain of individuals but as the "structural" extension of the primitive relations over a fixed domain of individuals. Compare again Hintikka on this point:

> Instead of asking whether a given model can be extended by introducing new individuals while the axioms stay satisfied, Carnap and Bachmann are asking whether a given model can be extended by assuming the relations involved in the axiom system to hold between more individuals among the members of the given fixed domain. (Hintikka 1991, 332)

Now, if understood in this sense, Carnap and Bachmann undoubtedly had misconceived the mathematical intentions behind the different extremal axioms they attempted to formalize. By looking at the way the minimal and maximal conditions are defined above via structure extensions, this claim also seems to be highly plausible. The crucial fact in the simple example described above is that the individual domain remains constant for all model extensions. In contrast to the standard account, the cardinality of the domain of individuals is axiomatically fixed (by axiom (4)) for all possible models of f(*R*). Interpretations of f(*R*) differ only in the structure of their relations. If one compares this e.g. to Hilbert's original formulation of the limiting condition expressed in his AC for geometry in *Grundlagen der Geometrie* the conceptual difference becomes obvious:[126] Whereas the extensibility in the completeness condition expressed here concerns the level of individuals of the theory in question, maximality in Carnap and Bachmann is defined as the non-extensibility of the axiomatically defined primitive relation *R* over a fixed domain of individuals.

However, as a will suggest in the remainder of this chapter, this is not a valid interpretation of Carnap's theory, despite the *prima facie* plausibility it gains from Carnap and Bachmann's example. To show this, it will prove worthwhile to further refine Hintikka's interpretive claims. In fact, there seems to be a sub-claim involved in (2) that can be put this way:

(2i) A FDC (or the "one-domain assumption") concerning the interpretation of a formal language implies a FDC concerning the models of an AS formulated in that language.

This is in fact a perfectly natural consequence of (2) if one assumes a standard conception of models, in which the interpretation of the language's quantifiers, i.e. the range of the bound variables, is actually part and parcel of the models. However, as I will show in the next section, this is not a sound assessment concerning Carnap's understanding of a model. So, to anticipate my argument: Hintikka interpretation fails since he misses to take into proper account Carnap's underlying heterodox conception of models. Where he is wrong is his assumption that *for* Carnap, a "one domain assumption" concerning a language necessarily involves a FDC concerning the models of a theory. With (2i) shown to be an inadequate assumption, Hintikka's claim (1) will also prove to be untenable.

[126] In the original version, the axiom reads: "It is not possible to add new elements to a system of points, straight lines, and planes in such a way that the system thus generalized will form a new geometry obeying all the five groups of axioms. In other words, the elements of geometry form a system which is incapable of being extended, provided that we regard the five groups of axioms as valid." (Hilbert 1902, 25) A structurally similar axiom of completeness specifying the non-extensibility of the system of real numbers can be found in Hilbert's axiomatization of the real number field in Hilbert (1900). Compare Awodey and Reck (2002).

2.4 A heterodox conception of models

What is Carnap's early conception of models in *Untersuchungen* and Carnap and Bachmann (1936)? Does it allow domain variation in some sense comparable to the modern understanding? Finally, how are models related to the interpretation of a given logical language, in which an axiomatic theory is formulated? In order to address these interpretive questions, let us start with Carnap's most detailed explication of formal models in the first part of *Untersuchungen*. In section 2.3 he introduces the notion in this way:[127]

If we write for f(R, S, T), in short f(R), and if R_1, S_1, T_1 are specific, e.g. arithmetic relations that are admissible values of the variables R, S, T, then we can also introduce an abbreviatory notation for the ordered system of relations R_1, S_1, T_1, say R_1. The *model* R_1 is then *a value of the model variable R*. Each of the relation variables R, S, T of the axiom system has a specific type (…); the sentential function for f(R, S, T) only yields a true or false sentence if the values of these types are inserted. If the three relations R_1, S_1, T_1 are admissible values of the variables R, S, T, i.e. if they can adopt the type of the three variables R, S, T (by declaring a specific domain of discourse in them to be the domain of individuals), we name the system of relations (R_1, S_1, T_1), or short R_1 an "*admissible model*" of fR, irrespective of whether R_1 satisfies fR or not; fR_1 is then in any case meaningful, i.e. either a true or a false sentence. An admissible model R_1 of fR is only then also a model of fR if fR_1 is not only meaningful, but also true. (Carnap 2000, 95)[128]

In Carnap 1930 the following shortened definition for models is given:

If fR is satisfied by the constant R_1, where R_1 is an abbreviation of a system of relations P_1, Q_1, …; R_1 is called a "model" of f. A model f(R_1) is a system of concepts of the basic system, generally a system of numbers (number classes, relations and so forth). (Carnap 1930, 303).[129]

Note that a central feature of modern model theory, the domain variability of a model, is not explicitly specified in Carnap's explications. Models are defined as "systems of numbers (number classes, relations and so forth)", i.e. as sequences of the form $\langle R, S, T, \ldots\rangle$, where

---

[127] See also Chapter 1, Section 2.3 for further discussion of Carnap's notion of a 'formal model'.
[128] See fn.36.
[129] "Wird fR durch die Konstante R_1 befriedigt, wobei R_1 Abkürzung für ein System von Relationen P_1, Q_1,.. ist, so heißt R_1 „Modell" von f. Ein Modell ist ein System von Begriffen der Grunddisziplin, meist ein System von Zahlen (Zahlklassen, Relationen u.dergl.)." (Carnap 1930, 303)

R, S, T, … stand for classes or relations of a given type. Apart from this rather intuitive qualification, no further indication is given with regards to his understanding of the domain of a model. In particular, unlike in the modern definitions of a model as a tuple $\langle D, I \rangle$, where D stands for the domain of the model and I for an interpretation function, there exists no notational sign indicating the domain of a specific model in the above definition.

Given this, the question arises whether of not Carnap conceived of something comparable to modern domain variability from his use of models in his theory of axiomatics. Intuitively, Carnap's overall motivation in *Untersuchungen* to provide a logical reconstruction of formal, i.e. reinterpretable axiom systems and its semantics alone seems to suggest that he aimed for a theory of models that allow VDC. Nevertheless, in the first part of *Untersuchungen* Carnap's position on the issue of domain variability is not fully conclusive. In fact, one can find prima facie conflicting evidence by looking at Carnap's remarks throughout the manuscript. On the one side, there is strong textual evidence pointing in the direction that Carnap actually intended domain variation in his treatment of models. Note first that in the definition above Carnap speaks of "declaring a *specific* domain of discourse in them [GS: the relations] to be the domain of individuals", which seems to indicate that the domain of individuals can be varied in a sense. Second, there are also several explicit remarks on the notion of "domain of discourse" (*Grundbereich*) in *Untersuchungen* pointing towards a VDC.

One is related to the specification of an isomorphism correlation between models by stipulating its "step number" or, which amounts to the same, the type-level of individuals assumed in a specific case. One of the ways this can be done is by indicating a domain of discourse to act as the domain of individuals for a specific model:

> The step number can in a specific case [be fixed] (…) by reference to a domain of discourse which will act as a domain of individuals for a model (…). (ibid, 120)[130]

One can thus in this way specify a domain of individuals for a specific model. It seems plausible that in the same way, one could also assign a different domain to this model.

This view finds additional support by considering to a second passage in which - in a slightly different context – the issue of a model's domain is discussed. In §2.11 Carnap introduces the notion of a "q-step model structure" and remarks that the structure of models with finite

[130] "Die Stufenzahl kann in einem bestimmten Falle (…) festgelegt sein: (…) durch Angabe eines Gegenstandsbereichs, der für ein Modell als Individuenbereich gelten soll (…)." (ibid, 120)

number of elements is presentable via a "q-step inventory list of a model" ("*q-stufigen Bestandsliste eines Modells*") in which all constants are substituted by variables (ibid. 121). Inversely, the resulting "schemes of inventory lists" ("*Bestandslistenschemata*") can be used for model building: one can get different models from interpreting, i.e. assigning *different* domains of discourse to them. The central passage here reads:

> (…) if a q-step model structure is given by an inventory list with variables, one can always specify a model which has this structure and whose individuals (relative to step q) belong to an arbitrary, stipulated domain of discourse by inserting arbitrary elements of the domain of individuals for the variables in the inventory formula. In short: (…) *if a finite (q-step) structure and a domain of individuals is given, one can specify a (q-step) model belonging to it*. (ibid, 122)

Carnap explicitly speaks of an "arbitrary, stipulated domain of discourse" here. This again seems to suggest VDC. Second, this remark on model construction in effect seems to come close to the modern conception of a model as assigning construction from a (non-empty) set of individuals to the expressions of a language. What is clearly missing is something like an interpretation function *I* that assigns elements and subsets of the domain to the non-logical constants of the language. This however, is insignificant for Carnap's account since he is working with a pure higher order language, i.e. a simple type theoretical language with an empty signature *L*, in his theory of axiomatics. Given this, models for an AS can be taken as assignments from a chosen domain of individuals to the variables of *L* or, in Carnap's own terms, the method of "inserting arbitrary elements of the domain of individuals for the variables" (ibid, 122).

Against this essentially modern reading of Carnap's conception of models (more precisely models' domains), several remarks in *Untersuchungen* can be mentioned that seem to indicate a FDC for models. They concern Carnap's use his logical language, more specifically the use of quantification in the formulation of axiom systems. One can find an insightful remark that is crossed out in the final version of the manuscript on an alternative to his version of formalizing axiom systems. It illustrates clearly Carnap's nonstandard conception of quantification involved in his project:

> A different conception on the nature of axiom systems which, however, does not seem acceptable to us, does not consider the basic relations of the axiom system, but the elements of these basic relations as the variables that transform the axiom system into a sentential function. The basic relations are

treated as constants. The elements do occur as variables, although *as bound variables* (...), those can also appear in a sentence since *they do not transform the axiom system into a sentential function*. (ibid, 89, my emphasis)[131]

The last sentence here is particularly insightful. In Carnap's own approach to formalize axiomatics theories are expressed as sentential functions and the primitive terms as free variables. What he argues in the passage above is this: If these variables were to be bound by quantifiers, then the free reinterpretability of the AS would be restricted to a fixed range, a fixed interpretation of the background language in use. The underlying assumption concerning the semantics of the language involved here becomes even clearer in an example of a simple transitivity axiom he gives for this alternative version:

I we express it with the help of the M-elements, the axiom is:

$$(x,y,z)\ ([M(x,y)\ \&\ M(y,z)] \rightarrow M(x,z))$$

Here x, y, z are not free variables, but bound by the universal operator in front. If M were now a constant, then the expression would be a sentence; it can only then be a sentential function, if M is a variable. (ibid, 89)[132]

This indicates that for Carnap, when formalizing an axiomatic theory, the logical language in use comes equipped with a (fixed) universe and not, as in modern semantics, allows changing its interpretations. Further textual evidence for such a fixed interpretation of the background logic, and thus a fixed range of the language's quantifiers can be found in a remark concerning the question of "decidability" (*Entscheidungsdefinitheit*) of the discussed systems:

Therefore, *not singular domains* are demarcated by the different axiom systems, but every deduction of any consequence of any axiom system is performed in the *one domain of logic*. (ibid, 149)[133]

[131] "Eine andere Auffassung über das Wesen der Axiomensysteme, die uns aber nicht zulässig erscheint, betrachtet nicht die Grundrelationen des Axiomensystems, sondern die Glieder dieser Grundrelationen als die Variablen, die das Axiomensystem zu einer Aussagefunktion machen, die Grundrelationen selbst aber als Konstante. Die Glieder treten allerdings als Variable auf, aber als gebundene Variable (…), solche aber können ja auch in einer Aussage vorkommen, sie machen also das Axiomensystem nicht zu einer Aussagefunktion." (ibid, 89)

[132] "Drücken wir die mit Hilfe der M-Glieder aus, so lautet das Axiom: $(x,y,z)\ ([M(x,y)\ \&\ M(y,z)] \rightarrow M(x,z))$. Hier sind x, y, z keine freien Variablen, sondern durch den vorn stehenden Alloperator gebunden. Wäre nun M eine Konstante so wäre der Ausdruck eine Aussage; er kann nur dann eine Aussagefunktion sein, wenn M eine Variable ist." (ibid. 89)

[133] "Daher werden durch die verschiedenen Axiomensysteme nicht einzelne Gebiete abgetrennt, sondern jede Ableitung irgendeiner Folgerung irgendeines Axiomensystems vollzieht sich in dem *einen Gebiet der Logik*." (ibid, 149)

Here again, what Carnap seems to be assuming in his theory of axiomatics is a universal, fully and invariantly interpreted background language, close to what Hintikka describes as the "one-domain thesis" in Carnap's thought.

Now, to be specific, the exact nature of the tension between remarks like this one and talk of "arbitrary, stipulated domain(s) of discourse" of a model cited above arises form lack of clarification on Carnap's side of the exact relationship between the semantics for axiomatic theories formalized *in* a language L and the semantics *for* L itself. Put differently, it concerns the way his theory of models is related to the interpretation of L. As we have seen, Carnap's remarks on this issue are not fully conclusive in the first part of *Untersuchungen*. Fortunately, Carnap and Bachmann, in their 1936 paper, get more explicit.

2.4.1 Models and analyticity

Recall the passage in Carnap and Bachmann (1936) quoted in Section 2.2.1, in particular the last sentence concerning the model's structure:

> The relations so determined we will call structures and will say that model M_1 has structure S_1 if 'S_1 (M_1)' is analytic. (Carnap and Bachmann (1981, 74), my emphasis)

The fact that the notions of structure and model are treated in a highly similar fashion in Carnap's 1928 manuscript and in 1936 should not lead one to overlook the substantial differences in their explication. The passage above witnesses a significant shift in Carnap's underlying conception from what has been termed a *Russellian* "monolinguistic" approach in 1928 (see Coffa (1991)) to an explicit metatheoretic stance formulated in *Logical Syntax of Language* (in the following *LSL*) (Carnap (1934)). Put differently, the conceptual framework within which Carnap's theory of models was developed had changed, by 1936, from a universalist approach to that of *LSL*. This new horizon is most evident in his above use of the metatheoretic notion of *analyticity* to state that a model M_1 satisfies a given structure: a "model M_1 has structure S_1 if '$S_1(M_1)$' is analytic." Similarly, models themselves are now defined *via* the notion of analyticity:

> Let 'M_1' be an abbreviation for a sequence of constants of the language S. We say that M_1 is a model of the axiom system '$F_1(M)$' if the sentence '$F_1(M_1)$' is analytic in S. (ibid, 67)

Compare this to the corresponding segment in Carnap's earlier explication of an "admissible model" of an axiom system in the 1928 manuscript: "An admissible model R_1 of fR is only then also a model of fR if fR_1 is *not only meaningful, but also true*." (Carnap 2000, 95, my emphasis). In Carnap (1930), the concept of truth used above is substituted by that of 'satisfaction': "If fR is satisfied by the constant R_1, where R_1 is an abbreviation of a system of relations P_1, Q_1, …; R_1 is called a "model" of f." (Carnap 1930, 303)

Now, the nominal difference between the 1936 definition and those in *Untersuchungen* seems to concern only the substitution of the terms "truth" (and "satisfaction" respectively) with that of "analytic in S" (viz. in a type-theoretic language like LII defined in *LSL*). Nonetheless, this conceptual background of the respective definitions makes all the difference. First, note that both concepts - truth and satisfaction - used in the earlier definitions are used in a strictly informal sense and thus in no sense comparable to Tarski's formally explicated versions given in (Tarski 1935). No further remarks can be found in *Untersuchungen* on how the notion of truth is understood or what conditions have to be met in order to speak of a theory to be true in a model. Carnap, *anno* 1928, simply does not have something corresponding to a formal concept of *truth in a model*.[134] In contrast, the explications of a model and a model's structure given in Carnap & Bachmann (1936) are based on the formal definition of the notion of 'analytic for LII' first published in a side paper to *LSL* (see Carnap (1935)).[135] With this in hand, Carnap's background semantic apparatus used in his definition of models in 1936 is essentially identical with Tarski's recursively defined notions of truth and satisfaction.[136]

Second, with the shift to the conceptual framework of *LSL*, there is also a change in the explication of the intended semantics of the background languages in use. In *Untersuchungen*, Carnap is not fully conclusive on the intended interpretation of the 'contentful basic system' ('*inhaltliche Grunddisziplin*'). Given his overall logicist program at that time and the two passages quoted in Section 4, one could argue that he is assuming a universal and fully interpreted logical language with quantifiers ranging over a *fixed* but *unspecified* universe of type theoretic objects (but see Section 4.3). In contrast, the ranges of the quantifiers in Carnap's languages LI and LII studied in *LSL* are stipulated explicitly.[137] It is crucial to note that here, as in 1928, Carnap's conception of the semantics of a logical language like LII is

[134] There is some evidence that truth in the axiomatic context is understood here by Carnap as "being tautological". Compare Bonk and Mosterin (2000, 38) on this point.

[135] In a footnote in the 1936 paper, Carnap and Bachmann explicitly refer to Carnap (1935) for the formal definition of 'analytic in II' used in their explication of models (see ibid, 84).

[136] For a closer comparison with Tarski's semantic notions see Coffa (1991, 288-293).

[137] See Carnap (2002, §6 and §26).

not comparable to a modern model-theoretic account, where quantifiers are subject to free reinterpretation (on a given type level). As Lavers stresses in a recent paper, a pure type-theoretic language like LII in *LSL* is used not as a formal language in the modern sense, but as a "meaningful formalism", i.e. a language with a fixed intended interpretation for each type level (see Lavers 2008, 10). Thus, there is a clear continuity between 1928 and 1936: logical languages are fully interpreted and the interpretation is fixed.

The real difference to *Untersuchungen*, however, is that where Carnap was most likely operating with a single universal language ranging over a universal domain of typed objects, there is now a move to a logical pluralism concerning the choice of formal languages *and* their specific interpretations. The intended semantics for a language like LII is still invariant, however not longer universal. Quantification in a language can be freely stipulated to range over specific mathematical systems and not the type-theoretical universe. For instance, the individual quantifiers of LII are set to range over "a domain isomorphic to the natural numbers" (Lavers 2008, 10). In contrast, in Carnap and Bachmann 1936, one can find the following remark on a (syntactically) similar language:

> We think of a language [GS: like LII], excepts that the domain of individual contains the natural numbers and also the points of a line of Euclidian geometry among which two points are distinguished. (Carnap & Bachmann 1981, 71)

Unlike in the 'basic system' in 1928 the individual variables here do not range over *all logical objects* of type 0, but instead over a stipulated classes of mathematical objects, namely natural numbers and a denumerable set of points. Similar holds for the higher-order variables in LII.

Having underlined these transitions in Carnap's approach – i.e. to shift from an informal notion of truth to a formal notion of analyticity (in LII) on the one hand, a full explication of the intended interpretation of LII on the other – we are finally in a position to better understand the exact nature of the connection between the semantics for a language and his conception of models. Put differently, the main effect of this shift is a new explicitness of the intrinsic relation between the domain of models and the semantics of the background language in use. Given the conceptual framework of *LSL* the link between models and truth conditions for a formula in a given interpretation becomes transparent. The dependency on the language's interpretation gets evident when considering the application of the truth condition expressed in Carnap's account of analyticity for a type-theoretic language like LII to his definition of models.[138]

To illustrate this, consider a simple example of an axiom system f(M), consisting of just one transitivity axiom (borrowed from Carnap's example above). The axioms (implicitly) define a single primitive term, expressed as a binary first-level relation variable M:

(1) $$(\forall x)(\forall y)(\forall z)\ ([M(x,y)\ \&\ M(y,z)] \rightarrow M(x,z))$$

A possible model of f(M) consists of the binary relation M_1. If M_1 is substituted for the free variable M then the interpreted axiom system f(M_1) turns into the closed formula

(2) $$(\forall x)(\forall y)(\forall z)\ ([M_1(x,y)\ \&\ M_1(y,z)] \rightarrow M_1(x,z))$$

that according to §43 in *LSL*, is analytic in LII iff

(3) $$[M_1(a,b)\ \&\ M_1(b,c)] \rightarrow M_1(a,c)$$

is analytic in LII for all numerals a, b and c in the substitution class of LII (in short D_{LII}). Similar truth conditions hold for cases in which the formulation of an axiom involves higher order quantification over relations or properties. Take for instance the second order version of the *axiom of induction* for arithmetic as g(z,S) and an interpreted version with a model g(S_1) as

(4) $$(\forall X)([X(z) \wedge (\forall x)(X(x) \rightarrow X(S_1(x)))] \rightarrow (\forall y)(X(y)))$$

where S_1 is an interpretation of the function variable S. Then g(0, S_1) is analytic in LII iff

(5) $$[X(0) \wedge (\forall x)(X(x) \rightarrow X(S_1(x))] \rightarrow \forall y(X(y)))$$

is analytic in LII for all possible assignments (in Carnap's terms "valuations") of relations on D_{LII}, i.e. all subsets of D^2 to the predicate variable X (assuming given substitutions for x and y). The main difference to the case of bound individual variables is that Carnap treats higher-order quantification extensionally in *LSL*. Instead of being restricted to a substitution class of predicate symbols in LII, the quantifier '($\forall X$) …' is supposed to range over all relations or properties of the set of numerals d of LII, thus e.g. over all subsets of D^n for an n-ary relation variable.[139]

Two points can be made here. First, this account of specifying the truth of an AS in a model in the (proto-) semantic framework of *LSL* seems irreconcilable with a current understanding of model-theoretic truth. In particular, note that in contrast to the modern account of *truth in a structure*, the question whether an axiomatic theory is true in a model is not reduced to truth

[138] For Carnap's detailed definition of 'analytic in LII' see Carnap (2002, §34).

[139] For further historical details concerning Carnap's move to an extensional treatment of higher-order quantification see Coffa (1991) and Awodey and Carus (2001).

conditions *relative to that model* here. Instead, it is conceptually linked to the fixed substitution class of numerals D_{LII} of LII, thus to the intended interpretation of the language. Instead of specifying truth conditions for formulas via the assignment of elements of and relations on a variable model's domain of discourse to the variables of the formal language Carnap states conditions of analyticity based on assignments from a fixed interpretation of the language.

Second, given the above examples, it is *prima facie* not clear how the domains of different models are related to the range of individuals of the language. Several interpretive options to understand this are possible: (a) In the first, all models' domains are supposed to coincide with D_{LII}. In this case, the result is a FDC concerning models suggested by Hintikka. Against this, one can hold that there is no reason intrinsic in Carnap's explication to assume that D_{LII} necessarily has to coincide with the domains of M_1, M_2, etc. for an axiom like (1). Moreover, such an understanding would be in conflict with Carnap's remarks cited in §4 and more generally with Carnap's overall attempt to reconstruct (Hilbertian) formal axiomatics plus its informal model theory.

(b) The second option would assume that with each model, the range of quantification is reinterpreted with the model's domain. Thus, with each new assignment to M in (1) the range of x, y, and z is interpreted to the domain of M. The earliest account of this – genuinely modern approach I found formulated in the first volume of Hilbert and Bernays' *Grundlagen der Mathematik*, published in the same year as Carnap's *LSL* (Hilbert and Bernays 1934). Here, Hilbert and Bernays give a presentation of an axiom system of Euclidian geometry with two predicates R (for *greater than*) and S (for *between*). The axiom system is expressed in a strikingly similar sense to Carnap's version as the propositional function $\mathbf{U}(R, S)$ and a model accordingly as a certain tuple $\langle R_1, S_1\rangle$. The central point for our discussion is a passage in which the specification of the models domain is explained:

> It should be noted that together with the assignment of the predicates one also has to fix the domain of individuals, on which the variables x, y, … are to refer. It is in a way introduced in the logical formula as a *hidden variable*. (ibid, 8-9)[140]

[140] "Zu beachten ist, dass man zugleich mit der Bestimmung der Prädikate auch den Individuenbereich festzulegen hat, auf den sich die Variablen x, y, … beziehen sollen. Diese geht gewissermaßen als *versteckte Variable* in die logische Formel ein." (ibid, 8-9)

Such a stipulation of the range of individuals with the 'assignment' ('*Bestimmung*') of the predicates *R* and *S* clearly implies a VDC concerning the individual domains of the interpretations of **U**(*R*, *S*). At the same time, it also implies a VDC for the language itself. In fact, models are here explicitly understood as the combination of the "individual domain" and the chosen 'value-ranges' ('*Wertverläufen*') of the primitive predicates, based on which "the satisfaction of the axioms" ("*das Erfülltsein der Axiome*") can be shown (see ibid, 12).

Note that this is not what Carnap has in mind in 1936. In fact, such an understanding is simply not consistent with Carnap's own definition where *analytic* means *analytic in* LII. The conditions for analyticity of a formula are explicitly linked to the interpretations of the language, in the case of individual variables to D_{LII}. Thus, any attempt to understand (3) (or (5)) as specifying the truth of an AS in a given model in something comparable to the modern sense will fail since the conditions for "analytic in LII" require the preservation of truth over all possible assignments from D_{LII} and *not* over a specific (but principally variable) model domain.

The crucial interpretive question we are thus confronted with is this: How can Carnap's account of truth conditions for a type-theoretic language like LII with a fixed, intended interpretation be reconciled with Carnap's scattered remarks in *Untersuchungen* and in Carnap and Bachmann (1936) cited in Section 2.4.1 where something comparable to a VDC for models is explicitly indicated? I suggest that the key to solving this tension lies in the acknowledgement of the genuinely heterodox character of Carnap's conception of models involved in his theory. It is to see where exactly his account differs from the modern architecture of models.

2.4.2 Domains as fields

Carnap's definition of models *via* analyticity given in 1936 seems to involve two separate semantic systems: the intended interpretation of the given language on the one hand, the models for a theory expressed that language on the other. This renders impossible the modern account in which the interpretation of the language, i.e. the range of its quantifiers, is intrinsic or "built into" the conceptual architecture of a model.[141] Note, however, that the apparent tension arising from this can be resolved once the domain of discourse of a model is understood as independent from the truth conditions defined above. More precisely, if the

[141] Compare Bays (2001) for a similar remark on Tarski's early conception of models.

domain is not assumed – as in modern understanding - to be identical with the range of the individual variables of LII but as a subset of D_{LII} gained by effectively restricting its quantifiers of LII. How does Carnap conceive such a restriction of the range of the (bound) variables to the domain of a specific model expressed in this language? To my understanding, the most plausible interpretation of his underlying conception is also the simplest one: consider again the above example of a transitivity axiom:

$$(\forall x)(\forall y)(\forall z)\ ([M(x,y)\ \&\ M(y,z)] \rightarrow M(x,z)).$$

What specifies the individual domain of an admissible model of this axiom is not the fixed intended interpretation of the language but the different assignments to the free variable M themselves, i.e. the relations M_1, M_2, etc. substituted for M. Where Carnap speaks of a model domain he most likely thinks of the field of the possible relation assigned to M. More generally, a specific sequence of relations $M_1 = \langle R_1, S_1, T_1, \ldots \rangle$ that satisfies an axiom system thus simultaneously fixes the domain of the theory. It simply consists of the union of the respective fields of R_1, S_1, T_1, …. Note that this conception of model domains is still closely linked to the semantics of the background language in use. It follows from Carnap's explication of models via analyticity that all (first level) relations of a sequence constituting a model are defined over the individual class of the language. In our example, any domain of a model for M is thus a certain segment, a subset of D_{LII}. Thus, while the range of the quantifiers in LII if fixed to D_{LII}, a specific assignment to the model variable M in fact restricts their effective range to the domain of the given model. For the subsequent discussion, I will dub this the 'domain-as-fields' view.

There exists extensive textual evidence in Carnap's writings on axiomatics in support of the view that he actually conceived models and model domains in this way. In Carnap and Bachmann (1936), in a passage on the 'complete isomorphism' between two relations, the authors mention the "fields of their field elements (i.e., for *certain segments of the individual domain of the language*)" (ibid, 73, my emphasis). A few lines further below, in the passage already cited in Section 2.2.1, concerning the isomorphism between "*n*-place models" it is stated that

(…) *n*-place relations can be defined over the field of this [GS: isomorphism] relation, i.e., *over the totality of constants of the basic, assumed language which can appear as elements of models* (…). (ibid, 74, my emphasis)

It is clear from this that Carnap understands a model's individual domain as a specific subset of the "totality of constants" of the specified background language.[142] Given this substitutional approach, models with different domains are constructed by carving out segments of the substitution class of individual constants, i.e. mostly numerals of the language. The specific domain of a model is here again conceived as union of the fields of individuals of the primitive relations constituting the model.

This identification of domain of discourse of a model with the (union of) field(s) is articulated more explicitly in a separate, but closely related work on axiomatics by Carnap's collaborator Bachmann (Bachmann 1934). In his inaugural dissertation on the logical foundations of arithmetic, Bachmann presents different possible axiomatizations of "elementary arithmetic" in a many-sorted language and a detailed study of their respective logical relations.[143] He discusses different axiom systems of the form f(x R α) with three free variables (for individuals, relations and classes respectively) that stand for the primitive terms of the theory (Bachmann 1934, 1). Models for the theory are defined here in a way similar to Carnap's account in *Untersuchungen*: "Each triple of values that transforms the axioms in true sentences when inserted for x R α is termed a *model of arithmetic*." (ibid, 1)[144] Note that - in contrast to the axiom system presented in Carnap and Bachmann (1936) – there exists a class variable α here that is conceived as an explicit notational sign for the domain of discourse, i.e. the number set of an arithmetical model of f(x R α). Moreover, according to Bachmann, α can be interpreted differently: as the "set of the natural numbers," as the "set of natural numbers greater then 37," as the "set of primes," etc. (see ibid, 4-5). Bachmann also discusses the question whether an axiom system of "basic arithmetic" f(R) can be constructed with just one primitive term, i.e. with one free variable that is "generally equivalent" to f(x R α).[145] The crucial point for our discussion is that he explicitly mentions a condition concerning the relation of R to α the models of that f(x R α) have to fulfill in order that the two theories become equivalent:

---

[142] A similar informal remark concerning the composition of models can be found in *Untersuchungen*: "(...) we speak in short of "models" of an axiom system und mean by this logical constants, i.e. "systems of concepts of the basic system" (and these are mostly systems of numbers)." (ibid, 94)

[143] This work, supervised by Heinrich Scholz, anteceded the collaboration with Carnap. Nevertheless, Bachmann was well acquainted with (published) Carnap's work on axiomatics. He refers to Carnap (1929) and Carnap (1930) repeatedly in the text (see e.g. Bachmann 1934, 47).

[144] "Jedes Wertetripel, dass in x R α eingesetzt die Axiome in wahre Aussagen überführt, heißt ein *Modell der Arithmetik*." (Bachmann 1934, 1)

[145] A discussion of the details of Bachmann's notion of 'general equivalence' of axiom systems would take us too far afield here.

The reason why the reduction to R as the single free variable was not possible so far lies in the fact that there were no sufficiently strong connections between R and α; for it was simply demanded by f(x R α) that α is contained in the field of R, not that α is identical with the field of R. We therefore add to out AS the condition F. $(\forall z)(\exists y)$ $(yRz \rightarrow z \in \alpha)$ (...). (ibid, 43, notation changed)[146]

Given this additional condition that any successor of a number is to be contained in the class α, Bachmann effectively shows that f(x R α) + F imposes that, in any model, the domain is identical with the field of the relation R. Given this extended AS, it is thus possible to provide a functional definition of α in terms of R (see ibid, 43-44). An additional primitive indicating the domain of a model besides R thus becomes redundant given that the domain is already stipulated with the assignment to R. A version of Bachmann's axiom system of "elementary arithmetic" with one variable R that defines a class of "progressions" will be discussed in Section 2.5.1. What is important here is to note that Bachmann's dissertation must be considered as the immediate background for the theory of axiomatics outlined in Carnap and Bachmann (1936). With this in mind, the remarks made in the passage above on the connection of α and R should be valued as direct evidence for the 'domain-as-fields' interpretation of the conception of models outlined above.

Decisive textual evidence in support of this view can also be presented from Carnap's own works. As mentioned in Section 2.4, Carnap is not fully explicit on his conception of models in the first part of *Untersuchungen*. Nevertheless, by drawing to Carnap's *Nachlass*, in particular to the manuscript documents for the projected second part of *Untersuchungen* it becomes clear that he already conceived models in essentially the same way in the late 1920s and thus years before the conceptual turn to *LSL*. In the two typed documents titled "Complete isomorphism" ("*Vollständige Isomorphie*") (RC 081-01-15) and "Domains of a model" ("*Bereiche eines Modells*") (RC 081-01-16), both dated from 1928, Carnap develops a clear and detailed discussion of the auxiliary notions underlying his theory of models. Carnap's method of explication here is to first describe the different domains of a given, *disinterpreted* axiom system in terms of a 'domain analysis' ('*Bereichsanalyse*'). Based on this purely structural description, the corresponding domains of a model are then stipulated in

[146] "Dass die Reduktion auf R als einzige freie Variable bisher nicht möglich war, liegt daran, dass zwischen R und α nicht genügend starke Bindungen bestanden; denn es wurde in f(x R α) nur verlangt, dass α im Feld von R enthalten ist, nicht dass α mit dem Feld von R identisch ist. Wir fügen daher zu unserem AS die Forderung F. $(\forall z)(\exists y)$ $(yRz \rightarrow z \in \alpha)$ hinzu (…)." (ibid, 43, notation changed)

a second, independent step. Briefly, the 'domain analysis' can be understood as a way to specify the type structure of the theory beginning with the relations of highest type and descending the relational hierarchy defined by an system until reaching the elements of type level 0. One thus starts with the 'base domains', i.e. the different domains of the primitive relations *R*, *S*, *T*, etc. defined by the theory. In case these base domains can be further reduced to domains of a lower type level, then the analysis proceeds until one reaches the non-reducible 'individual domains' of type level 0 of an axiom system (RC 081-01-15). An individual domain of a system is thus defined as the "domain from which no other domains descend" (ibid.). Given this, a "family tree of domains of an AS" ("*Stammbaum der Bereiche des* AS") is constructed by assigning each domain in the AS a fixed "type expression" ("*Typenausdruck*"). Note that up to this point, models are not mentioned. The domains of an axiom system are understood in a purely structural way so far. The analysis of domains concerns exclusively the type-theoretic structure defined by an axiom system.

This changes, however, in the following section on "Bereiche eines Modells" (RC 081-016) where Carnap describes, in analogy to the "family tree of domains" of an axiom system f(…) the compositions of domains of a given model *M* for f(…). Here again, models are conceived as the relations substituted for the model variables of f(…).[147] Carnap specifies the point in this way:

> Let *M* be an admissible model of the AS f(...). *M* assigns to each primitive term of f a logical constant; in order that this is possible, there always has to be an *n*-ary relation corresponding to each *n*-ary relation. It follows that *exactly one set in M* is assigned to each domain of f as its "corresponding domain". (ibid, my emphasis)[148]

The central notion here – not mentioned in Part 1 of *Untersuchungen* – is that of 'corresponding domains' of a model. The different domains (of different type levels) of a model are treated in strict analogy with the structural domains of the theory specified before. Here, however, as the passage above shows, they are treated extensionally, i.e. as 'classes' of (tuples of) individuals, relations on classes of individual, relations on relations on classes of individuals etc.. The relevant point for our discussion is the way in which Carnap understands

[147] Perhaps the most explicit statement of this understanding can be found in a remark where Carnap speaks of "the values *R* (*the models*)" of an AS f(*R*) (RC 081-01-12, my emphasis).

[148] "*M* sei ein zulässiges Modell des AS f(…). *M* ordnet also jedem Grundbegriff von f eine logische Konstante zu; damit dies möglich ist, muss einer *n*-gliedrigen Relation auch stets eine *n*-gliedrige Relation entsprechen. Daraus folgt, dass jedem Bereich von f genau *eine Klasse in M als "entsprechender Bereich"* zugeordnet ist." (ibid, 105 my emphasis)

the ‘individual domains’ of a model in this context. In a footnote in this section he makes the following insightful remark:

k_1 in M_1 is called an “individual domain“ in relation to AS if the corresponding domain k of the AS is an “individual domain“ or, in case k_1 corresponds to several domains k, k’, ... of the AS: if all these are individual domains of the AS. (ibid.)[149]

Thus, the individual domain*s* (note the plural) of a model are understood as the (possibly) different sets of individuals that can occur in a model. It is but a small step from this to consider *the* domain of individuals of a given model as the union of these sets of individuals. In fact, Carnap at several places in his manuscript explicitly speaks of the ‘union’ (‘*Vereinigung*’) of ‘corresponding domains’ of an axiom system and – on the level of extensional relations of the domains as “unions of equi-level relation domains” (“*Vereinigungen gleichliegender Relationsbereiche*”) (RC 081-01-18).

I consider these remarks in the projected second part of *Untersuchungen* in combination with those made in Carnap and Bachmann (1936), Bachmann (1934) to be conclusive evidence that Carnap, from 1928 onward, understood formal models of axiomatic theories as sequences of relations and a universe of a model as the union of the fields of these relations. This said, three points of commentary are in order here. First, what follows from this heterodox account of the domain-as-fields account is that – even if one assumes a fixed interpretation of the background language - models of an AS can have different, variable domains. Assuming this conception of formal models of an axiom system as the interpretations of its primitive terms, one can specify models with different domains in one language without violating the truth conditions codified in Carnap’s notion of analyticity. Since the relations’ fields stipulated as domains can be chosen freely as subclasses of a more comprehensive class of syntactical signs that form the individual constants of the *language in use*, Carnap’s approach is de facto structurally comparable to the modern variable domain conception of models. Thus, VDC for the models of an axiomatic theory is provided in Carnap’s account. What this amounts to is the possibility of assigning *extensionally different* relations M_1, M_2, etc. to the model variable M. Thus, if one understands an n-ary relation R as defined over D^n of LII and the extension of

[149] “k_1 in M_1 heisst ein “Individuumbereich” inbezug auf das AS, wenn der entsprechende Bereich k des AS ein „Individuumbereich“ [inbezug auf das AS] ist, oder falls dem k_1 mehrere Bereiche k, k’, … des AS entsprechen: wenn alle diese Individuumsbereiche des AS sind.“ (ibid.)

R as a set A of n-tuples the relation is true on, where $A \subseteq D^n$, then the domain of a model expressed by R consists simply of the set of individuals contained in A, i.e. the field of R. Within the limits of D_{LII} extensionally different relations can be formed as the models of an AS.[150]

Second, it should be noted here that variants of the 'domain-as-fields' view have recently been discussed extensively in an ongoing scholarly debate on the proper understanding of Tarski's notion of models in the 1930s.[151] In the light of the discussions there, several striking resemblances between Tarski's and Carnap' accounts become visible. In an instructive paper, Mancosu presents an interpretation of Tarski's understanding of "domains of discourse" of an axiomatic theory that is essentially identical to Carnap's approach (Mancosu (2006)).[152] Mancosu shows that at least for some formalizations of axiomatic theories given by Tarski, there is no prefixed "identity between the intended domain of the theory and the range of the quantifiers" which is antecedently fixed for the used background type-theoretic logic (ibid, 226). In order to allow the variation of the domain of discourse of a theory, the fixed range of quantification is in some cases effectively restricted by a non-logical domain-predicate additionally introduced to the language.[153] This convention used in Tarski's formal axiomatics certainly comes close to the suggested 'domains-as-fields' view for Carnap's theory. Nonetheless, one important difference to Tarski's approach concerning the method of logical formalization of axiomatic theories can be mentioned here. At least in some examples of Tarski's formalizations of a 'deductive theory', he is in fact operating with a logical language with a non-empty signature, i.e. a language including non-logical, schematically understood constants to express the primitive terms of an axiomatic theory. In Carnap's account, as we have seen, axioms systems are always expressed in a *pure* type-theoretic language without the use of non-logical constants. Here, the domain of a theory is fixed by the interpretation of the model variable, i.e. by the assignment of a specific extensionally understood relation (or a sequence of relations) to the model variable (or the variables expressing the disinterpreted

[150] For Carnap's notion of a relation's extension see Carnap (2000, 65-67) as well as Carnap (1929, 66).
[151] See Bays (2001), Gomez-Torrente (1996) and (2009), Mancosu (2006), Jané (2006) and the literature cited there for further references.
[152] In fact, Mancosu explicitly mentions – without going into specifics - "Carnap's conception of models" in *Untersuchungen* and describes it as "very similar to Tarski's". He further states that: "In general (…) one should devote a whole paper to the relationship between Carnap and Tarski on the notion of model and logical consequence." (Mancosu 2006, 215). At least regarding Carnap's notion of models, this suggestion is taken up in the present chapter.
[153] Compare also Bays (2001) for a similar interpretation of Tarski. For lack of space I cannot go into the subtle differences between Mancosu's and Bays' understanding of Tarski's domain conception, but see Gomez-Torrente (2009) for a detailed comparison.

primitive terms). Thus there is no quantifier relativization involved here via a non-logical domain-fixing predicate. However, as is also shown in, there are cases in Tarski's writings where he as well is operating with a *pure* simple theory of types as the background language for reconstructing axiomatic theories.[154] In these cases Tarski's method of domain specification is practically identical to the one suggested by Carnap. Compare Mancosu on Tarski's method:

(…) the 'universe of discourse' of the mathematical theory in question (…) will be given *by a class taken as the value of the variable corresponding to one of the primitives of the theory* (…) . (ibid, 224, my emphasis)

One could hold that there is one subtle difference compared to Carnap's account. In the latter's case, there is no specific single relation in sequence constituting a model, i.e. in $M = \langle R, S, T, \ldots \rangle$, explicitly reserved for specifying the domain of the theory. Thus, there is no functionally distinct variable used for relativizing the quantifiers in LII. However, in Mancosu (2006), at least in one example in Tarski's writings from the 1930s is discussed where a convention identical to Carnap's can be identified. In describing an axiom system defined in Tarski (1935a) with only one primitive binary relation *R*, Mancosu holds:

A model for such a theory is given by the finite sequence containing only the relation *R* (the 'universe of discourse' is implicitly defined by the field of *R*). (ibid, 224)

Given this, we can witness here a striking similarity between the two approaches to a logical formalization of axiomatic theories in 1920s and 1930s. Both authors seem to assume variants of the 'domain-as-fields' approach. This is particularly insightful given the fact that there seems to be no direct line of influence in either direction between Tarski and Carnap.[155] I take this conceptual similarity in dealing with formal models as an expression of a by then standard practice for the formalization of the semantics of axiomatic theories. Moreover, it clearly shows historical point that Carnap, even before Tarski's work in the 1930s (and, most importantly, Tarski's classical paper on models and logical consequence (Tarski (1936))), was

[154] Goméz-Torrente describes in detail "Tarski's pluralism", i.e. "his disposition to work within a variety of different frameworks" in his formalization of axiomatic theories (see ibid, 250-251).
[155] At least in the Carnap *Nachlass*, no documents can be found that would suggest an influence on Carnap from Tarski's side while working on *Untersuchungen*. This is also not surprising given the fact that the first points of personal contact between Carnap and Tarski took place in 1930. Compare Awodey and Carus (2001) on this point.

introducing a convention to treat formal models with variable domains in the axiomatic context.

Finally, note that this conception of 'domains-as-fields' is in Carnap's case perfectly consistent with Hintikka's sweeping assumption of a "one domain thesis" *concerning* the logical languages used by him. It can be consistently argued – along the lines of Hintikka - that for Carnap every "*formal language in use*" comes with a fixed interpretation and still hold that he conceived of the variability for his formal models. The reason for this is, as we have seen, that Carnap, in his theory of axiomatics, seems to draw a distinction between the range of the quantifiers of the 'contentual' background logic in which a theory is formalized and the domain of discourse of a theory captured in a specific model. Hintikka's attribution of the "one domain thesis" to Carnap's semantics for his language, his claim (2), thus has no stronger implications for his method of model construction. So, whereas it makes perfect sense to hold that Carnap's semantics involved in his axiomatic project is in crucial ways heterodox compared to the modern understanding of models it would be wrong to argue that he assumes a FDC for his models. Where Hintikka evaluation of Carnap's semantics misses track is in his conclusion that a "one domain assumption" concerning a language necessarily involves a FDC for the models in Carnap's theory of axiomatics. Given Carnap's heterodox treatment of models, claim (2i) that a FDC for the language in use implies a FDC for models shows to be a *non sequitur*.

2.4.3 VDC via type ascent

So far, our discussion has left unquestioned the assumption that Carnap, in his theory of axiomatics, is using fully interpreted background languages in which the range of quantification at each type level is invariantly fixed. This is essentially Hintikka's claim (2). According to him, for Carnap "a specification of the domain of individuals is part and parcel of the interpretation of any one language. If the domain of individuals is changed, we are dealing with a different language." (Hintikka 1991, 328) This "one-domain-assumption" is in fact an interpretational claim concerning Carnap's use of formal languages commonly shared in recent scholarship.[156] It is held that due to Carnap's logicist doctrine inherited from Frege

[156] Compare Hintikka (1999), Awodey (2008), Awodey and Carus (2001), and Reck (2004).

and Russell, quantifiers were conceived by him as logical constants with a fixed interpretation throughout his intellectual career.[157, 158]

Support for this view is usually provided via textual evidence from his later works, particularly from Carnap's works in his *Series of Semantics* in the 1940s (See Awodey 2008). Moreover, also the passage cited in this chapter in Section 4 on "*the one domain of logic*" can be considered as support for such a universal conception for logical languages with fixed range of quantification in *Untersuchungen*.[159]

Nevertheless, I want to argue here that this interpretive claim of fixed assignments to the quantifiers of each type level does not do proper justice to Carnap's *actual use* the simple type theoretic language in the axiomatic context, i.e. in his attempt to symbolize axiomatic theories, its metatheory and its interpretations. My point is that at least for Carnap's early work on axiomatics this understanding of his tacit semantics for STT is not adequate. The reason for this is the following: In his attempt to express the models of an axiomatic theory in this formal language Carnap is assuming and fruitfully implementing a kind of type flexibility concerning the semantics of his language, i.e. a certain flexibility concerning the type assignments to the variable expressions of a given syntactic level or category that is in direct conflict with the "one domain assumption".

Carnap's convention of a type flexible treatment of STT is based on a certain tolerance concerning the type status of logical objects assigned to the variables of the language. In particular, he allows a certain relativity concerning the level of individuals of a given language: for the 0-type-level variables, the actual range of quantification is not fixed to a specific class of elements, e.g. logical elements of type 0, but allows different interpretations also in (classes of) higher-typed constructs. In Part 1 of *Untersuchungen*, Carnap puts the point this way:

---

[157] This is considered as a premise in the logicist tradition since the seminal work of van Hejienoort (1967). Compare Goldfarb's often cited description of this point: "The ranges of the quantifiers – as we would say – are fixed in advance once and for all. The universe of discourse is always the universe, appropriately striated." (Goldfarb 1979, 352)

[158] A similar point was made for some of Tarski's works. Compare again Mancosu: "It is the type-theoretic framework that in the first place decides what the class of individuals (V) is. (…) The type-theoretic framework comes interpreted. In particular, the quantifiers play the role of logical constants and thus, (…), there is no reinterpretation of the quantifiers." (Mancosu 2006, 232)

[159] Compare also Awodey and Carus on this point in Carnap (2000): "What made Carnap's definitions seem natural to him is his "universal" conception of logic (…). There is only one logic, in this conception. Not only were all possible axiom systems fragments of this universal system; all logical analysis of and statements about this system has also to be stated within this same universal system." (Awodey and Carus 2001, 159)

Which elements are considered as "individuals" in respect to a given variable is not determined by the variable's type; at one point elements of this kind can function as individuals, at another point the elements of another kind. Accordingly, elements of a given kind can count as individuals at one time can count as relations at another time. (Carnap 2000, 107)[160]

What is considered as the class of individual for a given language like SST is thus relative to a given context of use.[161] The notion of 'individuals' is thus not an "absolute" but a "relative concept" whose extension in STT, i.e. the range of its first-order variables, has to be stipulated for each case of application of the language.[162] There is thus no unifying universal class of individuals in terms of a fixed class of type-theoretic objects taken as the level of individuals of the language. So, for instance, the individual variables of STT can range over the type 0 objects in one theory and over relations or functions on these, i.e. over type 1 objects in another. If applied to mathematical theories this means that in one case the language might range over a class of natural or real numbers as the individuals of the theory. In another case the language ranges over elements of a higher type that can be constructed from the class of real numbers (see Carnap and Bachmann (1936) and the discussion below).
Several consequences of this convention concerning the use for STT as an "applied logistic" are worth noting here. First, it should be clear that when Carnap is speaking of the context-relativity of individuals, he means different formalizations of axiomatic theories in STT. The use of flexible types is thus explicitly devised for the axiomatic context, i.e. for symbolizing axiomatic theories their models as well as certain metatheory for them.[163] Second, note that for a formal type-theoretic language, this effectively calls for a conventional stipulation of the interpretation of the type-levels of the language, i.e. a stipulation which elements of a typed universe are considered as the individuals of the language at a given time. Note that his type flexibility has the effect that the range of a given variable is de facto reinterpretable. The ranges of the variables of each level are not fixed in the sense of ranging over a certain specified invariant class of elements in the type-theoretic hierarchy. Instead, assignments to

---

[160] "Welche Gegenstände als "Individuen" inbezug auf eine bestimmte Variable anzusehen sind, ist durch den Typus der Variabeln nicht festgelegt; das eine Mal können Gegenstände dieser, das andere Mal Gegenstände jener Art als Individuen fungieren. Und demgemäß können Gegenstände einer bestimmten Art das eine Mal als Individuen gelten, das andere Mal als Relationen." (Carnap 2000, 107)
[161] Compare also Carnap (1929, 19).
[162] Note that a related account of "*relative* types of variables" dependent on the given context of application can already be found in the convention of typical ambiguity in *Principia Mathematica*. See e.g. Whitehead and Russell (1962).
[163] The main context in which this relativized version of "individuals in respect to step n" is used is in the explication of "complete isomorphism" between models. See Carnap (2000).

the variables are supposed to change when STT is applied to another theory. Given this, Carnap is in fact assuming a kind of VDC concerning his background language that is based on the idea of keeping flexible the semantic ties between the formal language and the type-theoretic universe. His convention allows that the relations between this type-theoretic hierarchy and the levels of the language can be treated as flexible and most importantly, as changeable. Variability is thus secured here in terms of variable type assignments to the syntactic expressions of the language.[164]

Finally, the comparison with Tarski's use of formal language to formalize axiomatic theories in his work in the 1930s can be taken up here again. In Mancosu (2006) - with hindsight to Tarski's case - an interesting distinction between two alternative ways to understand the semantics of a type-theoretic language is drawn. In both version, STT is fully interpreted and the interpretation is fixed. In the first, "strong" FDC, the individual quantifiers of STT range over a fixed class of logical objects, Tarski's class of individuals V.[165]

In the second, suggested account of a "weak" FDC, languages like STT are introduced as "specific interpreted languages" for a given axiomatic theory.

> Every theory comes equipped with its background theory of types and with its own interpretation of the theory of types. (…) Thus, there is a certain flexibility in choosing what the class of individuals is that will be assumed in the background. (Mancosu 2006, 233)

Thus, if STT is used for the formalization of an arithmetical theory, the individual variables are to range over the class of natural numbers. In case STT is used for symbolizing an axiom system of geometry, the first order variables range over points (or lines etc.).

Now, Mancosu offers two ways of explanation how this more tolerant version of FDC can be made consistent with the "universalist" assumption of an all-encompassing type-theoretical universe, based on a fixed class of individual objects. One is to suggest a class-theoretic

---

[164] Note that this flexible type theory has no stronger implications for the traditional logicist understanding of type- theoretic hierarchy described above. In particular, the convention is fully consistent with the view that the "logicist" Carnap assumed, *anno* 1928, a universal universe of type-theoretical objects. Moreover, it also does not conflict with the Russelian view that a logical language is universal in the sense of ranging over this all-encompassing universe of objects. What is does come in conflict with is the claim that for Carnap, in his logical approach to formal axiomatics, the universal background language STT comes fixed with a single, invariant interpretation in which each range of quantification for a given type-level is assigned a fixed class of type theoretical objects. Where the flexibility of types becomes virulent is in the assignment of these universes to the different type-levels of STT.

[165] Compare Mancosu (2006): „The theory of types comes already with the meaning of the quantifiers fixed and that determines the range of the individual variables in all possible theories that can be formulated over the theory of types. The range of individual variables (thus the class V) is taken to be the "real" universe of individuals." (ibid, 232)

understanding of the type-theoretic universe. The domain (of individuals) is considered here as a proper class "from which we can manage to carve out set-theoretical interpretations of higher and higher cardinality." (ibid, 235) According to the second reading, the individual class (henceforth IC) assigned to the individual variables of the language can be freely stipulated when the language is introduced for a given task (see ibid, 236). Thus, IC can be interpreted differently in different theoretical context by substituting different classes of mathematical objects to it. (Thus, it can consist of countable set of the natural numbers in on context and of the incountable set of real numbers in another). There exists no prior, in some sense logically privileged individual class of STT. Nevertheless, once the language is set up in this way for a specific mathematical formalization, its interpretation remains fixed. There is thus no way to change or reinterpret IC in a given formalization of a mathematical theory.

Now, this account of a weak FDC seems to come close in spirit to Carnap's understanding of languages like LII in his 1936 paper. One can in fact find passages in Carnap's work from the early 1930s that seems to suggest such a weak conception of FDC for his background language. In Carnap (1931), in providing a general explication of STT, he states that:

> To type 0 belong the names of the elements ("individuals") of the domain of thought that are treated in the respective context (for instance *a*, *b* …). (Carnap 1931, 96)[166]

In the paper "Die Mathematik als Zweig der Logik" from 1930, in the context of discussing the "existential character" of Russell and Whitehead's axioms of infinity, reducibility and choice, he makes the following remark:

> The axiom of infinity is not valid solely because of its form alone; it is valid, if at all, somewhat by coincidence. For some domains of individual it holds, for others it does not. Whether one can speak at all of an absolute domain of individuals, seems problematic. (Carnap 1930a, 306)[167]

These passages strongly suggest that Carnap understood the semantics for STT, if applied in the axiomatic context, in a way similar to the "weak" conception described by Mancosu. Moreover, the last sentence of the second passage could be taken as evidence that Carnap

---

[166] "Zum Typus 0 gehören die Namen der Gegenstände („Individuen") des Denkbereiches, der in dem betreffenden Zusammenhang behandelt wird (etwa *a*, *b*...)." (Carnap 1931, 96)

[167] "Das Unendlichkeitsaxiom gilt aber nicht aufgrund seiner Form allein; sondern es gilt, wenn überhaupt, gewissermaßen zufällig. Für manche Individuenbereiche gilt es, für andere nicht. Ob man von einem absoluten Individuenbereich reden kann, erscheint problematisch." (Carnap 1930a, 306)

seems to sympathize - contra the universalist picture - with a 'substitutional' explanation of weak FDC in Mancosu's sense, where IC is considered to be but a class structure that has to be filled or substituted with a specific class of mathematical entities.

Nevertheless, this similarity should not lead one to overlook where Carnap's account differs essentially from Mancosu's Tarski. One can see the former's convention of a type flexible interpretation of STT as a special kind of weak FDC. It is, however, not a substitutional understanding in Mancosu's sense that follows from this convention. The main body of evidence discussed above shows that Carnap, unlike Tarski assumed a fixed class of type 0 objects at the basis of the hierarchy of types. The first-order variables of the language can be interpreted, not via substitutions of different, non-related classes of elements, but by assignments of classes from different levels in the hierarchy of objects. Thus, any two interpretation of the individuals quantifiers of STT, say IC_1 and IC_2, stand in a fully explicable connection via their respective position in the objective type-theoretic hierarchy. The elements of IC_2 (say the real numbers) can be related to the elements of IC_1 (say the natural numbers) by making explicit the logical construction of the former in terms of the letter. With this convention, Carnap's weak FDC can be understood as a kind of a substitutional account that is consistent with a fixed universe. Moreover, this is also a better explanation of the possibility of domain variation given a fixed hierarchy than Mancosu's explanations since it follows naturally – at least for Carnap's theory – from the traditional 'universalist' assumptions concerning the type-theoretic hierarchy.

A second, more substantial difference to Mancosu's Tarski is this: recall that Mancosu explicitly holds in his explanation that once a specific IC is assigned to the language for a specific formalization, it remains fixed. Note that this is not necessary when siding with Carnap's convention of flexible types. Here, in a specific application of an interpreted language to an axiomatic theory, it is possible to vary the interpretation of the individual variables (and thus the range of individuals) of STT by assigning different classes from the type-theoretic hierarchy. Moreover, and more importantly, this type flexibility concerning the variables of the language opens up a possibility to consider a second kind of VDC for models of an AS formalized in STT that is independent from the 'domain-as-fields' approach sketched above. Recall from Section 2.4 Carnap's remarks on an "arbitrary, stipulated domain of discourse" of the models (ibid, 122) and of "declaring a specific domain of discourse in them to be the domain of individuals" (ibid, 95). In addition to the convention outlined above, Carnap conceives of a second version how model formation and the VDC indicated here can be gained. It is directly related to the type structure of the AS in question or, more precisely, to the type-theoretic architecture of models in Carnap's account. I will dub this approach as "domain variation via type ascent". In a set of notes from the projected Part 2 of *Untersuchungen* Carnap gets explicit on how the idea of type flexibility can be used for domain variation for models. Recall that in his presentation of a model's individual domain, the level of individuals for a given theory depends on the type-structure of the axiom system. The individual domain in M_1 can only be specified "in respect to AS". In the document titled "Domains of a model" (RC 081-01-16) Carnap further discusses the connection of the type-structure of an AS and the type level of the "corresponding domains" of a model. He points out that two models M_1 and M_2 of an axiom system f can have (individual) domains of different types if the type of the variable indicating the corresponding domain in f is 'undetermined' ('*unbestimmt*'):

There is notably as difference between *M* and f: *M* consists of constants, thus any two constituents of *M* therefore have a specific type-relation, they are either of the same or of a different type and if different, then in a certain relation to each other; in contrast, k_1 and k_2 are variables, there are 3 possibilities: of the same type, of different types, undetermined. (RC 081-01-16)[168]

The crucial point here is that for a given variable *R* of f, two admissible models can be constructed that are "of different type" ("*typenverschieden*"), given that *R* is undetermined. Given Carnap's notion of "corresponding domains" between two models M_1 and M_2, the undetermined type expression of the primitives of an axiom system allows that there can be corresponding domains K_1 and K_2 of M_1 and M_2 "in respect to AS that have a different type level" (see also RC 081-01-18). As Carnap makes explicit, this holds in particular for individual domains:

The type of the individual domains [GS: in an AS] does not at all have to be the same (since the AS can stipulate for 2 such domains either that they are type-identical or it lets the type relation undetermined (…). (RC 081-01-18)[169]

[168] "Es ist eben ein Unterschied zwischen *M* und f: *M* besteht aus Konstanten, irgendzwei Gebilde von *M* haben daher ein ganz bestimmtes Typenverhältnis, sie sind entweder von gleichem oder von verschiedenem Typ, und wenn von verschiedenem, so in bestimmtem Verhältnis; k_1 und k_2 hingegen sind Variablen, das sind 3 Möglichkeiten: typengleich, typenverschieden, unbestimmt." (RC 081-01-16)
[169] "Der Typ der Individuenbereiche [GS: in einem AS] braucht durchaus nicht derselbe zu sein (den das AS kann für 2 solche Bereiche entweder bestimmen, dass sie typengleich sind, ode es lässt das Typenverhältnis unbestimmt (…)." (RC 081-01-18)

It is not difficult to see the possible impact of this use of syntactically undetermined type levels for the question of domain variability of the models of f. It can be considered as a certain complementation or strengthening of the domain-as-fields convention. Under the latter convention, the variables symbolizing the theory's primitive terms of a given theory f(R, S, T), if left unspecified, can be (re-)interpreted by different models M_1 and M_2 where the respective model constituting relations R_1, S_1, T_1 and R_2, S_2, T_2 are extensionally different. Nevertheless, their respective fields consist of elements of the same type level. In other words, the type structure of the relations is held constant here, i.e. all relations have fields consisting of elements of the same type level. Now, adding the convention of flexible types, to class of possible models for AS can be extended to models like M_3 where the relations R_3, S_3, T_3 are not of the same type level as the corresponding relations in M_1 or M_2. This allows the reinterpretations of f with models whose domains consist of individual elements of different types.

Interestingly, Carnap and Bachmann are not silent on this issue in their 1936 paper. In fact, the idea of domain variation via type ascent is explicitly mentioned here. In the final section 6 of the paper one can find an explicit argument why such a strengthened form of domain variability might be productively used for some formalization of axiomatic theories. Carnap and Bachmann discuss "informal extensions" in mathematical practice based on models that are 'stepwise' distinct, i.e. not of the same type level:

> In actual informal thought, one has no hesitation in permitting systems of higher level as extensions and it is even quite customary to extend domains by constructing new domains from the elements of the original domain *through one or more steps of class construction*. (ibid, 83, my emphasis)

Note that 'systems' means models and 'domains' means model domains here. The above passage stresses that there are kinds of domain variation in informal mathematical practice that involve an informal ascent to higher-level constructions that, if recast in a type theoretic setting, involve a type ascent of domains described above. Thus, "class construction" in a type-theoretic framework simply means that for a given set T of n-typed objects, n+1-typed objects can be constructed as certain relations of functions on T. As one "classic" example for such informal kinds of model extensions Carnap and Bachmann mention the field of real numbers as the extension of the field of rational numbers gained "by regarding classes of fundamental sequences of rational numbers as elements" (ibid, 83). Thus, in respect to a given axiomatization of fields, one can view the above example as a typical instance of model variation via type ascent. Carnap and Bachman go on to outline a version of a "type-flexible presentation" ("*stufenmäßig bewegliche* Darstellung") (in contrast to the traditional presentation with rigid type distinctions) of axiom systems in STT aimed to logically reconstruct these kinds of domain variations in mathematical practice. The cornerstone is here again a certain flexibility concerning the types of the variables of the formal language:

The variables and those logical constants that in general belong to the type systems would (…) be *flexible as to type*, i.e., they would have no definite type but run through a denumerably infinite sequence of types, beginnings with the "base type" of the given sign. (ibid, 85, my emphasis)

Given this flexible conception of variables in STT, the effect for the interpretations of AS is this: "To a given axiom system then there may belong models of different levels." (ibid, 85) We have seen that this convention of symbolizing axiomatic theories in a non-type-rigid version of STT, devised explicitly to capture kinds of model variation in mathematical practice essentially goes back to Carnap's earlier ideas of flexible types from 1928.

To recapitulate: We have shown in Section 2.4.2 that Carnap's heterodox conception of models, in particular the 'domain-as-fields' convention shared with Tarski is perfectly consistent with the assumption of a FDC concerning the background language in use. Thus, in terms of Hintikka's claims, inference from (2) to (2i) is invalid in Carnap's case since he can perfectly well hold a VDC concerning the models of a theory and still hold a FDC concerning the formal language in which the theory is formalized. However, given Carnap's additional convention of a flexible type theory and of model variation via type ascent outlined above, Hintikka's sweeping claim (2) that the ranges of quantifiers of a given type level are fixed itself seems problematic at best. This is so at least when it is directed, as in Hintikka (1991), at Carnap's use of formal languages in the axiomatic context.[170] Given the "type-flexible presentation" of axiom systems, Carnap's conception of STT allows a kind of reinterpretability of its ranges of quantification that is in direct conflict with the "one domain assumption". Thus, to summarize, we can state that both claims made by Hintikka, (2i) and (2) ware shown to be untenable in the case of Carnap's axiomatics. The remaining question now is: what follows from this for Hintikka's first claim (1)?

---

[170] Let me emphasize this point: my claim here is not to say that (2) (or something equivalent) is wrong *tout court* as a characterization of Carnap's general conception of formal languages. It is to say that it is not a valid characterisation of Carnap's specific use of a language like STT in the axiomatic context, i.e. for the logical reconstruction of mathematical axiomatic theories.

2.5 Submodels and model extensions

Did Carnap, as Hintikka (1991) argues, completely misconceive the use of extremal axioms and thus the informal understanding of model extension in formal axiomatics? Does his formal explication misrepresent the basic intentions behind the completeness and restriction axioms introduced by Hilbert, Peano and Fraenkel?[171] Given the above refutations of Hintikka's claim (2) and his tacit sub-claim (2i), one should be more than skeptical about his claim (1). Nevertheless, in order to address these questions thoroughly, it will be necessary to see what exactly Carnap's understanding of submodels and model extensions tacitly assumed in his theory of extremal axioms was. In particular, it is to see whether he held a conception of model extensions that is comparable to modern account and this despite the continuous emphasis on model's structures and structure extensions outlined in Section 2.2.1. In fact, it seems that at the time of working on the *Untersuchungen* manuscript in 1928, Carnap was not quite certain of the exact relation between "structure extensions" and "model extensions". In a document of Part 2 in the *Nachlass*, we can find him asking himself:

> Should we consider "partial relation" and "families" (with respect to the partial relation) as well as "maxima" and "minima" (with respect to the partial relation) for the relations or the structures? (RC 081-01-05)[172]

The remarks on the extremal structures of an AS as well as his toy example presented in 1936 suggest that Carnap chose the second option. Nevertheless, closer inspection of how the extremal structures and the corresponding extremal axioms are presented in Carnap and Bachmann (1936) will show that the above made distinction between structure and model extensions turns out to be less consequential. In other words, the distinction alluded to in the passage above between structure extension and model extensions has no stronger implications for Carnap's overall theory if both terms are understood in sufficiently broad sense. My point in this section will be to show that –contra Hintikka's interpretation – Carnap assumed a version of model extensions in 1936 that is fully compatible with the informal uses of model extensions in mathematical practice he aims to capture. To see this, a closer look at the formal

[171] Compare Schiemer (2010) for a closer discussion of Carnap's mathematical influences for his account of extremal axioms.
[172] "Sollen wir "Teilverhältnis" und "Familien" (inbezug auf das Teilverhältnis) und "Maxima" und "Minima" (inbezug auf das Teilverhältnis) für die Relationen oder für die Strukturen betrachten?" (RC 081-01-05)

details of Carnap and Bachmann's auxiliary notions of submodel and extension of a model introduced will prove to be instructive.

Following their explications of minimal and maximal structures, Carnap and Bachmann (1936) go on to present their account of the corresponding extremal axioms in §4 of their paper: Informally, a maximal axiom imposes on a model M that "*there is no extension N of M such that F(N) holds* (…)." Accordingly, a minimal axiom imposes that "there is no *N such that F(N) holds an M is an extension of N* (…)" (ibid, 77). The notion of extensions used here is defined more formally in terms of the models of a theory: for any two models M_1 and N_1, N_1 is an extension of M_1 if M_1 is a "proper part" of N_1, i.e.

$$M_1 \subset N_1 . M_1 \neq N_1.$$

Given this, a further ramification of this kind of extension is mentioned. The conceptual motivation for it is to make a distinction between two types of extension based on whether or not the extended models are admitted that can be isomorphically embedded in the original model. In Carnap's own terms, whether or not it is axiomatically stipulated that the "structures of models are really distinct" (ibid, 77). In the first case, the relation of extension is defined via the notion of "proper part" as Carnap conceives it in the sense above. This leaves open the possibility that two models share what he terms a "dividable structure".[173]

In the second case of extension, structures of the two models are deliberately fixed to be "undividable", the admissible models are thus non-isomorphic – a condition explicitly "included in the definition of extension". These two concepts of extension, terminologically fixed as "model-extension" and "structure-extension", are given the following formal presentations (see ibid, 77):

$$\mathrm{Erw}_m(N; M) =_{df} M \subset N \,.\, M \neq N$$

$$\mathrm{Erw}_s(N; M) =_{df} M \subset N \,.\, {\sim}\mathrm{Ism}_v(M, N)$$

Note that the first condition for extensions is weaker than the second: the use "proper parts" leaves open the possibility that there exits a proper extension N_1 of a given model M_1 such that N_1 is isomorphic to M_1. This explicitly ruled out in the more restrictive case of second type that also rules out any isomorphic extensions of M_1.[174]

---

[173] A structure is called "dividable" if there exists a proper submodel M of a model N that satisfies the structure of M, put formally $(\exists M)(\exists N)$ $(S(M).\ S(N).\ M \subset N.\ M \neq N)$. It follows in this case that models are "isomorphic with a proper part of itself" (ibid, 74-75).

[174] Given these two types of extensions, Carnap & Bachmann distinguish between two versions of extremality axioms: in the case of maximality axioms, between "maximal model axioms" and "maximal structure axioms":

Surprisingly, this distinction between model-extension and structure-extension is not mentioned in Hintikka's commentary. Neither does he get more explicit about Carnap's specific understanding of the notions of "part" or "proper part" of a model, i.e. Carnap's use of the "inclusion" relation in his treatment of what in a modern terms amounts to a (proper) submodel featured in both definitions. Obviously, Carnap's specific understanding of '$M \subset N$. $M \neq N$' in this context turns out of be crucial for his tacit assumptions concerning the extensions (or restrictions) of a model and, in consequence, extremal axioms in an axiomatization.

Let us see how Carnap's notion of a "proper part" relates to the modern model-theoretic notion of a submodel. Here, given two models $U = \langle A, s\rangle$ and $W = \langle A', t\rangle$ of a language, U is considered to be a submodel or substructure of W (or inversely, W an extension of U), if the following conditions hold:

(i) $A \subseteq A'$

(ii) for every n-ary relation symbol $R_U = R_W \cap A^n$

(iii) for every n-ary function symbol, $f_U = f_W | A^n$

(iv) for every constant symbol, $c_U = c_W$.[175]

Is Carnap's notion of a "proper part" of a model structurally comparable to this? Is he assuming a condition like (i)? Note that according to Hintikka this cannot be the case since it would involve the possibility of actually changing the individual domain. This - according to his second claim - was simply inconceivable for Carnap. However, we have seen that this reading shows to be untenable as soon as one acknowledges the heterodox character of Carnap's conception of models. Where does this insight leave us in our understanding of his central use of submodels in his theory? Here again, a closer look at what Carnap actually says about the very notion of "proper parts" proves to be instructive. Recall again that models in *Untersuchungen* are understood as sequences of relations (of a specified type level) of the form $M_1 = (R_1, S_1, T_1)$. In an insightful footnote in §3 of Carnap & Bachmann (1936) the meaning of '$M \subset N$', is further specified:

$\mathrm{Max}_m(F; M) =_{df} \sim(\exists N)\ (M \subset N.\ M \neq N.\ F(N))$; $\mathrm{Max}_s(F; M) =_{df} \sim(\exists N)\ (M \subset N.\ \sim\mathrm{Ism}_v(M, N).\ F(N))$. The second axiom rules out extensions of a model M that do not share a "dividable structure" with M; the first and stronger axiom rules out extensions of *M per se* (see ibid, 77).

[175] Compare e.g. Enderton (2001, 95) for a similar definition.

Let ‘*u, v, …, w, F, G, …H*’ and ‘*x, y, …, z, J, K, …, L*’ be two series of variables, where ’*u*’, …, ’*z*’ are the individual variables and ’*F*’, …, ’*L*’ the predicate variables, which are abbreviated by ’*M*’ and ’*N*’; then ’$M \subset N$’ is an abbreviation for ’$u = x\ .\ v = y.\ ….\ w = z.\ F \subset J.\ G \subset K.\ …\ .\ H \subset L$’. (ibid, 75)[176]

Note that here, as in the official definition of a “formal model” cited in Section 2.4 an explicit notational specification of the model domains in the sense of condition (i) in the modern definition is missing. In fact, one might be tempted to read the series of individual variables ‘*u, v, …, w*’ and ‘*x, y, …, z*’ in the 1936 passage as indicators of the respective domains of the models. In that case, the crucial question would be how to make sense of the series of conjunctions of identity relations between the individual variables of the two structures ’$u = x\ .\ v = y.\ ….\ w = z$’. Now, if understood as a relation between the respective individual domains of *M* and *N*, a serious tension with our main interpretive claim that Carnap held a VDC for his models would arise. This is due to the fact that in this case, the domain of a model would be fixed to a specific set given the fixed sequence of individual variables. Moreover, the series of identity relations would clearly imply a nonstandard conception of submodels for it imposes on the fields of the two models to be identical or at least equinumerous. In this reading, closely in line with Hintikka’s critical remarks, Carnap theory clearly would miss a modern account of submodels since the domain restriction necessary for condition (i) simply could not be expressed. This interpretation of Carnap’s theory of submodels also gains *prima facie* plausibility in the light of Carnap and Bachmann’s point set example mentioned in Section 2.2.2 where the individual variables stand for the elements of a fixed finite set of individuals that stays invariant through all structure extensions.

Nonetheless, Carnap and Bachmann did not take the sequences of individual variables as indicating the respective universes of discourse of two models. This is so for several reasons: first, the use of ‘*u, v, …, w*’ and ‘*x, y, …, z*’ for presenting the models’ universes would limit the formal treatment to models with a finite domain. This is due to the fact that the sequences are restricted to a finite number of “basic signs”, i.e. variables of a specified level. Carnap & Bachmann explicitly mention this condition of the finiteness in 1936: “*M*” is an abbreviation for a *finite sequence of variables*.” (ibid, 70, my emphasis)

[176] A similar but simplified version of the notion of a “submodel” (“*Teilmodell*”) (note the explicit use of the term here!) can be found in his notes in Carnap’s *Nachlass*. Here, Carnap states that for two models *P*: *R*, *S*, *T* and *P’*: *R’*, *S’*, *T’* the submodel relation $P \subset P’$ expresses $R \subset R’$, $S \subset S’$, $T \subset T’$ (see RC-081-01-19/119, notation slightly changed).

Moreover, given this syntactic restriction, the attempt to understand the series of individual variables as a convention to indicate a model's domain becomes especially implausible with regard to Carnap's examples of actual model extensions in mathematics, including extensions of models with infinite domains (see the following Section). It would in fact be inconsistent with Carnap's presentation of models with infinite powers since here the fields of the relations are simply not presentable by a finite series of variables.

If '*u*, *v*, …, *w*' and '*x*, *y*, …, *z*' are not used as syntactic indicatory of domains, what are their respective roles in the presentation of a model's structure? I think the most plausible reading is to understand them simply as placeholders for the "primitive individuals" ("*Grundindividuen*") of an axiomatic theory. There is textual support for this view in *Untersuchungen* and Carnap and Bachmann (1936) in the discussion of mathematical axiom systems, most notably that of Peano arithmetic, where Carnap mentions the *basic individuals* of the theory in the sense above:

> The axiom system of elementary arithmetic in the Peano form with "0", "successor" and "natural number" as basic concepts can be represented as a predicate taking as arguments one individual variable, one two-place, and one one-place predicate variable of the first level." (ibid, 169)

We have seen that in Carnap's conception of axiomatics, unlike in the modern schematic account, Peano arithmetic is symbolized as a propositional function $f(x, S(x, y), N(x))$ in pure STT where x stands for the primitive constant *zero*, $S(x, y)$ for *successor* and $N(x)$ for the *class of natural numbers*. A model of f of the form $\langle 0, S_1, N_1\rangle$ then assigns a binary relation to S, a class to N and an individual constant to x. In my mind, it is in this respect that '*u*, *v*, …, *w*' - if satisfied by a model - expresses a finite number of individuals that are assigned to the individual variables of the AS in question.

Note that the example discussed in this passage is somewhat atypical in Carnap's formal reconstruction of axiom systems and its models. In the majority of cases, Carnap does not explicitly refer to the level of individual expressions related to a model. Commonly, models are presented as sequences of relations of this form (R_1, S_1, T_1, …), i.e. without any mention of individual constants. A reason for this convention can be found in a section for Part 2 of *Untersuchungen* titled "Reduction of the primitive concepts" ("*Reduktion der Grundbegriffe*") (RC 081-01-12). Here, Carnap argues that the primitive individual terms can usually be 'eliminated' as 'primitive terms' from the AS if a "structural description" can be given for them in terms of the other primitive relations of the AS.[177,178] What is important here is that this idea of the structural eliminability of primitive individuals explains why individual variables are in most cases simply not mentioned in a model's structure. Moreover, if the series of individual variables in the above passage *are* understood as (potentially eliminable) primitive individuals of the AS, they of course have no restricting effect on the domains of the models satisfying the AS. In this case, the series of identity relations between the individual variables of the two structures '$u = x$. $v = y$. …. $w = z$' simply hold that for any two models M_1 and N_1 of an AS there has to be a direct mapping between the respective assignments to the (finite set of) primitive individuals of the AS. Thus, it essentially expresses condition (iv) in the modern definition of submodels.

Given this, it remains to see how the inclusion relations '$F \subset J$. $G \subset K$. … . $H \subset L$' are understood in the context of models. Recall that, Carnap's use of the inclusion sign was already exemplified in the point set example discussed above: Here, F was a partial relation of J, G a partial relation of K etc. This relation between the constituents of two models is further specified in Carnap's manuscript in the *Nachlass*. In a document for the second part of *Untersuchungen*, the connection between a "proper submodel" ("echtes Teilmodell") Tm(P, P') and a "proper partial relation" ("echte Teilrelation") Tr(P, P') is made explicit (RC 081-01-07). Carnap states here that given two models with model variables $P = (R, S, T)$ and $P' = (R', S', T')$: "Tm (P,P') means: ((Tr(R,R') $\vee$ $R = R'$) & (S…) & (T…) & (Tr(R,R') $\vee$ Tr(S,S') $\vee$ Tr(T,T'))" (RC 081-01-07, notation slightly changed.) In addition, he states that "Tm^{R} (bzw. $\text{Tm}^{R,S}$ bzw. $\text{Tm}^{R,S,T}$) means: Tm, given that a proper partial relation holds for R (or for R, S or for R, S, T respectively), for the rest proper or improper relations." (ibid, notation slightly changed.) This makes evident that a submodel was conceived here as a set of (proper or improper) partial relations. More specifically a proper submodel is understood as the sequence of disjunctions of partial relations (of the relations constituting the original model).
Now, Carnap does not get more precise on his understanding of the notion of a partial relation here (nor in Carnap and Bachmann (1936)). Nonetheless, one can get a clear grip on his use of the notion by taking his related work into account. For example, in Carnap (2000), the following definition of a partial relation Trl(P, Q) is given (ibid, 67):

[177] Carnap, in the first first of *Untersuchungen* also mentions this possibility and refers to *Principia Mathematica* §122 for the presentation of such an elimination in the case of Peano arithmetic (see Carnap 2000, 88-89).
[178] Carnap's discussion of such descriptions of primitive terms would take us too far afield here. On this topic, compare also Carnap's notes on the notion of "Homotopie" in (RC 081-01-20 and -21).

$$\mathrm{Trl}(P, Q) =_{\mathrm{Df}} (u, v, \ldots z)[P(u, v, \ldots z) \rightarrow Q(u, v, \ldots z)]$$

In *Abriss,* written around the same time, a more detailed, ramified presentation is given that *grosso modo* corresponds to the modern version and explicitly refers to the "restriction of the field" of relation:

The partial relation of R which results from the restriction of the domain of R to the class α, is expressed by $\alpha\uparrow R$ (read as "R pre-restricted to α"); accordingly the restriction of the range: $R\downarrow\beta$ ("R post-restricted to β"); in case of both restrictions: $\alpha\uparrow R\downarrow\beta$, in case of the restriction of the field: $R\uparrow\downarrow\gamma$ ("R restricted to γ"). (Carnap 1927, 37, notation slightly changed)[179]

The crucial point to note here is that since models are conceived as sequences of n-ary relations by Carnap, all three kinds of restriction also apply to his conception of submodels. Most importantly, keeping in mind the 'domain-as-fields' convention only the third kind of restriction of a relation's field "to a class γ" necessarily involves a restriction to a subset of individuals of the original universe of a model. One can thus plausibly understand a submodel N_1 of model M_1 as a set of relations restricted to a subset of the field of the relations of M_1. Thus, in order to reconcile Carnap's definition of submodels with the modern definition one simply has to reformulate the conditions mentioned above: recall again that the models M_1 and M_2 - unlike U and W – share the same range of individuals of the given background language. Nevertheless – assuming the domains-as-field convention – an equivalent to condition (i) (and thus (ii) and (iii)) is met by the restriction of the field (more precisely, the restriction of the union of the fields) of M_1 to that of M_2. Thus, (i) is met in Carnap's conception by understanding the restrictions implied in (ii) (and (iii) to simultaneously result in a restriction of the domain of M_1 to the domain of M_2.

If submodels (or more specifically partial structures representing different admissible models) are understood in this way, Carnap's resulting theory of extremal conditions becomes consistent with the standard account. Nevertheless, this should not lead one to overlook the

[179] "Die Teilrelation von R, die durch Beschränkung des Vorbereiches von R auf die Klasse α entsteht, wird bezeichnet mit $\alpha\uparrow R$ (gelesen "R vorbeschränkt auf α"); entsprechend bei Beschränkung des Nachbereiches: $R\uparrow\beta$ ("R nachbeschränkt auf β"); im Falle beider Beschränkungen: $\alpha\uparrow R\downarrow\beta$, bei Beschränkung des Feldes: $R\uparrow\downarrow\gamma$ ("R beschränkt auf γ")." (Carnap 1927, 37, notation slightly changed)

fact that there remains a central difference to the modern account of extensions. Not all types of extensions conceived by Carnap involve an effective extension (or restriction) of the individual domain in the sense of (i). The finite point set example mentioned in §2.2 is a case in point. As we have seen, the extensions described here simply do *not* imply the addition of new elements to the field of (interpretations) of the primitive variable R, but include only extensions over an invariant field. The point to stress here is that all structure extensions in this example present instances of "pre-" and "post-restriction", i.e. versions of $M \subset N$ expressing a restriction of the domain or the range of N. Notably, these kind of restrictions are perfectly compatible with an invariant fixed field shared by both relations M and N. Thus, the interpreted structures R_a, R_b, and R_d are partial relations of R_c, R_e, R_g, and R_f that result from a restriction of the respective domain or range but *not* of the respective field. For instance, in $R_a \subset R_e$, R_a results from restricting the range of R_e, viz. {1, 3, 4} to the smaller class {1, 4}, provided that the fields of R_a and R_e stay invariant. Quite plausibly, Hintikka's misconception of Carnap and Bachmann's theory of model extensions is based specifically on this toy example. It is important, however, to see that the invariance of the models' domains presents a contingent rather than an essential feature of his theory. In particular, we have to see that the example of point set models with different structures does not – contrary to Hintikka's assumption – have any strong implications for his their general theory of model extensions. Carnap and Bachmann emphasize the fact that the example deals with "finite, therefore undividable, structures." (ibid, 78) In contrast, the examples of model extensions from real mathematics discussed in the paper all have "infinite structures" where the extension of a model *necessarily* involves the extension of set of individuals, i.e. the field of a relation.[180] This is to say that, from a mathematical point of view, Carnap and Bachmann's example of point set models is clearly misleading, since all extensions of mathematical models with infinite powers discussed in the remainder of the paper clearly and necessarily presuppose the extensibility of the respective domains of individuals. To see this, a closer look at the mathematical examples they give will be instructive.

---

[180] This move from the toy example to instances of model extension in mathematics corresponds to what Carnap terms in *Untersuchungen* the distinction between "finite" and "infinite" axiom systems, i.e. axiom systems with models of "infinite structures (structures of models with an infinite number of immediate and mediate elements)" (see Carnap 2000, 149-150)

2.5.1 Examples of mathematical model extensions

My point in the above section was this: if one grants that Carnap understood models in 1936 and before as sequences of relations and their domains as the respective union of fields, then his theory of extremal axioms shows to be perfectly consistent with the informal versions of completeness and minimal axioms of Hilbert, Fraenkel and others. This point can be strengthened by looking at examples of genuinely mathematical model extensions discussed by Carnap and Bachmann. In fact, the examples from mathematical axiomatics in Carnap & Bachmann 1936 show their clear conviction to capture mathematically relevant instances of model extensions (and thus also those instances including the extension of individual domain of a given model).

The most extensively discussed example in 1936 is an axiomatization of "elementary arithmetic" presented at the end of §4.[181] The axiom system defines a primitive binary relation R based on Russell's concept of progressions.[182] It consists three first-order axioms expressing the "endlessness" (b1), injectivity (b2), and the existence of a unique base element (b3) as well as one second-order "minimal structure axiom" (b4) (ibid, 79).[183] What is insightful for our concern here about example is their discussion of model classes for the two "base systems" ("*Rumpf-Axiomensysteme*") b1-b2(*R*) and b1-b3(*R*) and of the effects of adding a minimal axiom to those. In particular, it is shown that b1-b3 is satisfied by models with different model structures. The AS describes a dividable structure *Prog* (for '*progression*') as well as an infinite number of non-dividable structures, i.e. "zero to infinitely many cycles with from one to infinitely many elements" (ibid, 79).[184] Thus, models of b1-b3(*R*) can be instances of a progression or of more complex structures (i.e. progressions plus cycles of up to infinite elements, progressions of progressions etc.).[185] Nevertheless, all relations b1-b3(*R*) (without the minimal condition b4) have fields of equal power as the set of

---

[181] The same axiom system is also discussed in (RC 081-01-10) in Carnap's *Nachlass*.

[182] Progressions are defined in *Principia Mathematica* as a specific type of series, i.e. as a class of all inductive cardinals whose "generating relation" satisfies the following conditions: (i) being 1-1; (ii) having a first term not included in ran*R*; (iii) being denumerably infinite; (iv) being closed under its "generating relation" *R* (Whitehead and Russell (1962, 245). The intention behind this last closure condition is to exclude those 1-1 relations that result in independent chains of elements, i.e. infinite series with a serial number other than ω (see ibid, 245-248).

[183] In formal terms: (b1) $(x)(y)$ [$R(x,y) \supset (\exists z(R(y,z))$]; (b2) $(x)(y)(z)$ [$(R(x,y).\ R(x,z) \supset x = y).\ (R(x,y).\ R(z,y) \supset x = z)$]; (b3) $(\mathrm{dom}R - \mathrm{ran}R)\ \varepsilon\ 1$; (b4) $\mathrm{Min}_S(\text{b1-b3}; R)$ (see ibid, 79, notation slightly changed).

[184] The concept of "cycle" ("*Zykel*") is defined here as follows: "By a cycle with *n* elements we understand a one-one relation whose field consists of a single closed *R*-family with n elements. (…) In the limiting case $n = \infty$ there arise a one-one relation, with no first of last element, which consists of single open *R*-family." (Carnap 1929, 178)

[185] Compare also Russell (1993).

natural numbers. This changes if one considers the possible model class of b1-b2(*R*). As Carnap and Bachmann point out, the base system consisting of b1-b2(*R*) allows different structures and thus models consisting of finite circular relations, i.e. single cycles of one to finitely many elements *as well as* of models of mixed infinite structures (ibid, 80). Thus, whereas all models satisfying b1-b3(*R*) have denumerable infinite domains, b1-b2(*R*) also allows instances of *R* whose fields are subsets of the set of natural numbers.[186]

This point is further elucidated in Bachmann (1934). In discussing an similar theory of basic arithmetic, Bachmann points out that there are different "partial relations" of a progression *R* that come in one of three possible forms: (a) a finite segment ("*Abschnitt*") of a progression, (b) several (up to infinitely many) finite segments of *R* or (c) an infinite segment of *R*. (see Bachmann 1934, 49-50). Relations of the first type have finite fields and are therefore non-isomorphic to *R*. This illustrates that b1-b2(*R*) allows interpretations of *R* with different cardinalities. What is important to see is that here is that any extension of a model R_1 of type (a) to a model R_2 of type (c) necessarily involves the extension of the field of R_1 to a denumerable infinite domain.

Note also that if the base system b1-b3(*R*) is closed by a minimal structure axiom (b4) of the form $\sim(\exists N)(N \subset M.\ \mathrm{Ism}_V(M,N).\ F(N))$ and its model class is therefore restricted to interpretations of infinite dividable structure, model extensions and restrictions have to be understood in the standard sense (see Carnap and Bachmann 1981, 80). The remaining class of admissible models consists of an infinite number of interpretations of *R* that are progressions and as such equal in power to the natural numbers.[187]

One could make the following objection at this point: The cases of structure and model extensions discussed by Carnap and Bachmann for different base systems of elementary arithmetic show that they conceived of model extensions in the case of subsystems of arithmetic. However, (for whatever reason) this need not be so for stronger cases of mathematical model extensions, in particular extensions to systems with a non-denumerable universe. But again, there is textual evidence in Carnap and Bachmann (1936) that model extensions of these kinds are also understood in the sense outlined above, i.e. as extensions of

[186] In Carnap and Bachmann's own words: "The axioms "b1-b2(*R*)" are satisfied not only by the models of "b1-b3(*R*)" but also by single cycles and combinations of finitely many or infinitely many cycles." (ibid, 80)

[187] Nevertheless, universe of the progressions can be different as long as they can be isomorphically embedded into each other. For instance, the domain of $M_1 = \langle D_1, R_1\rangle$ might be the set of natural numbers starting from number 100, the domain of $M_2 = \langle D_2, R_2\rangle$ the whole set of the natural numbers. In this case, an M_1 is a submodel of M_2 in the proper sense since $D_1 \subseteq D_2$.

the field of the interpreted primitive terms of an axiomatic theory. One *locus classicus* for the reflection about models of different higher cardinalities is Hilbert's axiomatization in Euclidian geometry in *Grundlagen der Geometrie* (Hilbert 1910). As is well known, §9 of the book contains Hilbert's classic "model-theoretic" consistency proof for his axiom system. Consistency is in fact proved for two axiom systems. The first is the base system consisting of all axioms of the first four axioms groups plus the Archimedean axiom. The proof is based here on an interpretation of the primitive geometrical terms "point," "line," "between," etc. in a model with a countable infinite "domain" ("*Bereich*") Ω ranging over a certain field of algebraic numbers (see ibid, §2).[188] Consistency of the full axiomatization, including the *axiom of completeness*, is then proved in a second step via the interpretation of the primitives in an extended model with the domain Ω' ranging over all real numbers (see ibid, 26).

Carnap, in *Untersuchungen*, clearly had in mind this use of models when referring to Hilbert method repeatedly throughout the manuscript. In fact, at one place he explicitly discusses Hilbert's consistency proofs by the "presentation of a model" ("*Aufweisung eines Modells*"), in particular his "arithmetic model" for Euclidian geometry:

> Every model of the axiom system can be mapped to the well-known model in which points are considered as triples of real numbers and the remaining primitive terms are interpreted accordingly. (Carnap 2000, 141)[189]

Hence, the arithmetic model referred to here is Hilbert's second one, ranging over closed ordered field of real numbers. Unfortunately, there is no explicit discussion, neither in Carnap (2000) nor in Carnap and Bachmann (1936) of the different models (or model classes) for the base systems of Hilbert's axiomatization as in the case of elementary arithmetic. Nonetheless, there is direct support that the kind of model extensions from Ω to Ω', i.e. from a model with a denumerable infinite to a model with a non-denumerable domain were supposed to be presentable in their formal framework. In Bachmann (1936), a side paper to Carnap and Bachmann (1936), Hilbert's case is taken up again. Theorem 1 in the paper states that an extremal axiom Ext(A_1-A_n; *M*) is dependent on the base A_1-A_n(*M*) if and only if there exists

[188] A comparable algebraic model is used in Hilbert's "model construction" ("*Modellaufweis*") for the independence proof of the Archimedean axiom in §12 Here, however, the domain Ω(t) ranges over a denumerable set of algebraic functions (ibid, 29-33). Compare Awodey and Reck (2002, 13) for further details.

[189] "Jedes Modell dieses Axiomensystems lässt sich eineindeutig abbilden auf das bekannte Modell, bei dem die Tripel reeller Zahlen als Punkte genommen, und die übrigen Grundbegriffe entsprechend interpretiert werden." (Carnap 2000, 141)

an extension of the model M that satisfies the base system. In formal terms, the extremal axiom is independent if the formula (3*)

$$\sim (\exists M)(\exists N)\ A_1\text{-}A_n(M).\ A_1\text{-}A_n(N).\ \mathrm{Erw}(N; M)$$

is true (Bachmann 1936, 45). Bachmann's example for this is illustrative for our concern: he states that Hilbert's axiom of completeness is independent of its base axioms in this respect "since the algebraic model (*Grundlagen* §9) and the Cartesian analytic geometry are two models M_1 and M_2 for which the axioms I-VI and Erw(N_1; M_1) hold; therefore the negation of (3*) holds." (ibid, 45)[190] Given the fact that Bachmann (1936) is identical to Carnap and Bachmann's joint paper from the same year with respect to terminology and background definitions, I take this as conclusive evidence that Carnap too, in his theory of extremal axioms, conceived model extensions fully in line with the then standard informal mathematical practice. Given this, little support remains for Hintikka's claim (1) concerning the descriptive inadequacy of Carnap's theory in view of the mathematical versions of extremal axioms.

2.5.2 Extensions and language "transition"

A further argument and more systematic that Carnap conceived model extensions close to the modern sense can be based on Carnap and Bachmann's perceptive remarks on a possible limitation of their approach in the final §6 of the 1936 paper. These concern the strong dependency of a theory's models on the intended interpretation of the background language. In particular, we saw that in Carnap's conception, all possible domains have to be subsets of the range of individuals of the language in use. In the example discussed in Section 2.4.1, the domains of the models for f(R) are subsets of the substitution class of numerals D_{LII}. D_{LII} thus presents an upper limit to possible extensions of a given model of f(R) that is expressed in LII. More generally, given Carnap's account, the freedom to construct and extend models for an axiomatic theory formalized in a given language is restricted to its intended interpretation. When considering cases of mathematical model extensions as in Section 2.5.1, the question naturally arises how his reconstruction is supposed to capture extensions to models of higher infinite cardinalities in the case of analysis or geometry in the case the theories are expressed

[190] "(…) denn das algebraische Modell (Grundl §9) und die Cartesische analytische Geometrie sind zwei Modelle M_1 und N_1 für die die Axiome I-VI und Erw(N_1; M_1) gültig ist; also gilt das Negat von (3*)." (ibid, 45)

in a background language like LII with a denumerable range of individuals. This, in particular, if this language's interpretation is supposed to remain fixed.

Interestingly, §6 of Carnap and Bachmann (1936) clearly shows that the authors were fully aware of this possible weakness of their theory when compared to "informal extensions" in mathematics. This is again left unnoticed in Hintikka (1991). In fact, they acknowledge the language dependency as a "serious objection" against their account arguing that one can in principle devise cases where a (relatively) weak interpreted language effectively imposes constraints on model building (ibid, 83).[191] The mathematical example mentioned to illustrate this is insightful:

> For example, if we regard the Hilbertian axiom system for Euclidian geometry as a propositional function of first level in a language with a denumerably infinite domain of individuals, there can be models which have only higher level extensions. (ibid, 83)

Thus, working in a language like LII, Hilbert's models with domain Ω (and Ω(t)) can be constructed. They cannot, however, be extended to Hilbert's model with domain Ω' ranging over full Cartesian space. It is important to stress here that these remarks alone underline that when Carnap speaks of model extensions he actually means extensions of individuals and not extensions of relations on a fixed set of individuals. This point can be further strengthened by looking at the different proposals outlined in §6 to block the potential limitation of their formal reconstruction. The first, most simple solution is simply to stipulate that the language in use is to be sufficiently rich in order not to limit model building for a given theory (ibid, 83).[192]

The second option outlined in §6 to vindicate the theory of extremal axioms was already discussed in Section 2.4.3 of the present chapter. It consists in introducing a flexible theory of

[191] Bear in mind that what is meant in this context by the relative weakness (or richness) of a language is its type-levels or its syntactical resources, but the range of individuals, the cardinality of its intended interpretation. Compare Carnap and Bachmann on this point: "It is to be sure quite possible to do without the greater richness afforded by admitting extensions of higher levels if the background language is rich enough, especially *with respect to its domain of individuals*. When, however, the language exhibits a certain poorness it is quite possible that extensions of higher level exist but none of the same level. (ibid, 83, my emphasis)

[192] An explicit formulation of this option can be found in Bachmann (1936): "We presuppose that the domain of individuals of [a background language] Σ is rich enough in order to secure that the impossibility of extensions of models stipulated in the extremal axioms for the considered axiom systems is not a consequence of the poverty of the language. (Bachmann 1936, 39) Note that this condition seems to be implicitly assumed in Carnap's theory of extremal axioms in *Untersuchungen*. There, in 1928, the theory was based on a universal background language. With the turn to the framework of *Logical Syntax* outlined in Section 2.4.1, such a condition on the language's interpretation had to be explicitly stipulated via an additional premise.

types in use in order to be able to model "informal extensions" in mathematics that involve (informal) type ascents (see Carnap and Bachmann (1936, 84-85).

The third attempt to block possible limitations of a language deserves closer attention here. Carnap and Bachmann suggest that the above mentioned restriction to model building can be circumvented by allowing model extensions that effectively transcend the limits of the given language in which the theory is formalized. Thus, the idea is to allow extensions of the kind mentioned above via the "transition" to a semantically richer background language (ibid, 84). In the case of Hilbert's axiomatization of geometry, the extension from a model ranging over Ω to one ranging over Ω' can thus be expressed by a change to a language with a non-denumerable range of individuals. Clearly, the crucial challenge of this approach is to express this language transition for a given theory in the conceptual framework of *Logical Syntax*. More specifically, it is to modify Carnap and Bachmann's heterodox definition of truth in a model via analyticity in a given language (presented in Section 4.1) in order to allow for these changes of the background language. Concerning this point, the authors present an insightful, slightly altered condition of *analyticity*:

> (…) in the case where the object language is weak, we amend the definition of "analytic in S_1" in such a way that a sentence of the form "$(M)F_1(M)$" is called analytic only when it remains analytic under transition to a language S_2 with a larger domain of individuals. In this way the earlier discussion will remain valid. (ibid, 84)

Let us name the revised scheme expressed here *analyticity**. It can be rephrased in this way:

(*analytic)** A sentence φ is analytic* iff φ, if analytic in S_1, is also analytic in S_2, where S_2 is semantically richer than S_1.

Given this amended version, one can reformulate Carnap's original version of truth in a model by saying that M_1 is a model of F if $F(M_1)$ is *analytic** in the sense specified here.

Several points are worth to be mentioned here. First, note a number of subtle but significant differences to the original version. For starters, talk of the substitution class of individual constants is replaced by talk of the "domain of individuals" of a language. Thus, one can witness a tacit shift here in Carnap's thought from a substitutional treatment of individuals (in terms of numerals) to an extensional understanding of individuals. Moreover, note that unlike before the new analyticity*-scheme explicitly involves quantification over models. The model variable M is not treated as free, but as a bound variable here, holding for all models of F_1.

Given these variations, it is important to see what is actually implied in this definition and whether or not it presents a valid option to allow Carnap's version of model extensions. Carnap and Bachmann's intuitive motivation underlying the corrected definition of truth in a model was this: For a given theory F_1, there exists an extension M_2 of a model M_1 that satisfies F_1 but (unlike M_1) cannot be captured by S_1, since the domain of M_2 is more comprehensive than the range of individuals of S_1. Thus, S_1 is too restricted to express M_2. Nevertheless, M_2 is clearly considered an admissible model of F_1 in so far as it is an instance of the model variable M in $(M)F_1(M)$. The central point to see here is that the class of models of the theory expressed in $(M)F_1(M)$ is understood somewhat independently of a given language. It is the range of (M) that transcends the actual bounds of the original language S_1.

The notion of a "transition" to a richer language is introduced then in order to provide a formal valid account of how this quantification over all models can include all mathematically relevant cases of models (and model extensions), irrespective of the specific background language. How is the modified account of *analyticity** conceptually connected to the analyticity conditions for a fixed language outlined in Section 4.1 above?[193] Note that, in our example, the condition not only requires that M_2 is a model of F_1 if $F_1(M_2)$ is analytic in S_2. It expresses a stronger condition, namely that M_2 can be expressed as a model extension of M_1 despite the transition to S_2. Thus, the central difficulty in the amended definition is to redefine the analyticity condition for the expression '$(M)F_1(M)$' given the transition from S_1 to a semantically richer language S_2, where $\mathrm{fld}(M_2) \subseteq D_{S2}$. How can this additional condition for analyticity properly be understood? In order to see what the requirement stipulated here amounts to, we need to spell out in closer detail what effect lies behind this new condition

[193] Given the lack of further specification, one could be inclined to view the condition for analyticity* in Carnap and Bachmann (1936) as conceptually related to Carnap's technical notions of a "sub-language" and "translation" in *Logical Syntax* (Carnap 2002, §50 and §61 respectively). Briefly, a sub-language S_1 of S_2 defined as language for which following conditions hold: (i) every sentence of S_1 is a sentence of S_2 and (ii) every sentence analytic in S_1 is also analytic in S_2. S_1 is also a "conservative sub-language" of S_2 if , in addition to (i) and (ii), also the following condition holds: (iii) if a sentence φ is analytic in S_2 then, if contained in S_1, is also analytic in S_1 (See Carnap 2002, 179). Put informally, (ii) secures the conditional from analyticity in S_1 to analyticity in S_2. The third condition secures the semantic conservativeness of the language extension to the effect that S_2 is conservative over S_1 with respect to the class of analytical sentences. Accordingly, Carnap defines a "transformance" in §61 of *Logical Syntax* as a syntactic correlation f (of "expression classes") between two "isomorphic" language S_1 and S_2. A syntactic correlation f is then a "translation" of S_1 into S_3 if S_3 contains S_2 as a sub-language (See Carnap 2002, 223-224). Note, however, that a comparison of "translation" and "transition" would overlook the different nature of language extension involved in the two cases. In contrast to Carnap's sub-languages in *Logical Syntax*, language transitions as outlined above concern primarily the extension of the respective base of individuals. Extensions understood here not as new type level of predicates, but, as Carnap and Bachmann put it quite explicitly, as "a larger domain of individuals". In the context of *Logical Syntax* a translation holds for languages with an identical interpretation of the individual variables. In the 1936 case of a transition from S_1 and S_2 to crucial difference between the two languages is of course that S_2 is essentially richer in terms of its domain of individuals.

expressed here that ""$(M)F_1(M)$" is called analytic only when it remains analytic under transition to a language S_2". Carnap's refined analyticity condition suggests this for the context of his model theory: Consider first the analyticity of $(M)F_1(M)$ in S_1. The closed formula is analytic if for every instances of the model variables M, viz. for every model M_1, M_2, …, M_n expressible in S_1, $F_1(M_n)$ is analytic in S_1. With the transition to S_2, the class of expressible models can be enlarged to M_{n+1}, …, M_m. In this case, $(M)F_1(M)$ is analytic in S_2 if for any model expressible in S_2, $F_1(M_m)$ is analytic in S_2. Now, it is clear that the range of the relation variable M changes with the move from S_1 to S_2. Thus, the set of possible relations constructible from the respective range of individuals, i.e. the range of the metatheoretic sentence "for all relations of the individuals" is more comprehensive for S_2 than for S_1. The higher-order quantifier ranging over models is enlarged here to a richer class of models. Note, however, that this transition to S_1 does not alter the truth conditions for the theory in the models expressed in S_1. Given our above specification of Carnap's notion of a model, it is clear that this extended range of quantification has not restricting effect on the truth conditions for the class of models expressible in S_1. Truth in model remains stable for each element of this class since the (extended) range of individuals of the language is effectively restricted to the field of the respective individual model. It is in this sense that the additional conditions in analyticity* simply holds that for any given model M_n, if $F_1(M_n)$ is analytic in S_1 then it is also analytic in S_2.

Given this, how does this account of treating model extensions via language transitions relate to our initial interpretive question concerning the domain variability in Carnap's conception of models? On first sight, the definition of analyticity* could be read as evidence in support of Hintikka's claim of a FDC for the models of a theory (and model extensions thereof) after all. An argument in this direction could go along these lines: Assuming (as Carnap allegedly did) a FDC concerning models, the only way to allow proper extensions of a model in the sense of extensions of the individual domain (and not simply as extensions of the existing relations on a fixed set of individuals) is to actually change the formal background language in which the theory is defined. Thus, assuming FDC for models, a change of individual is only possible by changing the underlying language.

However, this argument would severely conflict with Hintikka's original interpretation of Carnap's account. For Hintikka makes the following stronger claim concerning the latter's conception of extension: Carnap not only was unable to formalize mathematical standard model extensions in a proper way due to his FDC of the underlying language. He in fact conceived the very notion of model extensions in a nonstandard (and mathematically irrelevant) way. In this picture, there is simply no place for extensions of the individual domain in Carnap's account, since extensions are expressed in terms of relations (or structures) and not in terms of individuals (see again Hintikka 1991, 332). In his interpretation of Carnap's theory model extension *is* the extension of the relations (better relations' structures) over an invariant domain.

Hintikka does not provide any significant textual support for his view. In particular, he does not mention the concluding section of Carnap and Bachmann (1936). This is not altogether surprising since the very nature of the remarks made here on how model extensions can transcend the limits of a given language simply makes no sense in his notion of what counts as an extension for Carnap. If one were to follow Hintikka's picture of Carnap's account, the move to model extensions via language transitions would not be conceivable. In contrast, it makes perfect sense if one sides with the interpretation of Carnap's conception of models outlined in the present chapter.

2.6 Conclusion

Summing up, the principal aim in this chapter was to provide a detailed survey of Carnap's early conception of models in his work on axiomatics. Based on several documents from Carnap's *Nachlass,* I proposed and defended a new interpretation of Carnap's reconstruction of extremal axioms and his early contributions to formal semantics. Given the archival evidence as well as the systematic arguments for the heterodox 'domains-as-fields' version of models and the convention of model variation via type ascent, I showed that Hintikka's interpretive claims concerning the FDC in Carnap's early use of model cannot be upheld. I argued further that once a more balanced account of Carnap's conception of model extensions (and restrictions thereof) is provided, also Hintikka's second claim that Carnap, in his theory of extremal axioms, misconceived mathematic practice turns out to be untenable.

Chapter 3: Carnap's extremal axioms and his mathematical influences

3.1 Introduction

Recent scholarship on Carnap's early theory of axiomatics (Carnap (2000)) has focused on his distinction between three metatheoretic notions of completeness: categoricity ("*Monomorphie*"), semantic completeness ("*Nicht-Gabelbarkeit*"), and syntactic completeness ("*Entscheidungsdefinitheit*"). In particular, Carnap's somewhat misguided views on the equivalence of these notions prior to Gödel's and Tarski's metatheoretic results in the 1930s have been discussed extensively.[194] So far, less attention has been drawn, however, to a forth kind of completeness concerning the interpretations of an axiom system that plays a prominent role in Carnap's axiomatics, namely the "completeness of the models" ("*Vollständigkeit der Modelle*") (Carnap and Bachman 1981, 82). In Carnap's view, completeness in this sense describes a certain maximal property of the admissible models, i.e. the non-extendibility of the models of an axiom system. In modern terms, for an axiomatic theory T and the class of models M of T, this maximality assumption can be defined in this way:

Model maximality The models in M are *complete* if they are *maximal*, i.e. if the universe of each member of M cannot be extended without violating the axioms of T.[195]

As is well known, this kind of maximality property was first expressed in Hilbert's famous *axioms of completeness* (henceforth AC) for analysis and geometry at the turn of the last century (see Section 1). Hilbert's axioms played a central role in a broader discussion (involving mathematicians like Baldus, Baer, Finsler, Bachmann, and Bernays) on the kind of maximality conditions they impose on an axiomatic theory.[196] Interestingly, the completeness of the models has an inverse, i.e. a less extensively discussed minimal property that can be paraphrased in this way:

---

[194] See Awodey and Carus (2001), Reck (2007), Goldfarb (2005), Bonk and Mosterin (2000). For modern-style definitions of the different notions of completeness discussed by Carnap see in particular Awodey and Reck (2002).
[195] See Ehrlich (1997) for an alternative definition of this maximality assumption.
[196] Compare Awodey and Reck (2002).

Model minimality The models in M are *minimal* if the universe of each model in M cannot be restricted without violating the axioms of T.

As Hilbert's AC was introduced to enforce model maximality, comparable minimal or restriction axioms were devised to impose the minimality of the intended models of an axiomatic theory. Their history, however, is more intricate than in the above case. As we will see in Section 3.3 and Section 3.4.2, the first explicit minimal principles can be found in Dedekind's proto-axiomatic treatment of elementary arithmetic and, in succession, in Fraenkel's axiomatization of set theory.
Carnap, in his own metatheoretic work on axiomatics in the late 1920s, was among the first to notice the analogous character of the two completeness properties. In his resulting theory of extremal axioms – first developed in the projected second part of his *Untersuchungen zur allgemeinen Axiomatik* and then, in published form in Carnap and Bachmann (1936) – he set out to provide a unified account of the two concepts of model maximality and model minimality. Specifically, his theory was devised as a formal explication of the two groups of mathematical axioms that share the feature of imposing such "extremal conditions" on the possible models of an axiom system.

The aim in this chapter is to provide a detailed and historically sensitive reconstruction of Carnap's theory of extremal axioms and its theoretical evolution between 1928 and 1936. The main focus will be set on the unpublished documents of Part Two of *Untersuchungen* from Carnap's *Nachlass* (RC 081-01-01 to 081-01-33) where his account is developed in most detail. In particular, Carnap's theory outlined there will be surveyed with an eye to two interpretive points not properly considered so far. The first concerns Carnap's mathematical sources, i.e. is the mathematical axioms based on which his formal versions were modeled. Here, especially, the case for minimal axioms is not fully clear. Notably, Carnap mentions the (second-order) axiom of induction (henceforth AI) in Peano arithmetic as a mathematical instance of his version of minimal axioms several times. Given that the induction axiom is usually not conceived as an extremal axiom as are Hilbert's ACs, this might raise general doubts about his formal reconstruction. One critical objection would be to say that Carnap's proclaimed unificatory approach is in fact comparing the incomparable. In a second, more critical reading, one might claim that Carnap's formalized version of extremal axioms not only misrepresents mathematical axioms like AI that are not minimal in his specific sense, but

that he also failed to capture the informal conceptions of limiting conditions on a theory's domain expressed in the relevant mathematical axioms.[197]
The second interpretive issue concerns Carnap's specific understanding of the relationship between his "completeness of the models" and the other metatheoretic notions of completeness discussed by him, most notably the categoricity of an axiom system. As we will see in Section 3.5, Carnap's general account of the relationship outlined in Part Two of *Untersuchungen* can be embedded in a broader debate between Fraenkel, von Neumann, and Baldus in the late 1920s on whether the addition of an extremal axiom results in a categorical axiomatization of a theory. Here again the critical question can be raised whether Carnap's formal presentation of extremal axioms presents an adequate setting for deciding on the issue.

My attempt here is to approach these interpretive points with special attention to the historical setting of Carnap's formal reconstruction and, in particular, to the mathematical influences that led him to devise a theory of extremal axioms in the first place. The main point I want to make is that the *prima facie* peculiarities of his approach can be resolved once Carnap's main influence for modeling minimal axioms, namely Fraenkel's *axiom of restriction* (henceforth AR) is studied in closer detail. Fraenkel's relevance for Carnap's theory of axiomatics, especially his three notions of axiomatic completeness was already mentioned in recent literature.[198] In this chapter, I aim to show that his influence on Carnap's conception of axiomatics had been more substantial than was assumed so far. It not only concerns Carnap's notions of meta-axiomatic completeness and his formal reconstruction of minimal axioms but also Carnap's overall conception of the semantic notions such as a "model" and a "domain" of a model that underlie his theory of axiomatics.

The chapter is organized as follows: I start with a short outline of Carnap's theory of extremal axioms – as presented in Carnap and Bachmann (1936) – and of the mathematical examples Carnap and Bachmann mentioned there (Section 3.2). Following this, Fraenkel's axiom of restriction for set theory is presented (Section 3.3). In Section 3.4, I discuss Carnap's original theory of extremal axioms as documented in his *Nachlass* (Section 3.4.1) and see how it relates to Fraenkel's views (Section 3.4.2). The case of a minimal axiom in arithmetic is discussed in Section 3.4.3. Finally, Carnap's views on the relation between completeness of

[197] See Hintikka (1991) for a related critique. See Chapter 2 for a critical discussion of Hintikka's interpretation.
[198] See, e.g., Reck (2007) and Bonk and Mosterin (2000).

models and categoricity will be discussed in Section 3.5. After reviewing a limitative result by Baldus (Section 3.5.1), Carnap's more general account of the relationship will be evaluated in this light (Section 3.5.2).

3.2 Extremal axioms in 1936

We have mentioned in the second chapter that the only systematic treatment of Carnap's extremal axioms in published form is Carnap and Bachman (1936). The paper starts with a provisional characterization of minimal and maximal axioms by highlighting their respective domain-fixing functions:

> Axioms of the sort of Hilbert's axiom of completeness, which ascribe to the objects of an axiomatic theory a maximal property in that they assert that there is no more comprehensive system of objects that also satisfies a series of axioms, we call a *maximal axiom*. The same axiomatic role as that of the maximal axioms is played in different axiom systems by *minimal axioms* which ascribe a respective minimal property to the elements of the discipline. Maximal and minimal axioms we call collectively *extremal axioms*. (Carnap and Bachmann 1981, 68-69)

Based on this informal account (and given the specification of several auxiliary notions such as that of a formal models, a model's structure, and isomorphism correlation), a formal explication of different kinds of extremal axioms is presented in §4 of the paper. The distinction between different types of minimal and maximal axioms is based here on a ramification of the notion of a model extension.[199] In particular, Carnap and Bachmann distinguish between "model extensions" and "structure extensions". In the former case the original models are admitted to be isomorphically embeddable in extended model, in the latter they are not. Here, the condition that model extensions are non-isomorphic to the original model is explicitly "included in the definition of extension" (ibid, 77). The two notions are formally presented in this way (see ibid, 77): [200]

$$\mathrm{Erw}_{\mathrm{m}}(N; M) =_{\mathrm{df}} M \subset N \,.\, M \neq N$$

$$\mathrm{Erw}_{\mathrm{s}}(N; M) =_{\mathrm{df}} M \subset N \,.\, {\sim}\mathrm{Ism}_{\mathrm{v}}(M, N)$$

---

[199] Compare Chapter 2 for a detailed discussion of Carnap's conception of models and model extensions.

[200] A second ramification of the notion of extension Carnap and Bachmann do not work out in detail is based on the question on whether or not model extensions should allow extensions *M* of *N* where *M* contains a proper part *T* isomorphic to *N* (ibid, 177). Compare Ehrlich (1997) for discussion of a similar point.

Based on these two types of extensions, Carnap and Bachmann present two versions of extremal axioms: “maximal” and “minimal model axioms“ and “maximal” and “minimal structure axioms“ (ibid, 77):

$$(1)\ \mathrm{Max}_m\ (F; M) =_{df} \sim(\exists N)\ (M \subset N\ .\ M \neq N\ .\ \mathrm{F}(N))$$

$$(2)\ \mathrm{Max}_s\ (F; M) =_{df} \sim(\exists N)\ (M \subset N\ .\ \sim \mathrm{Ism}_v(M, N)\ .\ \mathrm{F}(N))$$

$$(3)\ \mathrm{Min}_M\ (F; M) =_{df} \sim(\exists N)\ (N \subset M\ .\ M \neq N\ .\ \mathrm{F}(N))$$

$$(4)\ \mathrm{Max}_S\ (F; M) =_{df} \sim(\exists N)\ (N \subset M\ .\ \sim \mathrm{Ism}_v(M, N)\ .\ \mathrm{F}(N)).$$

The axiom schemes (1) and (3) are more restrictive than (2) and (4): in the case of minimal axioms, (3) excludes any submodel of a given model *M* of a theory, also, if there are any, those isomorphic to *M*. (4), in contrast, only rules out non-isomorphic submodels of *M* (see ibid, 77). This distinction between structure axioms and model axioms will be relevant for one of the mathematical examples of a minimal axiom, namely the induction axiom, Carnap discusses in greater detail (see Section 3.4.3).

Irrespective of this, it is clear that the intended effect of all four axioms was to impose a kind of completeness on the axiomatic theory to which they are added. In fact, after having presented the formal versions, Carnap and Bachmann made explicit the supposed effect for the case of a maximal axiom:

> The models of an axiom system that is closed by a maximal axiom possess a certain completeness property in that they cannot be extended without violating the original axiom system. (ibid, 82)

This is precisely the “completeness of the models” of an axiom system, here the model maximality described in the Introduction.

Turning to Carnap’s mathematical influences, i.e. the actual axioms on which his formal reconstruction of extremal axioms is based, his source is obvious in the case of maximal axioms. In fact, a version of Hilbert’s axiom of completeness for geometry is cited at the outset of the paper and again mentioned in the passage quoted above. In particular, Carnap and Bachmann refer explicitly to the original version of the axiom of completeness for geometry in Hilbert's second to sixth edition of *Grundlagen der Geometrie* (ibid, 68).[201] As is well known, the axiom of completeness was conceived there as a means to impose the model maximality for the Hilbert's base axiom system, i.e. his four groups of geometrical axioms plus the Archimedean axiom. As Ehrlich points out, the axiom was introduced to "distinguish "ordinary analytic geometry" from the remaining models of Archimedean Euclidian geometry" and, thus, fix the Cartesian space over the continuous ordered field R, $C^2(R)$, as the intended, single model of the axiom system (Ehrlich 1997, 57).

Carnap and Bachmann explicitly held that Hilbert's two AC (both for geometry and analysis) could be explicated in terms of their maximal model axioms, in particular via axiom scheme (1).[202]

The mathematical influences for Carnap's minimal axioms are less obvious. As mentioned above, Carnap referred to "Peano arithmetic" or "elementary arithmetic" on several occasions as one axiomatic theories to illustrate his theory of minimal axioms (see e.g. ibid, 79-80). This might seem peculiar to the modern reader since it is - on first sight - difficult to see how the central axiom of the modern version of Peano arithmetic, namely AI

$$(\mathrm{AI}^2) \qquad (\forall X)([X(0) \wedge (\forall x)(X(x) \rightarrow X(S(x)))] \rightarrow (\forall y)(X(y)))$$

is related to the minimal axioms (3) or (4) presented in Carnap and Bachmann (1936). In fact, the closer look at AI and Carnap's axiom schemes makes evident the conceptual gap between the modern presentation of AI in second-order logic and Carnap's own reconstruction. In the induction axiom, the second-order quantifier ranges over arbitrary relations on a specified universe of a model in which Peano arithmetic is interpreted. The primitives '0' (for *zero*) and '*S*' (for '*successor*') are schematic non-logical constants interpreted in the same universe. In Carnap's version, in contrast, the higher-order quantifiers in the minimal axiom schemes are not ranging over arbitrary relations on a given model's domain, but over models as such. We saw before that models were conceived by Carnap as higher-level n-ary relations, more specifically as n-tuples of relations of a specified type. Moreover, all models were defined in a fixed universal interpretation of the formal background language.[203] There are thus no direct

---

[201] There, Hilbert states his axiom in this way: "To the system of points, straight lines, and planes, it is impossible to adjoin other elements in such a manner that the system thus generalized shall form a new geometry obeying all of the five groups of axioms I-V; in other words, the elements of the geometry form a system which is not susceptible of [proper] extension, if all of the stated axioms are preserved." (Hilbert 1900, 25) quoted from: Ehrlich (1995), 171)

[202] Moreover, since the intended model in Hilbert's axiom system for the Euclidian plane, viz. C^2(R), is "*non-dividable"*, i.e. it does not have submodels isomorphic to the maximal model, the axioms can also be symbolized by a maximal structure axiom (2) (see ibid, 178).

[203] Compare 2.4.3 for a closer discussion of this issue.

equivalents in Carnap's formalization to the schematic terms in AI above. Instead, the primitives of an axiom system are expressed by free or in the case of extremal axioms, by bound higher-order variables.[204]

Given these conceptual differences, it is difficult to see what led Carnap to consider AI as an instance of his minimal axioms.[205] Moreover, one might get the impression that in contrast to the case of model maximality, the issue of model minimality was not fully clear to Carnap. The point I want to make in the following is that this is not the case. In fact, the most fruitful way to see what exactly Carnap understood by model minimality will be to take a closer look at the second mathematical example he gives for a minimal axiom. This is Fraenkel's meanwhile neglected axiom candidate for set theory, the *axiom of restriction* ("*Beschränktheitsaxiom*") (AR). In Carnap and Bachmann (1936) there is only a one-line remark on the axiom: "The restriction axiom in Fraenkel's axiom system for set theory can be formulated as a minimal-model axiom." (ibid, 78) In Carnap (1930), AR is also mentioned as a minimal analogue to the axiom of completeness (ibid, 307).

Now, despite the lack of a more extensive discussion of Fraenkel's axiom candidate in Carnap's published work, AR exercised a formative influence on Carnap's understanding of minimal axioms and the kind of minimality condition they impose on an axiomatization. As I will show in Section 3.4.1, AR was (beside AC) Carnap's main motivation for providing a formal reconstruction of extremal axioms in first place. Moreover, a closer look at Fraenkel's different versions of formulating AR will help to better understand Carnap's own version of minimal conditions expressed in his minimal axioms and how they relate to the case of arithmetic (and AI in specific). So, before turning to Carnap's *Nachlass* to illustrate this line of influence, I will briefly present Fraenkel's own views on his axiom candidate.

3.3 Fraenkel's axiom of restriction[206]

In the 1920s, Fraenkel proposed two additional axioms for Zermelo's axiomatic set theory. The axiom of replacement has meanwhile become a standard axiom of ZF. The second, AR, was devised to express a minimality condition for any set model satisfying the axioms set up

[204] For a more detailed account of Carnap's conception of formal logical languages see Chapter 1.
[205] In fact, it seems as though Carnap himself was not quite certain of the status of AI. In his only published report on Part 1 of *Untersuchungen* (Carnap (1930)), a small section is also dedicated to his extremal axioms. Here, AI is mentioned not as a minimal axiom, but as an example of a maximal structure axiom. See Carnap (1930, 307).
[206] This section follows closely Schiemer (2010).

by Zermelo's original system Z.[207] His attempts to devise such a restriction clause on set models of Z varied in the course of his work on the topic and eventually led to different, partly independent versions of minimal axioms for set theory. In the following I present a brief reconstruction of the evolution of Fraenkel's thought on AR from this period.

3.3.1 Fraenkel's motivation

The first mention of AR can be found in (Fraenkel 1922). Here, Fraenkel presents two arguments why AR should be added to Zermelo's theory. He first states that "Zermelo's concept of set is more comprehensive than seems to be necessary for the needs of mathematics (…)." (ibid, 223) Fraenkel mentions two kinds of possible sets in the set-theoretic "domain" ("Grundbereich") that are consistent with the existing axioms, however irrelevant for mathematical purposes. The first are "non-conceptual" sets, e.g. sets consisting of physical elements. The second are so called "non- well-founded" sets, i.e. such with infinite membership chains. Given their possibility within Zermelo's axiomatization, Fraenkel remarked that:

> Whereas sets of the first as of the second kind are not necessary for set theory considered as a mathematical discipline, it in any case follows from the fact that they have a place within Zermelo's axiomatization that the axiom system (..) does not have a "categorical character", that is to say it does not determine the totality of sets completely. (ibid, 234)

In the second edition of his monograph *Einleitung in die Mengenlehre* (Fraenkel 1924) a more structured presentation of his motivation for adding AR is given. Fraenkel argued here that the introduction of the additional axiom would lead to a "simplification of the set theoretic edifice" by ruling out non-well-founded numbers. As an independent argument – the property of completeness in terms of categoricity is referred to:

> Moreover, without such a restriction it is not within reach that our axiom system captures the totality of admissible sets *completely* as is desirable for the construction of every axiomatization. (ibid, 218)

Thus, one can identify here at least two related but different objections against Z: the possible existence of extraordinary sets and its non-categoricity. Moreover, it should be stressed that

[207] For a more detailed account of Fraenkel's axiom candidate see Schiemer (2010).

Fraenkel explicitly expressed here the assumption concerning the connection between the intended restriction of the domain of sets and the categoricity of the resulting theory (compare Schiemer (2010, 312-313)). This point was further highlighted in his lectures from 1925. Here he remarked that:

It means more than a mere flaw of our axiom system that the totality of all possible sets is not unequivocally fixed but that instead there are always narrower and more comprehensive interpretations of the concept of set that remain compatible with our axiom system. (Fraenkel 1927, 101)

According to Fraenkel's view in 1924, the non-categoricity of Zermelo's original axiom systems is mainly caused by the possible existence of "extraordinary sets". As a solution to this, Fraenkel proposed his AR which was conceived in analogy to Hilbert's AC:

(…) As is the case there, the mentioned deficiencies can be remedied by setting up a (…) last axiom, the "axiom of restriction" that imposes on the concept of set or more appropriately the domain [of sets], the *smallest extension compatible with the remaining axioms.* (ibid, 234)

Fraenkel's motivation for adding AR to Z is clear: to rule out non-intended and non-well-founded sets by restricting either the interpretation of the concept of set or the domain of set. This would also make the resulting axiom system categorical (see Schiemer (2010, 315)).

3.3.2 Versions of restriction[208]

Fraenkel's remarks on the intended effect of his axiom candidate do not go beyond the level of informal remarks. The most detailed explication can be found in (Fraenkel 1922b). Here he formulates two alternative versions of AR for an axiomatization of cardinal numbers that also hold for the case of set theory.

(1) AR as an axiom of closure: in the first version, AR is defined as expressing a minimality condition on sets: There exist no sets in a model of ZF apart from the ones implied by the given axioms. A similar version of this understanding of AR can also be found in Fraenkel (1924): "Aside from the sets imposed by the axioms [of Zermelo (1908)] there exist no

[208] This section closely follows Schiemer (2010).

further sets." (Fraenkel 1924, 219)[209] The mathematical idea underlying this is that of a closure of a given bases system under a specific operation. This implies a view of the intended domain (or model) as an algebraic closure that Fraenkel first introduced to axiomatic set theory. According to it, the intended set domain (or set model) is the closure of a base (in Fraenkel's case the empty set, in Zermelo's case a certain *Urklasse*) under the operations on sets defined in the axioms of ZF. AR is then understood as expressing a closure property on the set domain, i.e. as a kind of "restriction clause for closures" (Kanamori 2004, 515).

(2) AR as a minimal axiom: Fraenkel's second reading sketches the intended effect of the axiom. According to it, the axiom imposes a minimal model for the axiom system: "If the domain (*Grundbereich*) B contains a smallest submodel (*Teilbereich*) B_0 satisfying the axioms (…), then B is identical with such a smallest submodel B_0." (ibid, 163) This effectively rules out the existence of proper submodels of B_0 that satisfy ZF. In an added footnote, Fraenkel got more specific on the method of constructing such a minimal model:

> As is usual, a smallest submodel of the indicated character is to be understood as a model that is the intersection of all submodels of B with the property in question and that also possesses the property itself. (ibid, 163)

This states that the minimal model B of ZF can be constructed by the intersection of all possible models that satisfy the theory. Moreover, if such a minimal model exists, then the extended theory ZF+AR will result to be categorical (see Schiemer 2010, 316). (I will return to the last point in Section 3.5 below.)

3.4 Part 2 of *Untersuchungen*

As I mentioned above, the formative influence of Fraenkel's AR on Carnap's own theory of extremal axioms is not intelligible from Carnap and Bachmann (1936) alone. Neither is it traceable in the first part of *Untersuchungen* (Carnap 2000). Here, Carnap sets his main focus on the three metatheoretic notions of completeness (see Introduction) without mentioning the "completeness of the models" and his extremal axioms. However, things look different taking into consideration Carnap's unpublished material for the projected second part of his

[209] Compare also the third edition of *Einleitung* (Fraenkel (1928) for a similar definition.

manuscript on axiomatics in the *Nachlass* (RC 081-01-01 to -33). These documents clearly prove that Carnap was working on a general account of extremal axioms already in 1928. Moreover, they show that most of the concepts and results listed in the 1936 paper with Bachmann can be traced back to this earlier work. Finally and most relevant for our account, the importance of Fraenkel's AR (and the way he conceives its effect on the models of ZF) for Carnap's theory becomes evident from the existing documents of Part 2 of *Untersuchungen*. Generally, one can say that Carnap's account of extremal axioms in 1928 is more ramified than in 1936. In a "preliminary work plan" for Part 2 he mentions different types of axioms that yield the completeness property described above. For the first group, the "minimal" and "maximal axioms", Fraenkel's AR is mentioned as the main example of the former (RC 081-01-02). Beside these, Carnap mentions a different but related group of axioms dubbed "accessibility axioms" ("*Erreichbarkeitsaxiome*"). And here again, AR (next to Peano's AI) is mentioned as a typical mathematical instance.[210] The next section will be used to present Carnap's different groups of extremal axioms and to show how the distinctions he makes between them closely reflect Fraenkel's different versions of AR.

3.4.1 Different types of minimal axioms

Fraenkel's influence becomes visible in the way Carnap describes the first group of minimal axioms in his manuscript. In (RC 081-01-04) he presents a general account of "axioms of restriction" (Not surprisingly, their maximal counterparts are dubbed "axioms of completeness.") The following two informal definitions are given:

An axiom of restriction means fixing the minimal extension: the domain of elements is to have the smallest possible extension compatible with the other axioms. (RC 081-01-04/1)[211]

The axiom of restriction can be formulated in this way: „There exists no (proper) part of a system of elements (a model or a realization) that satisfies … the axioms. (RC 081-01-04/2)[212]

---

[210] Carnap's "preliminary work plan" is the single document of Part 2 of *Untersuchungen* that is published in Carnap (2000). In their commentary, the editors Bonk and Mosterin hold that "Fraenkel's suggestion [GS: of adding AR to ZF] (…) was probably the starting point for Part 2 of *Untersuchungen*" without, however, discussing any details of this line of influence that can be witnessed in Carnap's documents (Bonk and Mosterin (2000, 48)

[211] "Ein Beschränktheitsaxiom bedeutet Festlegung des Minimalumfangs: das Gegenstandssystem soll den kleinsten mit den übrigen Axiomen verträglichen Umfang haben." (RC 081-01-04/1)

[212] "Das Beschränktheitsaxiom kann so formuliert werden: „Es gibt keinen (echten) Teil eines Gegenstandssystems (Modells oder einer Realisation), der die Axiomen … erfüllt." (RC 081-01-04/2)

The formal versions of an axiom of restriction presented here are more or less equivalent to (3) in Carnap and Bachmann (1936). In addition, a slightly different formalization is suggested. For an axiom system f with only one primitive relation P, the axiom reads:

$$(P): P \subset R. \, \mathrm{f}P. \supset. \, P = R$$

Carnap paraphrased this version in this way: "In words: a) every partial relation of *R* that satisfies the other axioms coincides with *R*." And in a footnote he adds: "This formulation is only suited as a verbal description of the logistic formulas; otherwise the earlier formulations are more easily comprehensible." (RC 081-01-04/2)[213]

Now, Carnap's informal definitions clearly echo Fraenkel's characterization of his set-theoretical AR. In fact, Carnap's first reading of the axiom as "determining a minimal extension" that is consistent with the other axioms is an almost a verbatim adoption of Fraenkel's talk of "the *smallest extension compatible with the remaining axioms*" quoted above (Fraenkel 1927, 234). Moreover, there is also a strong similarity in the way the models of an axiomatic theory are conceived in the two cases. Following Carnap's second definition, an axiom of restriction imposes the minimal model as the single admissible model of an axiom system. In his third definition, closely following the "logistic formula", a minimal model is conceived as the minimal partial relation of the model relation *R*. We have already indicated above that Carnap's conception of formal models is clearly heterodox from a modern point of view since the models of a theory are conceived as (sequences of) relations. Submodels are thereof presented as partial relations of a given (model) relation.[214] What is important to see here is that this conception of submodels as partial relations is already present in Freankel's own treatment of AR (and of set models more generally). Note for example a section in Fraenkel (1928) on the general use of "models" and "pseudosystems" ("*Pseudosysteme*") of an axiom system for indirect consistency or independence proofs. Models are explicitly defined here as „interpretations differing from the usual interpretation of the primitive concept" ("*vom üblichen abweichenden Deutung gewisser Grundbegriffe*") (ibid, 342). The standard method of devising restricted models for consistency or independence proof is described in this way:

[213] "In Worten: a.) jedes (sic!) Teilrelation von R, die die anderen Axiome erfüllt, fällt mit R zusammen. (…) Diese Formulierung eignet sich nur zur Wortbeschreibung der logistischen Formeln; sonst sind die früheren Formulierungen leichter verständlich." (ibid.)

[214] For an extensive study of Carnap's nonstandard treatment of submodels compare again Chapter 2.

> In many cases the interpretation in question more or less restricts the extension of a primitive concept vis-à-vis its usual interpretation for the whole axiom system so that only a part of the objects originally falling under the primitive concept is accepted as existing. (ibid, 342)[215]

Model variation is thus conceived here as the modification of the standard or intended interpretation of the axiomatic primitives of a theory. In the case of set theory, this means that, for different independence proofs, the concept of set (or, equivalently the membership relation) is reinterpreted in the sense that it is "understood differently, generally in a narrower sense than in the "proper" set theory following form our axiomatic (system)." (ibid, 344) Applied to Fraenkel's version of AR for set theory, this effectively means that the intended model of ZF+AR is conceived as the minimal "interpretation" of the "basic relation" $\in$ defined by ZF. Now, it is exactly this view of a minimal model as the minimal partial relation of the axiomatic primitives that is also expressed in Carnap's third formal version of his axioms of restriction in 1928.

The parallels with Fraenkel's account do not end here. As mentioned above, Carnap presents a second kind of extremal axioms in his manuscript, the so-called "accessibility axioms" he considers as "related with (possibly a special case of) the axiom of restriction" (RC 081-01-10). An accessibility axiom presupposes "absolute" and "relative existence axioms" which state that "there exist elements in the system with these and that properties ("*Urelements*")" and that "if elements exist of this and that property then there also exist elements of that and those properties" respectively. Given this, an accessibility axiom is defined in this way:

> An accessibility axiom then states: every element of the system is either an element (according to the absolute existence axiom) or its existence can be deduced in finitely many steps from the elements with the help of the relative existence axiom. (ibid.)[216]

The underlying idea here is that an axiom system T fixes a certain "*accessibility class*" ("*Erreichbarkeitsklasse*") of elements constructible from a certain base set (imposed by the absolute existence axioms of T) via the iterative application of operations (expressed in the

[215] "In vielen Fällen schränkt die fragliche Deutung den Umfang eines Grundbegriffs gegenüber der üblichen Deutung im gesamten Axiomensystem mehr oder weniger ein, so dass nur mehr ein Teil der ursprünglich unter jenen Grundbegriff fallenden Objekte (…) als existierend zugelassen werden." (Fraenkel 1928, 342)
[216] "Ein EA besagt dann: jeder Gegenstand des Systems ist entweder selbst eine Gegenstand (gemäß dem aExA) oder seine Existenz ist aus der der Gegenstände mit Hilfe des rExA in endlich vielen Schritten ableitbar." (ibid.)

relative existence axioms of *T*). The accessibility axiom then secures that this class is also the single intended model of *T*. It states, in Carnap terms, that "every element is accessible" (RC 081-01-10/3).

Here again, the set-theoretic context of his reconstruction is quite definite. Note first that Carnap's discussion of absolute and relative existence axioms clearly reflects Fraenkel's own presentation of the standard set theoretical axioms in Fraenkel (1928). There, ZF includes a number of "unconditional existence axioms" (e.g. the axiom of infinity) and "conditional existence axioms" (e.g. the axioms of union, power set and choice).[217] Second and more generally, Carnap's picture of a theory's domain as an accessibility class closely mirrors Fraenkel's newly introduced conception of the set universe as an algebraic closure of a base set (in his case the empty set) under the operations codified in the axioms of ZF. Note that Carnap explicitly speaks of a "basic class" ("*Urklasse*") consisting of "*Urelements*" and of an "order of classes" ("*Stufenreihe von Klassen*") as a theory's domain. The latter begins with the basic class and yields higher and higher levels by the reiterated application of an "order-relation" ("*Stufenrelation*") that defines the group of operations specified in the relative existence axioms. The "union of these classes" is then the accessibility class; its elements are the "accessible" elements of a theory (RC 081-01-10/2). It is precisely this idea of a theory's domain as the accessibility class that one can also find in Fraenkel's closure conception of set models. Moreover, Carnap's version of an accessibility axiom as expressing the (closure) condition that "every element is accessible" clearly corresponds to Fraenkel's first version of devising AR. Recall from Section 3.2 that AR was devised there as a minimal condition on admissible sets. The intended domain of ZF+AR was conceived as the restriction to the sets introduced by the absolute existence axioms and by the "finite application" of the relative existence axioms of ZF.[218]

Given this obvious background in axiomatic set theory, it is quite surprising that Carnap, instead of explicitly discussing AR, draws to "Peano arithmetic" to further illustrate his theory of accessibility axiom (RC 081-01-10/3).[219] Recall that in the "preliminary work plan" for

---

[217] See Fraenkel (1928, §16)
[218] Compare Fraenkel (1928, 355).
[219] In Carnap (1934) Fraenkel's AR is extensively discussed as an accessibility axiom: "The axiom of restriction states that in the domain of sets treated in [the axiom system] *S*, say *B*, only those sets are considered whose existence is imposed by the other axioms." (Carnap 1934, 281)
„Das Beschränktheitsaxiom besagt, dass in dem in S behandelten Mengenbereich, etwa *B*, nur diejenigen Mengen vorkommen, deren Existenz durch die anderen Axiome gefordert ist." (Carnap 1934, 281)

Part 2 of *Untersuchungen* “Peano’s axiom of induction“ is juxtaposed to AR as an instance of an accessibility axiom. What was Carnap’s motivation for treating these two axioms as functionally similar? In order to approach this question let us make a step back and have a second look at Fraenkel’s own conception of AR and his motivation for setting it up in the way he did.

3.4.2 Fraenkel’s arithmetical background[220]

It is not completely clear from Fraenkel’s remarks on AR mentioned in Section 3.3.2 how he conceived the intended restrictive effect of his axiom on the possible set models of ZF. Nevertheless, there exists an insightful passage in Fraenkel (1928) that helps to see his mathematical background for AR. Fraenkel mentions here the “special character” of the axiom compared to the “existential” and “relational axioms” of ZF and observes the fact that “in both versions [of AR], the inductive moment is essential.” (Fraenkel 1928, 355) What is his intuition about the inductive character of AR here? As we have already seen, the concept of intersection plays a central role for the intended effect of the axiom candidate. A minimal model is conceived as the intersection of all possible models satisfying ZF. From a methodological point of view, this is a paring down-approach of defining a specific minimal structure by taking the intersection of all closed subsets of a given set.

There is a direct link between Fraenkel’s use of this method and the early understanding of induction principles in arithmetic. As is generally known, the paring down method was productively used by Dedekind to express the principle of mathematical induction as well to fix axiomatically the standard model of arithmetic.[221] I make a case in Schiemer (2010) that Fraenkel’s conception of a minimal model for set theory was very likely shaped in close analogy to Dedekind’s strategy of defining the natural numbers as a minimal set closed under induction. More specifically, I show that there is a striking similarity between Fraenkel’s treatment of AR and Dedekind’s “*theory of chains*” (“*Kettentheorie*”) introduced in Dedekind (1888) (see Schiemer (2010, 316-321)). This suggests that Fraenkel actually modeled his idea of set-theoretic restriction based on Dedekind’s approach.[222] Here, my aim is not to retrace this line of influence but to offer an explanation why Carnap, in formally reconstructing AR,

---

[220] This Section is based on Schiemer (2010).
[221] Compare Sieg and Schlimm (2005).
[222] See Schiemer (2010) for further details of this comparison between Dedekind’s approach to arithmetic and Fraenkel’s conception of set theory.

also perceived a methodological affinity between Fraenkel's axiom and the arithmetical axiom of induction. It is for this reason that a brief look at Dedekind's approach will be instructive.

Dedekind's foundational treatment of the natural numbers in Dedekind (1888) is based on a number of methodological results concerning the central concepts that allow the reduction of numbers to a logical basis (ibid, 351). One of these newly introduced concepts allowing him to devise the sequence N of the natural numbers is that of a *chain* (relative to mapping function φ and a system S): in modern terminology, a subsystem B of S is called a *chain* if it is closed under a mapping φ (ibid, 352). A system A_0 is defined as the *chain of A* ("*Kette des Systems A*") if A_0 is the intersection of all chains containing A (ibid, 353). What is important to note here is that Dedekind also conceives A_0 (give this underlying intersection of closures approach) as the smallest chain containing A, i.e. the smallest subset of S closed under φ. Again, in modern terminology, this effectively says that A_0 is the minimal closure of A under φ (compare Schiemer (2010, 317).[223]

Given this concept of a minimal chain, a set of axiomatic conditions is set up by Dedekind to characterize "simple infinite systems," i.e., systems isomorphic to the system of natural numbers (1, N, φ). One of the four conditions, condition β, states that the admissible system has to be a minimal chain containing the base 1: $N = 1_0$ Thus, in Dedekind's account of the natural numbers, 1 is the base element and the sequence N is the intersection of all sets containing 1 that are closed under the successor operation. Note the difference of this early axiomatic presentation of arithmetic to standard (second-order) Peano arithmetic: In Dedekind's version, principle β is used as an alternative to an explicitly formulated induction axiom. Instead of introducing AI as an axiom, Dedekind proves the principle of mathematical induction with the help of his concept of a minimal chain.[224]

It is this minimal chain condition expressed in Dedekind's original axiomatic system of arithmetic (and the paring down method underlying it) that most likely motivated Fraenkel's own axiom of restriction for set theory. The alleged "inductive character" of AR thus can be best explained with reference to Dedekind's original indirect treatment of arithmetical

[223] As Sieg and Schlimm (2005) point out, Dedekind is not explicit about the minimality of chains of A in Dedekind (1888). There exists, however, have pointed out, a note in Dedekind's earlier unpublished manuscript "Gedanken über Zahlen" in which this issue is explicitly mentioned: "(A) [i.e. the chain of A] is the "smallest" chain that contains the system A" (quoted from Sieg and Schlimm 2005, 144). Compare again Schiemer (2010) for a fuller discussion of this point.

[224] Dedekind (1888, 361); compare Sieg and Schlimm (2005, 145).

induction via minimal chains. Moreover, Fraenkel's background in Dedekind arithmetic will also help us to understand Carnap's more general remarks about the role of minimal axioms in arithmetic in a proper way.

3.4.3 Carnap's versions of arithmetic.

Our initial question was: Why does Carnap repeatedly refer to arithmetic for illustration of his extremal axioms despite the fact that the AI obviously has a different structure than his formal reconstructions of minimal axioms? The point to see is that Carnap – just as Fraenkel – is in fact *not* considering Peano arithmetic (in its modern canonical form) when speaking of "basic" or "elementary arithmetic". Instead he has in mind alternative axiom systems including such related to Dedekind's β that *are* structurally comparable to his account of minimal axioms. Thus, even though Carnap mentions AI as an instance of accessibility axioms in particular and of minimal axioms more generally, it is in fact not the main focus in his discussion of the axiomatization of basic arithmetic.

In the notes on Part 2 of *Untersuchungen* there are two versions of a "Peano axiom system", both including alternatives to the axiom of induction. In Carnap's published work up to Carnap and Bachmann (1936), most importantly his logic manual *Abriss der Logistik* (Carnap 1929), one can identify all in all three alternative versions of arithmetic:

(a) (Second-order) **Peano arithmetic**: An "original form" of Peano arithmetic with three primitives (*Nu*, *Za*, *Nf*) standing for '*Zero*', '*Natural Number*', and '*Successor*' respectively. The axiom system includes a second-order version of the induction axiom (see Carnap 1929, 74).

(b) **Ancestral arithmetic**: The axiom system defines a single primitive relation (*Vg*) (for "antecessor" ("*Vorgänger*")). An accessibility axiom (AA) imposes that "every number is accessible from the base element (the Zero) in a finite number of steps (…)." (ibid, 75). Carnap's formal presentation of the axiom in *Abriss* makes central use of the concept of the "*succession* of *Vg*" ("*Vg-Nachkommenschaft*"):[225]

(AA) $\mathrm{E}!B'Vg. \supset. Za = Vg^{*}{}'B'Vg$

[225] Compare Carnap (1929, 34-35).

It can be paraphrased by saying that the set of natural numbers *Za* is the “succession of *Vg*” from the base element *B* of *Vg*. Translated in modern terms, this simply states that the set of natural numbers is the transitive closure of the base element under *Vg*.[226] Essentially the same version of the axiom system can also be found in (RC 081-01-10).[227]

(c) **Basic arithmetic**: A version of “Peano’s AS (without the induction axiom)” where the accessibility axiom in version (2) is substituted by a minimal structure axiom in the sense specified in Section 2 above. This axiom system can be found in (RC 081-01-11). It is also discussed more extensively in Carnap and Bachmann (1936, 79-80).

Now, Carnap holds that the three versions are equivalent axiomatic presentations of the same theory of the natural numbers. More importantly, for him, there are direct correlations between the axioms of the respective systems: the accessibility axioms in (b) “corresponds” to the induction axiom in (a) (Carnap 1929, 75). The axiom expressing a closure condition via the notion of “succession of *Vg*” in (c) corresponds to the union of the accessibility axiom and two other axioms in (b) (stating the existence of a base element of *Vg* and the endlessness of *Vg* respectively). Finally, from Part 2 of *Untersuchungen*, it is clear that Carnap takes the accessibility axiom in (b) to be equivalent to a minimal structure axiom used in (c) (see RC 081-01-11). So, in Carnap’s understanding, the induction axiom, the accessibility axiom, the closure axiom and the minimal structure axiom have (more or less) equivalent roles in the axiomatization of basic arithmetic.

Three points have to be emphasized here. First, it is clearly in connection with ancestral arithmetic that the status of AI as an accessibility axiom becomes intelligible. In fact, one can understand AA as conceptually motivating the induction axiom. The idea expressed in AA is that the intended model of ancestral arithmetic is gained by the transitive closure of Zero under successor operation. On closer inspection, this is of course also implicit in AI:

$$(\forall X)([X(0) \wedge (\forall x)(X(x) \rightarrow X(S(x)))] \rightarrow (\forall y)(X(y)))$$

[226] Compare Smith (2008) for a modern presentation of ancestral arithmetic. Here, the accessibility axiom is defined for the signature (0, *S*) (for ‘Zero’ and ‘Successor’) and reads: $(\forall x)(x = 0 \vee S^*0x)$ (see ibid, 4).

[227] In *Abriss*, there is a forth, “most simple” version of arithmetic, also with a single primitive relation (*Vg*), but with only two axioms. The first axiom states that *Vg* is 1-1. The second axiom effectively states that that domain of *Vg* is the closure of the base element under *Vg* (see Carnap 1929, 75).

The last part of the formula, i.e. the consequent $\forall y(X(y))$, expresses that all elements of the intended domain hold in a set (given that it contains Zero and is closed under Succession).[228] Thus, any subset A of a model's domain for PA that is assigned to X that satisfies these two conditions in the antecedent is also the intended domain of PA.

Second, note that in the light of these equivalences of the alternative axiomatic systems the minimality of the intended model of Peano arithmetic becomes clear. In particular, the comparison with the equivalent versions of AI makes evident its intended role in the axiomatization: to fix the system of natural numbers as the minimal subset of all sets closed under induction.[229] In all four of Carnap's axiomatic systems this minimal property of the standard model of basic arithmetic is implied: via accessibility and an implicit minimal closure condition in (b), via containment in all inductive sets in (a), and finally via an explicit minimality condition expressed in the minimal structure axiom in (c).[230]

Carnap's discussion of the alternative axiomatic systems of arithmetic actually explains why he conceived AI as a minimal axiom despite the obvious differences described above. We indicated that Fraenkel, when mentioning the "inductive character" of AR, did not think of AI but rather of something like Dedekind's axiomatic condition β: $N = 1_0$. A similar point holds for Carnap. His immediate conceptual background for his axiomatic presentation of arithmetic in (b) and the closure condition expressed therein is not Dedekind's notion of a chain. Nevertheless, it is built on something essentially similar, namely Frege's concept of an "ancestral" of a relation introduced in his *Begriffsschrift*[231] and, more directly, on the notion of an arithmetical "progression". Carnap mentions explicitly in *Abriss* that his version (b) is effectively based on Russell and Whitehead's formal notion of a "progression" from *Principia Mathematica* §122 (see Carnap 1929, 75).[232] Now, it is generally known that the closure condition expressed in Russell's progressions as well as in Frege's ancestral of a relation is in

---

[228] The same point was recently stressed by Smith when discussing the relation between PA and ancestral arithmetic: "Now note, however, the crucial assumption here: the successors of zero are the only natural numbers. This assumption evidently underlies our acceptance of induction. (…) it is an idea which has to be available to anyone who fully grasps PA." (Smith 2008, 3)

[229] Compare Hintikka (1986) on this point.

[230] The minimal structure axiom explicitly adds to the picture the additional condition that this model minimality holds for the intended models of arithmetic up to isomorphism.

[231] See Frege (1879, Prop.76)

[232] Progressions are defined in *Principia Mathematica* as a specific type of series, i.e. as a class of all inductive cardinals whose "generating relation" satisfies the following conditions: (i) being 1-1; (ii) having a first term not included in range of R; (iii) being denumerably infinite; (iv) being closed under its "generating relation" *R* (Russell and Whitehead (1962, 245). The intention behind this last closure condition is to exclude those 1-1 relations that result in independent chains of elements, i.e. infinite series with a serial number other than ω (see ibid, 245-248).

fact identical to the one expressed in Dedekind's notion of a chain.[233] There exists, however, a subtle difference between the three approaches to arithmetic that is relevant here. Dedekind, as we have seen, explicitly mentions that the chain 1_0 is the smallest of possible chains. This minimality property is not made explicit in the other two accounts that lay the ground for Carnap's axiom system (b). In particular, no explicit mention was made by Frege or Russell of the model minimality implied in their versions of ancestral arithmetic. It was for this reason that Carnap - in connecting the dots - introduced the third version (c) including an explicit minimal axiom in order to emphasize the model minimality already implicit in his predecessors' formal theories of arithmetic.

Finally, with the *Russellian* (and *Fregean*) background of Carnap's conception of arithmetic in mind, the apparent conflict between the specific "metatheoretic" character of his minimal axioms and his mathematical sources (specifically AI) can be dissolved. The axioms (or axiomatic conditions) on which Carnap's formal reconstruction is based are unlike AI clearly metatheoretic from a modern point of view. They contain, as do Carnap's minimal axioms, a minimal closure approach that is nowadays expressed in a theory's metatheory and not in the axioms themselves.[234] Thus, Carnap's minimal axioms (as his mathematical sources) can be criticized for their anachronistic metatheoretic character, i.e. mainly for quantifying over possible models of a theory.[235] They cannot, however, be criticized for not adequately representing their mathematical sources. A closer look at the actual influences for his formal reconstructions - namely Freankel's AR and a *Russellian* interpretation of Peano arithmetic - has shown that Carnap is in fact capturing the informal conceptions of limiting conditions on a theory's domain expressed there quite closely. Thus, the possible claim mentioned in the introduction that Carnap's theory of extremal axioms fails to capture actual axiomatic practice has shown to be untenable.

I turn now to the second interpretive issue mentioned at the outset of the chapter, the question on Carnap's views on the relation between "completeness for models" and "monomorphism" and on how they relate to a larger then contemporary debate on the issue.

---

[233] See e.g. Demopoulos and Clark (2005, 140).

[234] See Shapiro (1991) for further discussion of the use of minimal closures in Dedekind's work (ibid, 98) as well as in Frege's ancestral arithmetic (ibid, 129).

[235] See Hintikka (1991, 331-332) and, for an alternative view, Shapiro (1991, 185-186). For an early critique of this nature see Baldus (1928).

3.5 Extremal axioms and categoricity

A central motivation for Carnap's discussion of minimal axioms (and extremal axiom in general) in *Untersuchungen* was to show that the addition of such an axiom results a categorical (in Carnap's terms "*monomorphic*") axiomatization.[236] In the case of Peano arithmetic, adding an accessibility axiom or a minimal axiom to the base system enforces model minimality on the possible interpretations. In Carnap's view, this kind of completeness of the models also implies the categoricity of the resulting system. We saw in Section 3.1 that a similar motivation for introducing a minimal axiom to the axiom system of set theory can also be traced in Fraenkel's discussion of his AR. Recall that in Fraenkel (1924) and later on, the axiom candidate is introduced in order to yield an axiomatization that "captures the totality of admissible sets *completely* (…)," where completeness is understood as categoricity (ibid, 218).[237] Fraenkel never made an attempt at a formal categoricity proof for his set-theoretic system. Nonetheless, it is clear from his informal remarks that for him the minimality of the set models would imply the categoricity of ZF+AR. Here, as with Carnap's example of basic arithmetic, Fraenkel viewed the categoricity of the resulting system of axiomatic set theory was an immediate result of adding AR to ZF.

More generally, I think it is save to say that in the theoretical accounts of formal axiomatics up to the 1920s, the categoricity of a theory was conceived as a direct consequence of the completeness of models expressed by extremal axioms. Nevertheless, in the second half of the decade a more refined discussion of these two notions of completeness set in (with the participation of von Neumann, Zermelo, Skolem, Finsler, Bernays, Bear, Baldus, and Carnap among others). The main focus in this debate was dedicated to Hilbert's AC and Fraenkel's AR and their effects on the respective axiomatic theories.[238] However, to a lesser extent, also the more general logical relation between extremal axioms (and the kind of completeness imposed by them) and the categoricity of the mathematical theories was discussed. In this respect, two contributions stand out. One is Carnap's own attempt to characterize and classify extremal axioms in the *Untersuchungen* of 1928 that was least partly motivated by the wish to clarify their effect on an axiomatic system. The second explicit discussion of this point can be

[236] In (RC 081-01-04), after stating the informal versions of restriction and completeness axioms, he mentions that "both happens to yield monomorphism." Compare Reck (2007, 195-196).

[237] For Fraenkel's treatment of three types of completeness, i.e. semantic completeness, syntactic completeness and categoricity, as well as their relationship and later reception, by Carnap, see Awodey & Reck (2002).

[238] For the immediate, mostly critical reception of Fraenkel's axiom candidate see Schiemer (2010).

found in an article of the mathematician Richard Baldus also dated from 1928. In the remaining two sections, their contributions to this metatheoretic question will be compared.

3.5.1 A limitative result by Baldus

One central limitative result on the general impact of extremal constraints on the models of a theory was developed in Baldus (1928). The main target of Baldus' discussion is Hilbert's geometrical AC in *Grundlagen der Geometrie*.[239] Baldus discusses different versions of Hilbert's axiom (both the original version quoted in Section 3.2 above and the axiom of line completeness introduced in a later edition of the book). He shows that the later axiom of line completeness can be substituted by an equivalent continuity (or *Cantorian*) axiom.[240] A theory consisting of Hilbert's axioms groups I to IV, the Archimedean axiom, the continuity axiom but not the (Euclidian) axiom of parallels is termed (following Bolyai) "absolute geometry". In his paper, Baldus presents two important metatheoretic results on the axiomatic theory of absolute geometry. First, it is shown that the general completeness of models can be proved as a meta-theorem of the axiom system. In Baldus' own terms:

> Theorem of completeness. It is not possible to add in thought any kind of things ("points," "lines," "planes") to an interpretation of the axioms of absolute geometry (…) so that the axioms again hold in the extended interpretation and that the axiomatic relations between the elements (points, lines, planes) of the original interpretation are preserved under the extension. (ibid, 324)[241]

This completeness theorem in fact expresses the same kind of model maximality as Hilbert's AC does. Baldus motivation for his reconstruction is this: By substituting Hilbert's 'meta-axiom' by a less problematic axiom of continuity, the maximality constraint on the possible models of the theory does not have to be expressed within the theory itself. It can instead be transposed to the (semantic) metatheory of the axiom system.[242]

---

[239] Beside the discussion of AR, there are also a number of critical side-remarks on Fraenkel's AR, in particularly on its problematic metatheoretic character. For a fuller discussion see Schiemer (2010). Compare also Hintikka (1991).
[240] See Baldus (1928, 323).
[241] "Vollständigkeitssatz. Es ist nicht möglich, zu einer Deutung der Axiome der absoluten Geometrie (…) irgendwelche Dinge („Punkte", „Geraden", „Ebenen") so hinzuzudenken, dass in der erweiterten Deutung wieder die Axiome der absoluten Geometrie gelten und dass die axiomatischen Beziehungen der Elemente (Punkte, Gerade, Ebene) der ursprünglichen Deutung bei der Erweiterung erhalten bleiben." (ibid, 324)
[242] Note that this idea goes further than Hilbert's (and Bernay's) own approach in later editions of *Grundlagen* to substitute AC by an (still 'metatheoretic') axiom of line completeness and present full completeness as a theorem. Compare Awodey and Reck (2002, 12-17) on this point.

It is Baldus' second and main result concerning absolute geometry that clearly limits (or better, puts into proper perspective) the previously conceived relation between completeness of models and the categoricity of a theory. In Hilbert's axiomatization of Euclidian geometry the maximality assumption expressed in AC implies the categoricity of the theory. Baldus shows that this consequence does not hold in the case of absolute geometry. Here model maximality (in Baldus' specified sense as the non-extensibility of the domain of points) does no suffice to effect categoricity. Thus, Baldus (as far as I know) was the first to show in print that the two kinds of completeness – model maximality and categoricity - are not generally equivalent notions:

> An axiom system can be complete in the sense of the axiom of completeness without that all its interpretations are isomorphic, i.e. there exists a 1-1 correlation between them that preserves the axiomatically fixed concepts and relations. (ibid, 328)[243]

There can be cases, he argues, in which two models of a theory are both maximal in terms of the set of points but where a 1-1 correlation between the respective domains does not preserve the structure of the models (see ibid, 327-328). Baldus' result is highly intuitive in the light of non-Euclidian geometries, in particular hyperbolic geometry. The axiom system of absolute geometry was in fact originally set up by Bolyai as a neutral base theory to which different inconsistent axioms of parallels could be added, thus yielding different non-Euclidian theories.[244] Baldus shows that the completeness property for models holds for any theory that includes absolute geometry, i.e. Euclidian geometry or hyperbolic geometry, but not elliptic geometry. In both cases the continuity axiom fixes that the possible models have as the domain Cartesian space over the continuous ordered field *R*, viz. $C^2(R)$.

Model maximality is thus a necessary, but not a sufficient condition for the categoricity of a theory. Only the addition of an axiom that provides a unique characterization of the space defined by absolute geometry would render the resulting theory categorical (see ibid, 329). Since the possible axioms for this role, namely the Euclidian axiom of parallels and a hyperbolic alternative are inconsistent with each other, i.e. impose inconsistent structures on

---

[243] "Ein Axiomensystem kann im Sinne des Vollständigkeitsaxioms vollständig sein, ohne dass dessen sämtlich Deutungen isomorph sind, d.h. unter Erhaltung der axiomatisch festgelegten Begriffe und Beziehungen elementweise ein-eindeutig aufeinander beziehbar." (ibid, 328)
[244] Compare also Torretti (1984, 66).

its models, the base theory of absolute geometry underlying both theoretical extensions could simply not be categorical.[245]

The general moral to be drawn from these counterexamples presented in Baldus (1928) was that model maximality (and minimality thereof) is not equivalent to categoricity. More precisely, it is the conditional from model maximality to categoricity that does not hold in general. Constraints in addition to the extremal constraints concerning the possible models of a theory have to be specified so that the introduction of an extremal axiom yields a categorical theory. Baldus, apart from his discussion of the two examples of a complete, non-categorical theory did not get more specific on how these conditions can be understood in a more general setting. Here, Carnap's *Untersuchungen* enters into the picture.

3.5.2 Carnap's understanding in 1928

It was already mentioned that one of Carnap's main motivations for his formal reconstruction of extremal axioms was to get a clearer picture of the logical relations between the different notions of metatheoretic completeness, including those of "monomorphism" and "completeness of the models". How, then, does his account relate to Baldus limitative result? According to Hintikka (1991), the relation is one of complete misunderstanding of the original problem on Carnap's side: Due to a tacit "universality assumption" in Carnap's thought he presupposed in his formal models an invariant and fixed domain of discourse. Since the domain of a model cannot be extended or restricted in this conception, Carnap simply missed the point of extremal conditions (expressed by Hilbert and later by Baldus in terms of the *non-extensibility* of a model's domain) (ibid, 332). Hintikka observes further that since "the bone of contention was thus a profound difference in the conceptual assumption of the two parties, it is perhaps not surprising that is was never consciously acknowledged by the disputants, either by Baldus or by Carnap." (ibid, 333) Now, whereas Hintikka's last observation is correct, I take the reason for this to be that Carnap and Baldus *are* in fact (both in Carnap and Bachmann (1936) and in Carnap (2000)) addressing the same problem as Hilbert and Baldus. I showed in Chapter 2 that Hintikka's reading is not a valid characterization of Carnap's version of extremal conditions. Moreover, I argued there that

---

[245] A second example for a model maximal, non-categorical theory mentioned by Baldus is the theory of algebraically closed, absolute algebraic fields without a characteristic mentioned to him by Emmy Noether. See (ibid, 328-329).

once the heterodox character of Carnap's conception of models is understood properly, also his formal reconstruction of extremal axioms shows to be in accord with mathematical practice.

What remains to be studied here is the actual relation between Baldus' limitative result and Carnap's account. Here again, a closer look at Carnap's unpublished work on the topic proves to be insightful. There is archival evidence that Carnap was, already in 1928, seriously engaged in assessing the more general relation between the two types of completeness. For instance, in a letter to Fraenkel dated March 26, 1928 he writes:

> I have also started to analyse the question of the relation between the axiom of restriction, the axiom of completeness and monomorphism. The treatment of the axiom systems as propositional functions greatly facilitates the investigation. (RC 081-01-26)[246]

Unfortunately, it seems impossible to decide on the basis of Carnap's documents in the *Nachlass* whether or not he was aware of Baldus (1928) when working on the subject in his manuscript for Part 2 of *Untersuchungen*.[247] Nevertheless, a closer look at what Carnap thought about the relation of completeness of models and categoricity in his projected Part 2 is quite revealing. It shows that Carnap was in fact fully aware that the addition of an extremal axiom to a given base system was not a sufficient means to yield categoricity.

Before turning to Carnap's own results on this point, recall again his convention for discussing extremal constraints on the models of an axiom system. His last remark in the letter to Fraenkel that the conception of axiom systems as propositional functions simplifies the study of extremal axioms is central here. The main concept introduced by him to study extremal constraints is that of a model's structure.[248] In *Untersuchungen*, a model's structure characterizes a model in the sense that it fixes its isomorphism class (Carnap (2000, 74)). Accordingly, an axiom system - expressed as a propositional function f(*R*, *S*, *T*) with primitive relations *R*, *S*, *T* - has a "structure number" referring to the numbers of different structures it defines. Each of these structures fixes one isomorphism class of models. An axiom system

---

[246] "Ich habe auch angefangen, die Frage des Zusammenhanges zw. Berchr.Axiom, Vollst.Ax und Monomorphie zu untersuchen. Die Behandlung der AS als Aussagefunktionen erleichtert die U. sehr." (RC 081-01-26)
[247] In an attached note to the manuscript titled "For Axiomatics, to read" (RC 081-01-30) he mentions only Baldus (1927), a paper on the Archimedean axiom in geometry, but not Baldus (1928). Overall, I take it to be rather unlikely that Carnap was aware of Baldus (1928) when working on his manuscript.
[248] Compare 1.3.2 for a more detailed account of Carnap's 'structuralist' conception of axiom systems; see 2.2.1 for his notion of a 'model structure'.

with only one structure is categorical ("monomorphic"), a system with more than one structure non-categorical ("polymorphic") (ibid, 128).

We also saw in Section 2.2.2 that the "extremal conditions" imposed by minimal and maximal axioms were also cashed out in terms of structures and structure extensions (and restrictions respectively): In Part 2 (as well as in Carnap and Bachmann (1936)), a segment T_{SF} of this class of structures of an axiom system can be classified by the relation of a being a "proper substructure".[249] "*Beginning structures*" of f are those belonging to the domain of T_{SF}. "*End structures*" of f are those belonging to the range of T_{SF}. Finally, "*isolated structures*" are the structures of f not belonging to T_{SF} (see (Carnap and Bachmann 1981, 75-76), RC 081-01-08). Accordingly, the "*minimal structures*" of f imposed by a minimal axiom restrict the possible models of f to instances of the beginning- and isolated structures; "*maximal structures*" imposed by maximal axioms fix the models to be instances of either the end- or the isolated structures of f (see (ibid, 76); RC 081-01-08).[250]

Given this structural account of axiom systems, we can now turn to Carnap's understanding of the general logical relation of completeness of models and categoricity. In a document titled "Monomorphism by BA [axioms of restriction] or VA [axioms of completeness]," he discusses the general conditions which a base axiom system has to fulfill so that an extremal axiom implies the categoricity of the resulting system (RC 081-01-09). For the case of maximal axioms, the base system in question has to possess among the different possible structures either:

(i) a single undividable isolated structure and no undividable end structure *or*
(ii) a single undividable end structure and no isolated structure (RC 081-01-09).

[249] A structure S is called a "*proper substructure*" of an existing structure *T*, if *S* and *T* are not identical and there are at least two models *M* and *N*, where *M* is a "proper part" of *N* and *M* satisfies *S* and *N* satisfies *T*. See RC 081-01-06 and Carnap and Bachmann (1981, 75).

[250] Note that Carnap's reconstruction of extremal condition is cast in terms of model's structures and structure extensions and not, in terms of models and model extension. It is important to emphasize here again that this does not commit him to a nonstandard understanding of the informal use of models and model extensions found in Hilbert, Baldus, and others. As I show in Section 2.5, Carnap's notions of structure extensions and extremal structures fully correspond to their informal counterparts in Hilbert and others, provided that Carnap's heterodox account of models is understood properly.

Put in terms of models, a maximal axiom adds up to a categorical axiomatization if the base axiom system is satisfied by either exactly one maximal model or one structurally non-extensible model.[251, 252]

How does this account relate to the results in Baldus (1928)? It is clear in both cases that the addition of an extremal axiom was not conceived as a sufficient condition to yield the categoricity of a theory. For Carnap, a maximal axiom effects categoricity only if it restricts the set of admissible models to either an isolated or a single maximal model. This is essentially what Baldus expresses in his example of absolute geometry. Translated in Carnap's terms, the point to see there was that the axiom system allows different and independent proper substructures, say a "Euclidian" and a "hyperbolic" substructure. Baldus' axiom of continuity thus restricts the class of models to the respective maximal models, i.e. to the models of the respective maximal structure of each substructure. Since the condition (ii) of a *single* maximal structure is not met here, adding a maximal axiom to the base system does not imply categoricity. As we have seen, it only gets categorical if an additional axiom, say the axiom of parallels, is added to the theory. In this case, the resulting base axiom system (taken without a maximal axiom) possesses only one proper substructure, namely the "Euclidian" structure. Since Baldus' axiom of continuity restricts the possible models to the model(s) of the maximal structure, the theory suffices to capture its interpretations up to isomorphism.

Given this, Carnap's account in 1928 can be considered as a formal equivalent to Baldus result. In fact, it allows getting a better grip on the difference described informally by Baldus between the different inconsistent structures within absolute geometry. Note that the issue of extremality constraints and categoricity is taken up again in Carnap and Bachmann (1936), however this time with explicit reference to Baldus (1928). Carnap and Bachmann, in discussing Baldus' examples note here that whereas the inference from categoricity to completeness of models is trivial,

---

[251] An analogue case can be found a few pages later for the case of maximal structure axioms. Here, adding the an axiom yields categoricity if among the possible models there are either (i') a single isolated structure (dividable or undividable) and no end structure or (ii') a single end structure (dividable or undividable) and not isolated structure (RC 081-01-11).

[252] Note that this result is closely related to Carnap's second notion of completeness extensively discussed in Part 1 of *Untersuchungen*, namely "non-forkability" ("*Nichtgabelbarkeit*"). It states, in modern terms, that a theory T is semantically complete if there is no sentence φ expressible in T such that both $T \cup \{\varphi\}$ and $T \cup \{\neg\varphi\}$ are satisfiable (see Awodey and Reck (2002, 4)). Given this, one could say that for Carnap, a theory is categorical iff it is complete in the models and *non-forkable*.

The converse does not follow, however. The nonexistence of extensions for the models of an axiom system does not at all imply the monomorphism of an axiom system. (Carnap and Bachmann (1981, 82))

For Carnap in 1936, the inference in this direction might fail due to the existence of "incomparable models, i.e., such that neither can be isomorphically mapped into the other." (ibid, 82) Given the case of absolute geometry (as well as Noether's algebraic example), the authors go on to specify a general condition under what circumstances the addition of a maximal axiom implies categoricity:

An axiom system closed by a maximal axiom is monomorphic if among the models of the body axioms *there is one into which all others* (that do not belong to endlessly increasing structure sentences of the body axioms) *can be isomorphically mapped*. (ibid, 82, my emphasis)[253]

Two remarks on this passage are in order here. First, there is a slight change in the treatment of model maximality compared to Carnap's original account in *Untersuchungen*. Instead of speaking of end (and isolated) structures, the maximality of a model is expressed here in terms of its embedding properties.[254]

Second, and more important for our discussion, note how the stronger notion of completeness of models *for a categorical theory* is stated here. Model maximality in this sense can be paraphrased in this way: a model M is (up to isomorphism) the *unique model* of T that admits no proper extension to a model M' of T. Two characteristics of M are expressed here: (a) uniqueness and (b) non-extensibility. Now, the uniqueness condition had traditionally been understood as a direct consequence of the non-extensibility of a maximal model, for instance in Hilbert's axiom system for Euclidian geometry. However, as we have seen, this inference does not hold in absolute geometry.

A similar example can be given for the complementary case of minimal axioms. The relation between (a) and (b) is not completely settled in Carnap's original example of minimal axioms, Fraenkel's AR. Already in von Neumann (1925), the earliest critical discussion of Fraenkel's axiom candidate, doubts were raised concerning the model minimality effected by AR. More

---

[253] Compare also the following related passage: "For example, in monomorphic, closed axiom systems maximal axioms are equivalent with the positive requirement that every model of the axiom body can be isomorphically mapped into *M*." (Carnap and Bachmann 1981, 82)

[254] See Ehrlich (1995) for a closer discussion of this point.

specifically, von Neumann argued that Fraenkel's proposed method of devising a model via the intersection of all possible models need not necessarily lead to a single minimal model satisfying the other axioms and thus to a categorical axiomatization (ibid, 405).[255]

It is in this light, I think, that Carnap's and Baldus' more general limitative results on extremal axioms have to be seen: both authors stress the fact that whereas (b) is generally secured by an extremal axiom, (a) is not. Condition (b) can also hold for structurally different models of an axiom system. Now, in Carnap and Bachmann (1936), the conceptual distinction between (a) and (b) is drawn explicitly.[256] Nonetheless, we have seen that already in 1928 this distinction is clearly reflected in Carnap's account of extremal structures. In addition to the Carnap's remarks in (RC 081-01-09) discussed above, this becomes obvious from a distinction he draws throughout the manuscript between "deducing" an extremal axiom from a given axiom system and "applying" it to the axiom system (e.g. RC 081-01-05). Briefly, a maximal axiom can be deduced in the case that the system possesses *one or more* maximal (or isolated) structures. In these cases, the completeness of models is not necessarily an exclusive property of a single model. Thus, according to Carnap, it makes no sense to "apply" a maximal axiom in order to effect categoricity. In contrast, a maximal axiom should only be applied to a system if it helps to fix one unique maximal model as the intended model of the theory (as is the case in Euclidian geometry) (see ibid.).

3.6 Conclusion

As pointed out by Hintikka, there was a "subsequent deterioration of the discussion" on the general character of extremal axioms after Baldus (1928) and Carnap and Bachmann (1936) (Hintikka 1991, 334). This is an effect, however, of Carnap's misguided semantic understanding of extremal condition. To the contrary, in this chapter I aimed to show that Carnap - in his *Untersuchungen* from 1928 and later on - not only presented a formal account of extremal axioms that closely captures the informal conceptions of model maximality and minimality expressed in his mathematical sources, most importantly Hilbert's AC and

---

[255] According to von Neumann, this is due to the fact that the range of the generalization over all (sub-) models of a theory involved in the intersection approach depends on which background set theory is assumed. Different (or better differently interpreted) background systems might allow different ranges of submodels. Given Fraenkel's paring down method of specifying a minimal model, this may lead to different results when different systems are assumed as the background theory (ibid, 405).

[256] Compare again how model maximality is introduced in 1936: "*The models of an axiom system* that is closed by a maximal axiom possess a certain completeness property in that *they* cannot be extended without violating the original axiom system." (ibid, 82, my emphasis)

Fraenkel's AR. Moreover, he (together with his later collaborator Bachmann) also established a more refined picture of the connection between the completeness of models and categoricity that was in full agreement with Baldus result. Finally, whereas Carnap's type-theoretic account of extremal axioms has to considered as clearly heterodox and problematic from a modern logical perspective, it would be a mistake to make the infer ex-post that is was not conceived in earlier days as a valid formal explication. Two references should suffice to make this point. In a later paper on Hilbert's AC (Bernays (1955)), Bernays after mentioning the problematic logical character of the axiom states that: "That this complication does not, on the other hand, hinder the logistic formalization of the axiom of completeness was shown by R. Carnap and F. Bachmann (…) [in Carnap and Bachmann (1936)]." (ibid, 219)[257]

Second, in a Fraenkel's *Foundations of Set Theory* (Fraenkel & Bar Hillel 1958), Fraenkel referred to Carnap's attempts to formalize AR as a minimal axiom in HOL, concluding that "True, recently (…) Carnap proposed a vindication of this axiom of restriction, and Carnap formulated it symbolically, as an axiom of a minimal model (…)." (ibid, 90)[258] In an attached footnote to the passage cited above he stressed this notion of a restriction condition in Carnap's formal presentation as an adequate version of his own informal treatment of AR: "The pith of the axiom is then the demand that no "partial relation" ε should fulfill the conditions expressed by the other axioms." (Fraenkel & Bar Hillel 1958, 90)

These two references by Bernays and Fraenkel from the 1950s clearly do show that Carnap's formal account of extremal axioms was, at least before the eventual turn to contemporary model-theoretic semantics, generally considered as a serious contribution to the (meta-)theory of formal axiomatics.

---

[257] "Dass andererseits diese Komplikation die logistische Formalisierung des Vollständigkeitsaxioms nicht hindert, wurde von R. Carnap and F. Bachmann gezeigt (…) [in Carnap and Bachmann (1936)]." (ibid, 219)

[258] The formalization referred to here is '$(H)[(x)(y)(Hxy \supset Exy).\ \mathrm{Kon}(H) \supset (x)(y)\ (Hxy \leftrightarrow Exy)]$' from (Carnap 1954). As already in *Untersuchungen* an Carnap and Bachmann (1936), Fraenkel's account of a minimal property codified in AR is symbolized here by the use of higher-level binary relations (representing different possible membership relations) and the notion of a minimal partial relation here: "There exists no proper partial relation of E that also satisfies the properties stated in axioms A1 to A8 [GS: i.e. ZF]." (Carnap 1954, 154) Compare also Schiemer (forthcoming).

Outlook

Carnap's project of *Untersuchungen* was eventually abandoned as a direct reaction to Gödel's incompleteness theorems and to Tarski's metalogical work (see Awodey and Carus (2001) and Goldfarb (2005)). With the subsequent turn to a purely syntactical investigation of mathematics in *Logical Syntax of Language* also the explicit semantic framework of his "general axiomatics" was given up. This holds in particular for his notion of a formal model. The concept is not mentioned in Carnap (1934). More surprisingly, also his later work on semantics, most notably the three-volume book project *Series in Semantics*, remains silent on the notion. Instead of discussing formal models, Carnap introduces here the substitute concepts of "state of affairs," and "state descriptions" (see Carnap (1942) and (1947)).
It is only in his very late work, in his 'Replies and Expositions' in the Schilpp-volume (Schilpp (1963)) that the notion is eventually reintroduced in print. In §10 of the book - titled "My conception of semantics" - the following definition of 'model' is given:

> A model for a language (in the extensional sense of "model" customary in mathematics, as in the definitions by Tarski, Kemeny, and others) is an assignment of extensions of the following kind: To every type of variables a class of entities of this type is assigned as the range of values, and to every primitive constant of the type system an extension of the same type is assigned. (ibid, 902)

The conception outlined here essentially conforms to the way formal models are understood today. Interestingly, following a remark on the 'structure' of a model, Carnap also refers to his early work on axiomatics here. In an attached footnote he adds: "For more exact definitions, especially with respect to axiom systems, see Carnap and Bachmann [1936]."
This reference gives the impression of a certain continuity between his original theory of models for axiomatic theories and the late standardized account mentioned in 1963. However, in light of the results of the present work, this shows to be more of a retrospective idealization than an adequate account of the evolution of this notion in his work. One of the main objectives in the dissertation was to point out where exactly his early theory is heterodox compared to modern model theory. In particular, it was to show that the conception of models underlying *Untersuchungen* and also Carnap and Bachmann (1936) differs significantly from the above specification of "model(s) for a language".
The results gained in the present work open up possibilities for further research in Carnap scholarship. In particular, it would be worthwhile to retrace Carnap's evolving views on

semantics in general and on models in particular in his subsequent work up to Schilpp (1963). As I see it, at least three pressing questions are still open for closer discussion: (1) Which influences led to the shift from Carnap's substitutional conception of models to a modern extensional understanding in terms of an "assignment of extensions" outlined above? (2) At what point did he change his 'universalist' view of a pure type-theoretic language in favor of the use of schematically understood languages with non-logical terminology? (3) Finally, when did Carnap adopt the modern notion of domain variation for models. Addressing these questions would call for a closer study to the few documents (mostly in informal correspondence and in discussion notes from the *Nachlass*) where the notion of models resurfaces in Carnap's later philosophy. This includes (among other things) his notes of the Harvard discussions with Tarski and Quine from 1940/41 as well as his correspondence with Kemeny and Bar-Hillel on issues in inductive logic in the early 1950s. Investigation of these discussions will allow a deeper understanding of the conceptual transitions in Carnap's thinking about formal semantics throughout his intellectual career. This, however, is work for another day.

Acknowledgements

An early version of the first chapter of the present work was presented at the Wissenschaftsphilosophisches Kolloquium (VCI) as well as at the Doktorandensymposium 2010/ÖGP in Salzburg. Earlier drafts of the second chapter were presented as talks for the SoCal HPLM Group at UC Irvine, at the Logisches Café at the University of Vienna, at the GAP.7 in Bremen as well as the Epsa2009 in Amsterdam. I thank the members of the respective audiences for useful comments and suggestions. I am especially indebted to my supervisor Richard Heinrich as well as to Michael Friedman, Erich Reck, Ilkka Niiniluoto, Friedrich Stadler, Martin Kusch, Arianna Betti, and Günther Eder for valuable feedback.

References

Awodey, S. and Carus, A. (2001). Carnap, completeness, and categoricity: the Gabelbarkeitssatz of 1928, *Erkenntnis* 54: pp.145-172.

Awodey, S. and Klein, C. (eds.) (2004). *Carnap brought home: The view from Jena*, Open Court

Awodey, S. and Reck, E. (2002). Completeness and categoricity. Part 1: Nineteenth-century axiomatics to twentieth-century metalogic, *History and Philosophy of Logic* 23: pp.1-30.

Awodey, S., Reck, E., and Gabriel, G. (eds.). (2004). *Frege's Lectures on Logic: Carnap's Jena Notes, 1910-1914*, Open Court

Awodey, S. (2007), Carnap’s quest for analyticity: the *Studies in Semantics*. In: M. Friedman and R. Creath (ed.) (2007): pp. 226-247.

Awodey, S. and Carus, A. (2007), Carnap’s dream: Gödel, Wittgenstein and the *Logical Syntax*, *Synthese*.

Bachmann, F. (1934). *Untersuchungen zur Grundlegung der Arithmetik mit besonderer Beziehung auf Dedekind*, Frege und Russell. Münster.

Bachmann, F. (1934-35). Einige Modifikationen zum SCHOLZschen Begriff der exakten Wissenschaften, *Semesterberichte zur Pflege des Zusammenhaltes von Universität und Schule aus den mathematischen Seminaren* (special print), vol. 6.

Bachmann, F. (1936). Die Fragen der Abhängigkeit und der Entbehrlichkeit von Axiomen in Axiomensystemen, in denen ein Extremalaxiom auftritt, Actes du Congrès International de Philosophie Scientifique 7 (1935), Paris: pp.39-53.

Badesa, C. (2004). *The Birth of Model Theory. Löwenheim’s Theorem in the Frame of the Theory of Relatives*, Princeton: Princeton University Press.

Baldus, R. (1928). Zur Axiomatik der Geometrie 1: über Hilbert’s Vollständigkeitsaxiom, *Mathematische Annalen* 100: pp.321-333.

Baldus, R. (1938). Über nicht beweisbare und doch entbehrliche Axiome, *Mathematische Annalen* 44: pp.321-329.

Bays, T. (2001). On Tarski on models, Journal of Symbolic Logic 66: pp. 1701-1726.

Behmann, H. (1927). *Mathematik und Logik*, Leipzig and Berlin.

Bernays, Paul. (1922). Die Bedeutung Hilberts für die Philosophie der Mathematik, *Die Naturwissenschaften* 10: 93-99. Engl. transl. in: Mancosu, P. (1998): 189-197.

Bernays, P. (1955). Betrachtungen über das Vollständigkeitsaxiom und verwandte Axiome, *Mathematische Zeitschrift* 63: pp.219-229.

Blanchette, P. (2009). The Hilbert-Frege controversy, *The Stanford Encyclopedia of Philosophy*, E. Zalta (ed.), URL = http://plato.stanford.edu/entries/frege-hilbert/

Bohnert, H.G. (1975). Carnap’s logicism. In: J. Hintikka (ed.). *Rudolf Carnap, Logical Empiricist*. Dortrecht, pp.183-216.

Bonk, T. and Mosterin, J. (2000). Editors’ Introduction, in: Carnap, R. (2000): pp.1-54.

Burge, T. (2005), *Truth*, *Thought*, *Reason: Essays on Frege*, New York: OUP.

Burgess, J. (2005). *Fixing Frege*, Princeton: PUP.

Carnap, R. (1927). Eigentliche und uneigentliche Begriffe, *Symposion* 1: pp.355-374.

Carnap, R. (1929). *Abriß der Logistik*. Wien.

Carnap, R. (1930). Bericht über Untersuchungen zur allgemeinen Axiomatik, *Erkenntnis* 1: pp.303-307.

Carnap, R. (1930a). Die Mathematik als Zweig der Logik, *Blätter für deutsche Philosophie*, 4: pp.298-310.

Carnap, R. (1931). Die logizistische Grundlegung der Mathematik, *Erkenntnis* 2: pp.91-105, translated in P. Benacerraf and H. Putnam (eds.) (1983). *Philosophy of Mathematics - Selected Readings* (2nd ed.). Cambridge: CUP: pp. 41-52.

Carnap, R. (1934/2002). *Logische Syntax der Sprache*, Vienna: Springer, translated as: *The Logical Syntax of Language*, London: Open Court.

Carnap, R. (1935). Ein Gültigkeitskriterium für die Sätze der klassischen Mathematik, *Monatshefte für Mathematik und Physik*, 42, pp. 163-190.

Carnap, R. and Bachmann, F. (1936/1981). Über Extremalaxiome, *Erkenntnis* 6: pp.166-188, translated as: On extremal axioms, Bohnert, H. (ed.), *History and Philosophy of Logic* 2: pp. 67-85

Carnap, R. (1939). Foundations of Logic and Mathematics, *International Encyclopedia*, 1/3. Chicago: Chicago Press.

Carnap, R. (1942). *Introduction of Semantics*. Cambridge MA: Harvard University Press.

Carnap, R. (1947). *Meaning and Necessity*. Chicago: Chicago Press.

Carnap, R. (1954), *Einführung in die Symbolische Logik*, Vienna: Springer.

Carnap, R. (2000). *Untersuchungen zur allgemeinen Axiomatik*. Bonk, T. and Mosterin, J. (eds.). Darmstadt.

Carnap, R. (2003). *The Logical Structure of the World*, London: Open Court.

Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap*, Cambridge: CUP.

Corcoran, J. (1980). Categoricity, *History and Philosophy of Logic*, 1: pp.187-207.

Dedekind, R. (1888). *Was sind und was sollen die Zahlen*, in R. Dedekind (1932), *Gesammelte mathematische Werke*, vol 3, Braunschweig: Vieweg& Sohn: pp.335-391.

de Jong, W. and Betti, A. (2010). The classical model of science: a millennia-old models of scientific rationality, *Synthese* 174/2.

Demopoulos, W. and Clark, P. (2005). The logicism of Frege, Dedekind, and Russell, in: Shapiro, S. (ed.) (2005). *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford: OUP: pp. 129-165.

Ehrlich, P. (1995). Hahn's Über die Nichtarchimedischen Grössensysteme and the development of the modern theory of magnitudes and numbers to measure them, in: J. Hintikka (1995).

Ehrlich, P. (1997). From completeness to Archimedean completeness, *Synthese* 110, pp.57-76.

Enderton, Herbert B. (2001). *A mathematical Introduction to Logic*. San Diego: Hartcourt/Academic Press.

Ferreirós, J. (2007). *Labyrinth of Thought: a History of Set Theory and its Role in Modern Mathematics* (2nd ed.), Basel: Birkhäuser.

Ferreirós, J. & Gray, J. (eds.). (2006). *The Architecture of modern Mathematics: Essays in History and Philosophy*. Oxford: Oxford University Press.

Fraenkel, A. (1919). *Einleitung in die Mengenlehre*, Berlin: Springer.

Fraenkel, A. (1922a). Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre, *Mathematische Annalen*, 86: pp.230-237.

Fraenkel, A. (1922b). Axiomatische Begründung der transfiniten Kardinalzahlen, *Mathematische Zeitschrift*, 13: pp.153-188.

Fraenkel, A. (1924). *Einleitung in die Mengenlehre*, (2nd ed.), Berlin: Springer.

Fraenkel, A. (1927). *Zehn Vorlesungen über die Grundlegung der Mengenlehre*, Leipzig/Berlin: Teubner.

Fraenkel, A. (1928). *Einleitung in die Mengenlehre*, (3rd ed.), Berlin: Springer.

Fraenkel, A. and Bar-Hillel, Y. (1958). *Foundations of Set Theory*, Amsterdam: North Holland Publishing Company.

Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle a. S.

Frege, G. (1980). *Philosophical and Mathematical Correspondence*, Gabriel, G and MacGuinness, B. (eds.), Oxford: Blackwell.

Friedman, M. (1999). *Reconsidering logical positivism*. Cambridge: Cambridge University Press.

Friedman, M. and Creath, R. (eds.) (2007). *The Cambridge Companion to Carnap*, Cambridge: CUP.

Goldfarb, W. (1979). Logic in the twenties: the nature of the quantifier, *Journal of Symbolic Logic* 44: pp.351-368.

Goldfarb, W. (2005). On Gödel's way in: the influence of Rudolf Carnap, *Bulletin of Symbolic Logic* 11/2, pp.185-193.

Goméz-Torrente, M. (2002). The problem of logical constants, *The Bulletin of Symbolic Logic*, 8/1: pp.1-37.

Goméz-Torrente, M. (2006). Tarski on logical consequence. *Notre Dame Journal of Formal Logic*, 37/1: pp. 125-151

Goméz-Torrente, M. (2009). Rereading Tarski on logical consequence, *Review of Symbolic Logic*, 2/2: pp. 249-297

Grattan-Guinness, I. (2000). *The Search For Mathematical Roots 1870-1940. Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell to Gödel*. Princeton: Princeton University Press.

Hahn, H. and Carnap, R. and Gödel, K. and Heyting, A. (1931). Diskussion zur Grundlegung der Mathematik, *Erkenntnis* 2: pp.135-151.

Hallett, M. (2008). Reflections on the purity of method in Hilbert's *Grundlagen der Geometrie*, in P. Mancosu (ed.) (2008): pp. 198-256.

Heck, R. (2007). Frege and semantics, *Grazer Philosophische Studien*, 75: pp.27-63.

Hilbert, D. (1899). *Die Grundlagen der Geometrie*. In: Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen. Leipzig: Teubner: pp.1-92.

Hilbert, D. (1910). *The foundations of geometry* (2nd ed.). Chicago.

Hilbert, D. & Bernays, P. (1934). *Grundlagen der Mathematik* (vol. 1). Berlin: Springer.

Hintikka, J. (1988). On the development of the model-theoretic viewpoint in logical theory, *Synthese* 77:1-36. Reprinted in: Hintikka (1996): pp.104-139.

Hintikka, J. (1991). Carnap, the universality of language and extremality axioms. *Erkenntnis* 35: pp.325-336.

Hintikka, J. (1992). Carnap's work in the foundations of logic and mathematics in a historical perspective, *Synthese* 93: pp.167-189.

Hintikka, J. (1995). Standard vs. nonstandard distinction: a watershed in the foundations of mathematics, in Hintikka, J. (1995a): pp.21-44.

Hintikka, J. (ed.) (1995a). *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, Dordrecht

Hintikka, J. (1997). *Lingua Universalis vs. Calculus Ratioocinator: An Ultimate Presupposition of Twentieth-Century Philosophy*, Selected Papers II. Dortrecht: Kluwer Academic Publishers.

Hintikka, J. (forthcoming). What is the axiomatic method?, *Synthese*.

Hodges, W. (1997). *A Shorter Model Theory*. Cambridge: CUP.

Jané, I. (2006). What is Tarski's common concept of consequence, *The Bulletin of Symbolic Logic*, 12: pp.1-42

Kanamori, A. (2004). Zermelo and Set Theory, *The Bulletin of Symbolic Logic*, 10/4, pp.487-553.

Kanamori, A. (2007). Gödel and Set Theory, *The Bulletin of Symbolic Logic*, 13/2, pp.153-188.

Koellner, P. (2010). Strong logics of first and second order, *The Bulletin of Symbolic Logic*, 16/1, pp.1-36.

Kreisel, G. (1972). Informal rigour and completeness proofs, in: I. Lakatos (ed.) (1972). *Problems in the Philosophy of Mathematics*. Amsterdam: North-Holland: pp.138-171

Lavers, G. (2008). Carnap, formalism, and informal rigour, *Philosophia Mathematica* 3/16: pp.4-24.

Linnebo, O. (2004). Frege's proof of referentiality, *Notre Dame Journal of Formal Logic*, 45/2: pp.73-98.

Mancosu, P. (2006). Tarski on models and logical consequence. in: Ferreirós, J. and Gray, J. (eds.) (2006): pp.209-237.

Mancosu, P. (ed). (1998). *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s*. Oxford: OUP.

Mancosu, P. (ed). (2008). *The Philosophy of Mathematical Practice*, Oxford: OUP.

Mormann, T. (2007). Geometrical leitmotifs in Carnap's early philosophy, in: M. Friedman and R. Creath (eds.) (2007): pp. 43-64.

Niiniluoto, N. (2003). Carnap on Truth, in T. Bonk (ed.), *Language, Truth and Knowledge* (Vienna Circle Institute Library), Dortrecht: Kluwer Academic Publishers.

Parsons, C. (1990), The structuralist view of mathematical objects, *Synthese,* 84/3.

Reck, E. and Price, M. (2000). Structures and structuralism in contemporary philosophy of mathematics, *Synthese*, 125/3: pp.341-383.

Reck, E. (2004). From Frege and Russell to Carnap: Logic and Logicism in the 1920s, in: S. Awodey and C. Klein (2004): pp.151-180.

Reck, E. (2007). Carnap and modern logic, in: Friedman, M. and Creath, R. (2007): pp. 176-199.

Russell, B. (1993). *Introduction to mathematical philosophy*. (2nd ed.). New York: Dover Publications.

Schiemer, G. (2010), Fraenkel's axiom of restriction: axiom choice, intended models and categoricity, in: B. Löwe and T. Müller (eds.), *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice*, London: College Publications.

Schilpp, P.A. (ed.) (1963). *The Philosophy of Rudolf Carnap*, London.

Shapiro, S. (1991). *Foundations without Foundationalism - A Case for Second-Order Logic*, Oxford: OUP.

Shapiro, S. (2000). *Thinking about Mathematics – The Philosophy of Mathematics*, Oxford: OUP.

Shapiro, S. (2005). Categories, structures, and the Frege-Hilbert controversy: the status of meta-mathematics, *Philosophia Mathematica*,13/1: pp.61-77.

Sieg, W. and Schlimm, D. (2005). Dedekind's analysis of number: systems and axioms, *Synthese*, 147, pp.121-170.

Smith, P. (2008). Ancestral arithmetic and Isaacsons's thesis, Analysis, 68: pp.1-10.

Stadler, F. (2001). *The Vienna Circle – Studies in the Origins, Development, and Influence of Logical Empiricism*. Wien/NewYork: Springer.

Stein, H. (1988). *Logos*, Logic, and *Logistiké*: some philosophical remarks on the 19th-century transformation of mathematics, in: Aspray, W. and Kitcher, P. (eds.) (1988). *History and Philosophy of Modern Mathematics*. Minnesoty: University of Minnesota Press: pp.238-259.

Tappenden, J. (1997?). Metatheory and mathematical practice in Frege, URL = http://www-personal.umich.edu/~tappen/METATHEORY_REVISED_PROOFS.pdf

Tarski, A. and Lindenbaum, A. (1934-35). On the limitations of the means of expression of deductive theories, reprinted in: A. Tarski. (1956): pp. 384-392.

Tarski, A. (1936). The concept of logical consequence, reprinted in: A. Tarski. (1956): pp. 409-420.

Tarski, A. (1956). *Logic, Semantics, Metamathematics*. Corcoran, J. (ed.). Indianapolis: Hackett.

van Heijenoort, J. (1967). Logic as calculus and logic as language, *Boston Studies in the Philosophy of Science* 3: 440-446. reprinted in: Hintikka, J. (1997): 233-239.

van Heijenoort, J. (ed.). (1967a). *From Frege to Gödel: A source book in mathematical logic, 1879-1931*. Cambridge: HUP.

von Neumann, J. (1925). An axiomatization of set theory, reprinted in: van Heijenoort (1967): pp.393-413.

Wagner, P. (ed.). (2009). *Carnap's Logical syntax of language*, Palgrave Macmillan

Zermelo, E. (1908). Investigations in the foundations of set theory I, repr. in: van Heijenoort (1967), pp.200-215.

Webb, J. (1995). Tracking contradictions in geometry*:* The idea of model from Kant to Hilbert, in: Hintikka, J. (1995): pp.1-20.

Wehmaier, K. (1997). Aspekte der Frege-Hilbert-Korrespondenz, *History and Philosophy of Logic*, 18: pp.201-209.

Whitehead, N. A. and Russell, B. (1962). *Principia Mathematica* (3 volumes). Cambridge.

Wolenski, J. (ed.) (1999). *Alfred Tarski and the Vienna Circle*, Dordrecht: Kluwer Academic Publishers

Archival material used:

The Rudolf Carnap Collection, Philosophisches Archiv/Universität Konstanz

Abstract (German and English)

In jüngerer Zeit hat sich ein verstärktes Interesse an den historischen und technischen Details von Carnaps Philosophie der Logik und Mathematik entwickelt. Meine Dissertation knüpft an diese Entwicklung an und untersucht dessen frühe und formative Beiträge aus den späten 1920er Jahren zu einer Theorie der formalen Semantik. Carnaps zu Lebzeiten unveröffentlichtes Manuskript *Untersuchungen zur allgemeinen Axiomatik* (Carnap 2000) beinhaltet ein Reihe von erstmals formal entwickelten Definitionen der Begriffe ‚Modell', ‚Modellerweiterung', und ‚logischer Folgerung'. Die vorliegende Dissertation entwickelt eine logische und philosophische Analyse dieser semantischen Begriffsbildungen. Darüber hinaus wird Carnaps frühe Semantik in ihrem historisch-intellektuellen Entwicklungskontext diskutiert. Der Fokus der Arbeit liegt in der Thematisierung einiger interpretatorischer Fragen zu dessen implizit gehaltenen Annahmen bezüglich der Variabilität des Diskursuniversums von Modellen sowie zur Interpretation seiner typen-theoretischen logischen Sprache. Mit Bezug auf eine Reihe von historischen Dokumenten aus Carnaps Nachlass, insbesondere zu dem geplanten zweiten Teil der *Untersuchungen* wird erstens gezeigt, dass dessen Verständnis von Modellen in wesentlichen Punkten heterodox gegenüber dem modernen Begriffsverständnis ist. Zweitens, dass Carnap von einer ‚nonstandard' Interpretation der logischen Hintergrundtheorie für seine Axiomatik ausgeht. Die Konsequenzen dieser semantischen Annahmen für dessen Konzeptualisierung von metatheoretischen Begriffen werden näher diskutiert. Das erste Kapitel entwickelt eine kritische Analyse von Carnaps Versuch, die axiomatische Definition von Klassen von mathematischen Strukturen mittels des Begriffs von ‚Explizitbegriffen' formal zu rekonstruieren. Im zweiten Kapitel werden die Implikationen von Carnaps frühem Modellbegriff für seine Theorie von Extremalaxiomen näher beleuchtet. Das letzte Kapitel bildet eine Diskussion der konkreten historischen Einflüsse, insbesondere durch den Mengentheoretiker Abraham Fraenkel, auf Carnaps formale Theorie von Minimalaxiomen.

In recent years one was able to witness an intensified interest in the technical and historical details of Carnap's philosophy of logic and mathematics. In my thesis I will take up this line and focus on his early, formative contributions to a theory of semantics around 1928. Carnap's unpublished manuscript *Untersuchungen zur allgemeinen Axiomatik* (Carnap 2000) includes some of the first formal definitions of the genuinely semantic concepts of a model, model extensions, and logical consequence. In the dissertation, I provide a detailed conceptual

analysis of their technical details and contextualize Carnap's results in their historic and intellectual environment. Certain interpretative issues related to his tacit assumptions concerning the domain of a model and the semantics of type theory will be addressed. By referring to unpublished material from Carnap's *Nachlass* I will present archival evidence as well as more systematic arguments to the view that Carnap holds a heterodox conception of models and a nonstandard semantics for his type-theoretic logic.

Given these semantic background assumptions, their impact on Carnap's conceptualization of certain aspects of the metatheory of axiomatic theories will be evaluated. The first chapter critically discusses Carnap's attempt to explicate one of the crucial semantic innovations of formal axiomatics, i.e. the definition of classes of structures, via his notion of "*Explizitbegriff*''. The second chapter analyses the impact of Carnap's early theory of model for his theory of extremal axioms. The final chapter reviews the mathematical influences, most importantly by the set theoretician Abraham Fraenkel on Carnap's specific formalization of minimal axioms.

Curriculum Vitae

Georg Schiemer

University of Vienna georg.schiemer@univie.ac.at

Department of Philosophy phone: +43/699/18254069

Universitätsstrasse 7

1010 Vienna, Austria

Positions

2009 (winter/spring quarter): Visiting Scholar at the Department of Philosophy/FSI, Stanford University

01/2008 – 02/2010: Research Associate at the Department of Philosophy, University of Vienna

01/2008-12/2009: DOC-Fellow of the Austrian Academy of Sciences (ÖAW)

Education

05/2007 – 07/2010: Doctoral program at the Department of Philosophy, University of Vienna (Dissertation Project: "Carnap's early semantics")

05/2007 M.A. (with distinction), University of Vienna

(MA thesis: "Moderne Axiomatik, Krise der Anschauung und Quantifizierung," 219 p.)

03/2005 M.A., Vienna University of Economics and Business

(MA thesis: "Konkurrenzpositionierung im Wiener Museumsmarkt", 232 p.)

2004 (summer term) Studies in Philosophy, Université Sophia Antipolis, Nice

10/1997-05/2007 Studies of Philosophy, University of Vienna

03/1998-03/2005 Studies of Economics and International Commerce, Vienna University of Economics and Business

06/1997 Graduation, BG XIX Gymnasiumstrasse (with highest distinction)

1994 – 1995 Bush School, Seattle WA

Grants

2009 Visiting Scholarship, Stanford University/University of Vienna – Advanced Graduate Student Exchange Program

2008-2009 Recipient of a DOC-Fellowship of the Austrian Academy of Sciences (ÖAW)

2008-2010 Three Dissemination – grants for international conference participations.
2004 (winter term) Erasmus mobility grant (Université Sophia Antipolis, Nice)

Teaching

“Wahrheit“, Seminar (with Richard Heinrich), Department of Philosophy, University of Vienna, Winter Term 2009/2010
“Formale Axiomatik und Anschauung”, Seminar, Department of Philosophy, University of Vienna, Summer Term 2010

Publications

Georg Schiemer. (2010). Fraenkel’s Axiom of Restriction: Axiom choice, Intended Models, and Categoricity, in B. Löwe, T. Müller (eds.). *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice*, College Publications, London, Texts in Philosophy, (peer reviewed)

Languages

German (native), English, reading knowledge in French

Personal Information

Date of Birth: 17/02/1979
Children: one son (Leo)
Private Address: Grünentorgasse 35/4,
1090 Vienna, Austria