Carnap's Formal Philosophy of Science

Hans Halvorson

August 1, 2022

While it is questionable whether philosophy ever makes progress, the way that philosophy has been practiced has changed significantly over time. These changes in philosophical method are especially interesting when they are directly linked with changes in the methodology of the sciences. For example, mathematical methods have developed rapidly over the past two hundred years, resulting in a new paradigm for philosophical rigor and genuinely new methods of philosophical inquiry. The possibility of philosophical methodology benefitting from developments in mathematics was clearly recognized by Bertrand Russell (1917) — who, nonetheless, did not himself carry out any systematic program of formal philosophy of science. The task was left, rather, to Rudolf Carnap, who made numerous proposals for the use of exact (i.e. mathematical) methods to increase our understanding of science. In this article, I give a sampling of the ideas and proposals of Carnap's formal philosophy of science. It is hoped that this article will be a resource not only to historians interested in Carnap, but also for contemporary philosophers who wish to build further on Carnap's ideas.

The primary formal tool for Carnap is the "new logic", i.e. the symbolic logic that had been developed by, among others, Boole, Frege, Russell, and Hilbert. However, we should avoid the error thinking that Carnap saw symbolic logic as something distinct from, or standing above, the rest of pure mathematics. (For example, advocates of the "semantic view of theories" frequently claim that the received view of theories — which they identify with Carnap's view — restricts itself to formal logic.) In fact, as Carnap's earliest work already shows, he was well-versed with a broad range of pure mathematical methods, including set theory, topology, differential geometry, abstract algebra, etc. For Carnap, symbolic logic is just one of the mathematical tools of which philosophers can avail themselves. What's more, when it comes to symbolic logic itself, Carnap never restricts himself to classical first-order logic. He is quite happy to work with non-standard logics, higher-order logics, type theories, etc. The overall spirit of his method is: if X is a tool that scientists themselves use, then let's see if philosophers can use X to understand science. Of course, Carnap, like Frege and Russell before him, was more interested in mathematical than empirical methodologies.

In the remainder of the article, I will proceed chronologically through some of Carnap's major proposals for formal philosophy of science. The first idea, proposed in the *Der logische Aufbau der Welt*, is the notion of a constitution system. The philosophical motivation for this idea comes from Russell's external world program, and specifically the notion of a "logical construction". The formal approach of the *Aufbau* also has a Russellian motivation, but in this case, from Russell's ramified theory of types, where there is a set of rules for generating types at level $n+1$ from types at level n. In more contemporary terms, we would say that Carnap is developing a meta-theory for how theories can be extended by means of definitions or other type constructions. As Carnap mentions in Chapter A, section 2, he is motivated by the fact that mathematicians have successfully constructed higher number systems (e.g. rational, real, and complex numbers) from the natural numbers, thereby showing that sentences about these higher number systems can be translated into sentences about natural numbers.

Why undertake this kind of construction? Early anglo-american interpreters of Carnap, such as Quine, largely took him to be driven by an empiricist agenda (e.g. reduce scientific theories to empirically verifiable statements) or by an ockhamist agenda (e.g. minimize ontological commitments). A case can certainly be made that Carnap had such motives, but we should also recall the historical context, where there were pressing practical questions about the relationship between different scientific theories and different sub-fields of science. For Carnap, a construction system goes some way to answering these practical questions. A successful construction of a from b and c shows that sentences about a can be translated into sentences about b and c, thus allowing for unambiguous communication across scientific theories or fields. Carnap was to return to the notion of "translation" repeatedly throughout his career, although his terminology for the notion was not consistent.

Carnap himself strictly distinguished his early work on the *Aufbau* from his later *Wissenshaftslogik* program. However, in highsight there are clear signs of continuity between the programs. The latter program is announced in Carnap's *Logische Syntax der Sprache*, where Carnap says that all subsequent philosophy (of science) should be the logical analysis of the language of science. Here Carnap envisions the philosopher of science standing in relation to the practice of science as the mathematical physicist stands in relation to the physical world. That is, the philosopher of science builds formal models (a rational reconstruction) of that part of scientific practice that can be illuminated by means of formal modelling. For Carnap, the creation of symbolic logic makes it possible to formally model the language of science — where "language" here is mean to include both grammar and logic. The goal of *Syntax* is not to carry out this program in detail with respect to specific sciences, but to provide a template for this kind of theorizing, and to explain the philosophical upshot of looking at science from this perspective.

In contemporary terminology, we would say that *Syntax* is an exercise in logical metatheory, where Carnap's guiding light is Hilbert's metamathematics. As a result, much of *Syntax* is devoted to the seemingly mundane exercise of classifying elements of the syntax of a scientific theory. Behind this apparently mundane exercise, however, lies Carnap's creative theoretical choices with regard to the classificatory concepts of the metatheory itself – a branch of logic that was still in its early infancy in the 1920s. The job of the philosopher of science, claims Carnap, is to propose a suitable framework for for thinking *about* science. What's more, the philosopher of science can dare to think — as Newton, Maxwell, and Einstein thought — that the relevant phenomena admit a precise, mathematical description.

The meta-theoretical vocabulary that Carnap develops in *Syntax* includes terms that are standard from meta-logic: name, variable, term, relation symbol, proof, etc.. However, Carnap also suggests that the philosopher of science has need for other, less precisely defined, meta-theoretical concepts, such as that of a "descriptive term". Here Carnap is trying to cash out an intuition that not every term in a scientific theory is intended to denote an element of physical reality; some terms, in contrast, play an auxiliary or organizing role. In Carnap's mind, the terms of pure mathematics — e.g. names for sets, numbers, etc. — are not intended to be descriptive, and so they should be classified as "logical terms". Nonetheless, he declares that this distinction is fluid, as for example, the metric tensor of a spacetime theory might either be taken to be descriptive or as logical, depending on whether it functions as background framework (as in special relativity) or as a dynamical variable (as in general relativity).

Another controversial choice that Carnap makes in *Syntax* is in equipping a language with its own consequence relation, which he says is generated by L-rules (purely logical rules) and P-rules (including perhaps contingent physical laws). He mentions that even L-rules are up for grabs, but that it is convenient to consider a collection of languages that share the same L-rules. Of course, it is precisely this specification of L-rules that was subject of attack in Quine's *Two Dogmas of Empiricism*.

Carnap takes up the topic of translation and interpretation in Part IV(d) of *Syntax*. Here he first defines the notion of a *syntactic correlation* between two languages, which is a map from syntactic elements of the first language to syntactic elements of the second language. (Quine was later to use the term "reconstrual" for a similar notion.) Carnap then goes on to define the notion of a *transformance* between two languages, which is a syntactic correlation that transforms the consequence relation of the first language to the consequence relation of the second language. Given that Carnap's languages include their own consequence relations (which might include P-rules), logicians today would call his languages "theories" and his transformances "translations". Finally, Carnap says that languages are isomorphic if there is a transformance of the one to the other whose converse (qua relation) is also a transformance – which makes his notion of isomorphism stricter than those that have been entertained by logicians in recent years. For example, Pelletier and Urquhart (2003) define an isomorphism as a translation that has an inverse *up to logical equivalence* (in the relevant theory); and Visser (2006) entertains even more liberal notions of isomorphism. Given his principle of tolerance, Carnap would likely have said that there is no correct notion of isomorphism, and that each person is free to adopt the notion that she finds most useful.

Carnap seems not to have further developed his explicit account of transformances between languages (i.e. translations between theories), which is surprising, given that he could easily have made explicit proposals about notions such as reduction (of one theory to another) or equivalence (of two theories). As concerns the latter question, it seems that Carnap preferred to think about equivalence in terms of invariants, rather than in terms of the maps (i.e. translations) between theories. As is well known, Carnap's metasemantics committed him to the view that sentences are equivalent if they have the same observable content. With that view in the background, Carnap treated theories as equivalent if they had the same observable consequences, a notion that has subsequently been dubbed "empirical equivalence".

Carnap devotes the final chapter (Part V) of *Syntax* to explaining the upshot of this formal method for doing philosophy of science. Most notably, he says that this formal approach can help to diagnose cases of metaphysical pseudo-questions — which are often generated by a confusion of material and formal modes of speech. The distinction between these modes of speech is, according to Carnap, an idealization: our speech typically mixes the two modes freely together. However, taking every declarative sentence to be an assertion that things simply are that way (i.e. material mode) ignores the fact that the one asserting the sentence has freedom to make choices about how to formulate it. For example, a person might choose a language in which there are exactly two distinct predicate symbols; and such a choice is not tantamount to an assertion "there are exactly two distinct properties". So, when a theorist asserts something about her language (or descriptive framework), then such an assertion needs to be interpreted in a different way than her assertions of a world-directed nature. Of course, Carnap's picture here was soon to come under severe criticism from Quine.

The *Wissenschaftslogik* program was stalled to some degree by the Second World War, Carnap's immigration to the United States, and most relevantly, by Carnap's ongoing debate with Quine. However, Carnap himself continued to propose new approaches to bringing formal methods to bear on the questions of philosophy of science. Of Carnap's later-life efforts in this direction, the most notable is his use of the Ramsey sentence method to analyze a theory into analytic and synthetic parts. In this case, a theory's primitive predicates are divided into two groups, the observational and the theoretical so that the theory can be represented by a single sentence:

$$T\left(O_1, \ldots, O_n; E_1, \ldots, E_m\right),$$

where the O_i represent the observable predicates and the E_i represent the purely theoretical predicates. Carnap then says that the content of the theory is represented by the Ramsey sentence

$$\exists X_1 \ldots \exists X_m \mathrm{T}\left(O_1, \ldots, O_n; X_1, \ldots, X_m\right),$$

while the analytical (i.e. conventional) part of the theory is represented by the sentence

$$\exists X_1 \ldots \exists X_m \mathrm{T}\left(O_1, \ldots, O_n; X_1, \ldots, X_m\right) \rightarrow T\left(O_1, \ldots, O_n; E_1, \ldots, E_m\right),$$

which has since come to be called the *Carnap sentence* of the theory.

Carnap's use of the Ramsey sentence to represent the content of a theory has been deeply influential and has been the subject of much discussion. Indeed, David Lewis' famous "How to define theoretical terms" (Lewis 1970) was directly inspired by Carnap; and through Lewis, Carnap inspired both functionalism in the philosophy of mind and a certain strand of structural realism in philosophy of science. Nonetheless, it is questionable whether Carnap's use of the Ramsey sentence has provided much value for philosophy of science itself. First, many philosophers think that the method is ill-conceived, as it relies on a questionable distinction between observational and theoretical vocabulary. And even if the observational-theoretical distinction can be maintained in theory, it would be impossible to sort the predicates of any interesting scientific theory, making the method of purely abstract theoretical interest. Second, it has been argued by William Demopoulos (2022) and Neil Dewar (2019), among others, that a theory's Ramsey sentence is too weak to capture that theory's content.

Carnap's way of thinking and speaking are influenced by formal models, even when his arguments are not explicitly formal – as in his discussion of external and internal questions in "Empiricism, semantics and ontology." For Carnap, the choice of a language L specifies answers to the external questions, while internal questions are posed in L itself. Or to be more precise, the answer to an internal question is a L-sentence. For example, if one adopts the language of Peano arithmetic, then one has thereby answered the external question "are there numbers?", while the internal question "are there infinitely many prime numbers?" remains open until, at least, the specification of axioms.

Carnap's talk of internal and external questions harks back to his discussion of *Allwörter* (translated as "universal words") in *Syntax* §76. To understand what Carnap is doing here, it is important to note that he is essentially working in a many-sorted version of logic, where sort symbols denote distinct domains of objects. In informal theorizing, this sorting is done by using different variables, or simply by using different names such as "natural number" or "complex number". Carnap calls these names *Allwörter,* and he claims that positive existence claims using *Allwörter* are analytically true, as a result of the choice of the language.

Seeing Carnap's use of many-sorted logic to underwrite the analytic-synthetic distinction explains why Quine was so concerned to argue that any many-sorted theory is equivalent to a single-sorted theory, and then to propose – for the sake of simplicity – that we use only single-sorted logic. For if sort symbols are interchangeable with predicates, then analytic claims are interchangeable with synthetic, thereby destroying the distinction (Barrett 2017).

Whenever assessing Carnap's formal philosophy of science, we should remember that logic, and mathematics more generally, was under rapid development in the twentieth century. When Carnap began his career, there was no logical semantics in the sense that we understand it today, and so it is understandable that Carnap's earliest work has a strong syntactic emphasis. Carnap was not opposed to semantics when done mathematically, and when Tarski and the Polish logicians invented that field, Carnap followed their work eagerly and began to find uses for it in philosophy of science. The same goes for the notion of a translation between theories. While Gödel and other logicians occasionally invoked the notion of a translation, the concept did not receive an official definition until perhaps even the 1970s with the work of Leslaw Szczerba. Thus, it is not surprising that the notion of translation does not play a more central role in Carnap's analysis of science.

During the second half of the twentieth century, mathematics continued to develop at a rapid pace, and Carnap – caught up in his debates with Quine – seems to have drifted away from actual science and mathematics. This is unfortunate, because some of the new developments in mathematics might actually have helped inject new life into Carnap's program. For example, it does not seem that Carnap followed the exciting developments in algebraic topology, with the discovery of translation schemes between geometric objects (e.g. topological spaces) and algebraic objects (e.g. groups, rings), leading eventually to the invention of category theory (Eilenberg 1945). However, Carnap was well aware that mathematical methods evolve, and he would have encouraged philosophers to adapt their methods accordingly.

Bibliography

Eilenberg, Samuel and MacLane, Saunders. 1945. "General theory of natural equivalences." *Transactions of the American Mathematical Society* 231-294.

Russell, Bertrand. 1917. *Mysticism and Logic and Other Essays.* George Allen & Unwin Ltd.

Pelletier, Francis and Urquhart, Alasdair. 2003. "Synonymous logics." *Journal of Philosophical Logic* 259-285.

Visser, Albert. 2006. "Categories of theories and interpretations." In *Logic in Tehran*, by Ali and Kalantari, Iraj and Moniri, Mojtaba Enayat, 284-341. Cambridge University Press .

Lewis, David. 1972. "How to define theoretical terms." *Journal of Philosophy* 427–446.

Demopoulos, William. 2022. "Logical empiricist and related reconstructions of theoretical knowledge." In *On theories*, by William Demopoulos, 18–63. Harvard University Press.

Dewar, Neil. 2019. "Ramsey equivalence." *Erkenntnis* 77//99.

Barrett, Thomas and Halvorson, Hans. 2017. "Quine's conjecture on many-sorted logic." *Synthese* 3563–3582.