What scientific theories could not be

Hans Halvorson*

December 9, 2011

Abstract

According to the semantic view of scientific theories, theories are classes of models. I show that this view — if taken literally — leads to absurdities. In particular, this view equates theories that are distinct, and it distinguishes theories that are equivalent. Furthermore, the semantic view lacks the resources to explicate interesting theoretical relations, such as embeddability of one theory into another. The untenability of the semantic view — as currently formulated — threatens to undermine scientific structuralism.

1 Introduction

The twentieth century saw two proposed formal explications of the concept of a "scientific theory." First, according to the *syntactic view of theories*, a theory is a set of axioms in a formal (usually first-order) language. This view was so dominant during the first half of the twentieth century that Hilary Putnam dubbed it the "received view." But during the 1960s and 1970s, philosophers of science revolted against the received view, and proposed the alternative *semantic view of theories*, according to which a theory is a class of models. Thus Bas van Fraassen states that, "...if the theory as such, is to be identified with anything at all — if theories are to be reified — then a theory should be identified with its class of models" (van Fraassen, 1989, 222). Within a few short decades, the semantic view has established itself as the new orthodoxy. According to Roman Frigg (2006, 51), "Over the last four decades the semantic view of theories has become the orthodox view on models and theories." One only has to glance at recent writings on the philosophy of science to verify Frigg's claim: the semantic view is now assumed as the default explication of the notion of a scientific theory.

The received view was an attempt to give a precise explication of a vague concept. The view was, accordingly, judged by exacting standards; and it failed to meet these standards. It would be natural to assume, then, that the semantic view fares better when judged by these same standards — else why do so many philosophers see the semantic view as superior

*Department of Philosophy, Princeton University. hhalvors@princeton.edu

to the syntactic view? Sadly, philosophers have been too quick to jump onto the semantic bandwagon, and they have failed to test the semantic view as severely as they tested the received view. In this paper, I put the semantic view to the test, and I find that it falls short. In particular, I show that the semantic view makes incorrect pronouncements about the identity of theories, as well as about relations between theories. Consequently, the semantic view must be fixed, as must any any position in philosophy of science that depends on this inadequate view of theories.

2 What is at Stake

If you were to make a list of the top ten most important philosophical questions, I doubt that the syntactic vs. semantic view of theories would be near the top. But of course, logical connections do exist between one's preferred explication of the concept of a scientific theory and one's views on the big questions about the nature of science and of scientific knowledge — these connections just lie a bit below the surface. Thus, I devote this first section to bringing these connections to the surface, in hopes that it will motivate the reader to engage seriously with the more arcane discussion to follow.

First, I recall why some philosophers claim that the scientific realism-antirealism debate hinges in part on the tenability of the semantic view of theories. Second, I discuss some consequences of the the semantic view of theories in the philosophy of the particular sciences.

2.1 Scientific realism and antirealism

Versions of the semantic view were already present in the work of the Dutch philosopher Evert Beth as well as in the early work of Patrick Suppes. But these philosophers did not press the semantic view into the service of a particular philosophical agenda. The semantic view first became philosophically charged in the 1970s, when Bas van Fraassen used it to rehabilitate antirealism in philosophy of science.

At times, van Fraassen has indicated that his version of antirealism stands or falls with the semantic view of theories — or at least that his version of antirealism leans upon the semantic view of theories. For example, in responding to a criticism of the observable-unobservable distinction (which is presupposed by van Fraassen's antirealism), van Fraassen and Muller ascribe blame to the syntactic view of theories: "...we point to a flaw in these and similar criticisms [of the observable-unobservable distinction]: they proceed from the *syntactic view* of scientific theories whereas constructive empiricism is and has always been wedded to the *semantic view*" (Muller and van Fraassen, 2008, 197). Thus, the syntactic view supposedly provides premises for an argument against constructive empiricism; and rejecting the syntactic view allows one to neutralize these objections.

The semantic view has not only been thought to help constructive empiricism. Some (such as Ronald Giere and Fred Suppe) have also found the semantic view to be helpful for elaborating a realist philosophy of science. But perhaps the most interesting and non-trivial application of the semantic view is in developing a structural realist philosophy of science.

Recall that structural realism is the view that (stated loosely) what is important in a scientific theory is the structure that it posits or describes. In particular, suppose that T is a scientific theory that we believe to be true. What sort of attitude is this belief in T? In old-fashioned realism, believing T means believing in the existence of the entities in its domain of quantification, and believing that they stand in the relations asserted by the theory. But, as we very well know, old-fashioned realism makes it look like we change our minds about ontology during every scientific revolution. Thus, structural realism counsels a modified attitude towards T, namely we should believe that the world has the structure that is posited by T.

Since James Ladyman's seminal article of 1998, many structural realists have hitched their wagon to the semantic view of theories. Ladyman urges that, "the alternative 'semantic' or 'model-theoretic' approach to theories, which is to be preferred on independent grounds, is particularly appropriate for the structural realist" (Ladyman, 1998, 417). Ladyman then suggests that structural realists adopt Giere's account of theoretical commitment: to accept a theory means believing that the world is *similar* or *isomorphic* to one of its models. For example, a model of the general theory of relativity is a four-dimensional Lorentzian manifold; thus, believing the general theory of relativity means believing that spacetime has the structure of a four-dimensional Lorentzian manifold. In the words of Paul Thompson, "the application of the model(s) to a particular empirical system requires the extra-theoretical assumption that the model(s) and the phenomena to which they are intended to apply are *isomorphic* ..." (Thompson, 2007, 495). Others, such as van Fraassen, claim that isomorphism cannot hold between a model and the world, because "being isomorphic" is a relation that holds only between mathematical objects. Nonetheless, van Fraassen and all other semanticists claim that a theory is adequate to the extent that one of its models "represents" the world.

2.2 The semantic view applied to particular sciences

The semantic view of theories has trickled down into the consciousness of the next generation of philosophers of science. Many of these next-generation philosophers of science style themselves as "philosophers of X," where X is some particular science — for example, philosophers of physics, philosophers of biology, philosophers of psychology. Moreover, these philosophers imbibed the semantic view with their mother's milk, and their *Ausbildung* influences, for better or for worse, their judgment of issues in their subdisciplines. In this section, I remind the reader of some of the more obvious ways in which the semantic view manifests itself in the philosophy of the particular sciences.

2.2.1 Philosophy of biology

The semantic view of theories has played a visible and central role in the philosophy of biology since the 1980s. Already in 1979, John Beatty mounted a criticism of the "received view" of evolutionary theory (Beatty, 1979, 1980), and in her 1984 PhD thesis "A semantic approach to the structure of evolutionary theory," Elisabeth Lloyd claims that, "...a semantic approach to the structure of theories offers a natural, precise framework for the characterization of contemporary evolutionary theory. As such, it may provide a means with which progress on outstanding theoretical and philosophical problems can be achieved" (Lloyd, 1984, p. iii). Suffice it to say that some of the most important recent work in the philosophy of biology has rested upon, or drawn upon, the semantic view of scientific theories.[1]

2.2.2 Philosophy of psychology

The semantic view of theories has also changed the landscape in the philosophy of psychology, which is centrally concerned with questions of how the mind can be reduced to the brain — rephrased in the lingo of philosophers of science, of how naive folk theories of the mind can be reduced to neuroscience. But when we ask what it means to say that one theory is reducible to another, the answer we give will depend on our conception of what a "theory" is. As pointed out by Jordi Cat, "the shift in the accounts of scientific theory from syntactic to semantic approaches has changed conceptual perspectives and, accordingly, formulations and evaluations of reductive relations and reductionism" (Cat, 2007). As a specific example of Cat's claim, John Bickle (1993) applies the semantic view of theories to support a claim that neuroscientific eliminativism is "principled." See also (Hardcastle, 1994). Similarly, in a recent discussion, Colin Klein (2011) argues that multiple realizability arguments depend for their plausibility on the syntactic view of theories, and that from the perspective of the semantic view, these arguments are unmotivated.

2.2.3 Philosophy of physics

Up to this point, I have attempted only to describe cases where philosophers have explicitly claimed that the semantic view of theories makes a difference for some other philosophical thesis or position. But now I want to make my own claim about the importance of the semantic view: in application to the philosophy of physics, the semantic view of theories has led to *false* conclusions. One such conclusion is:

> *Model isomorphism criterion for theoretical equivalence:* If theories T and T' are equivalent then each model of T is isomorphic to a model of T'.

To see what is meant by this criterion, let's look at a couple of cases where it has been tacitly invoked.

First, Jill North applies a sort of isomorphism criterion when she argues that Hamiltonian mechanics and Lagrangian mechanics are inequivalent theories. She says that, "the equivalence of theories is not just a matter of physically possible histories, but of physically possible histories through a particular statespace structure. Hamiltonian and Lagrangian mechanics are not *equivalent* in terms of that structure. This means that they are not equivalent, period" (North, 2009, p. 79). In other worlds, the statespaces of Hamiltonian and Lagrangian

[1] See also (Lloyd, 1994) and (Thompson, 1983, 1989). For a recent review and further sources, see (Thompson, 2007).

mechanics are non-isomorphic; therefore the two theories impute different structure to the world; therefore the two theories are inequivalent.

Similarly, Erik Curiel applies a version of the model isomorphism criterion to argue that Hamiltonian and Lagrangian mechanics are inequivalent, or more particularly, that Hamiltonian mechanics does not have the resources to describe all the facts that Lagrangian mechanics describes. Curiel says:

> "...the family of kinematically possible evolutions of a dynamical system, in so far as they are characterized by interactions with no prior assumption of a geometrical structure ... cannot be naturally represented as Hamiltonian vector fields on phase space, for by definition an affine space is not isomorphic to a Lie algebra over a vector space. It follows that there is no analogous structure in the Hamiltonian representation of a system isomorphic to a dynamical system's family of interaction vector fields ..." (Curiel, 2009, 20)

In other words, Lagrangian mechanics imputes affine structure to the world; but Hamiltonian mechanics does not impute affine structure; therefore these theories are inequivalent.

The model isomorphism criterion should seem obviously correct to a structural realist who elaborates that position in terms of the semantic view of theories. For according to semantic structural realism, to accept a theory is to believe that the world is isomorphic to one of its models. Thus if two theories posit different structure — e.g. one posits affine structure, and one posits Lie structure — then they cannot both provide accurate representations of the structure of the world.

But if you think about it for a moment, you will see that this view cannot be correct. For example, Heisenberg's matrix mechanics is equivalent to Schrödinger's wave mechanics. But a matrix algebra is obviously not isomorphic to a space of wavefunctions; hence, a simple-minded isomorphism criterion would entail that these theories are inequivalent. So, something goes seriously wrong if we take the semantic view of theories seriously.

Preliminary Precisifications

Before I begin my argument against the semantic view of theories, I should clarify the terms that I will be using.

The semantic view of theories claims that:

(S_1) A theory is a class of models.

In the earliest articulations of the semantic view, the word "model" was taken to denote some sort of mathematical object. Many philosophers of science now disagree that models should be mathematical objects. Those views will not be subject to the critique that I develop in this paper. My critique is aimed at views that try to explicate the concept of a scientific theory using the concepts of contemporary mathematics.

So, within the bounds of mathematics, what is a model? We begin with the standard "elementary" concept due to Alfred Tarski. If L is a (one-sorted) first-order language, then a *L-structure* consists of a set S (the domain of quantification) as well as an assignment $R \mapsto [[R]] \subseteq S \times \cdots \times S$ for each n-place predicate symbol R of L. A *first-order theory* in L consists of a set T of sequents. Here a sequent is of the form:

$$\varphi \vdash_{\overline{x},\overline{y}} \psi,$$

where $\overline{x}$ is a sequence of variables containing all the free ones in φ, and $\overline{y}$ is a sequence of variables containing all the free ones in ψ. I assume that the reader is familiar with the definition of when an L structure $[[\cdot]]$ *satisfies* a sequent. If $[[\cdot]]$ satisfies all sequents in T, then it is said to be a *model* of T.

When semanticists say that a theory is a class of *models*, then they must not intend exactly the Tarskian definition of "model" — because then their definition would be circular. (A theory would be a class of models ... of a theory.) But to a first approximation, the semanticists are just saying that:

(S_2) A theory is a class of L-structures, for some language L.

This second definition might still be unacceptably language-bound in the eyes of van Fraassen:

> "The impact of Suppes' innovation is lost if models are defined, as in many standard logic texts, to be partially linguistic entities, each yoked to a particular syntax. In my terminology here the models are mathematical structures, called models of a given theory only by virtue of belonging to the class defined to be the models of the theory." (van Fraassen, 1989, 366)

So, van Fraassen would have us revise the definition of "model," or more accurately, of "structure": structures are not mappings from languages to (the category of) sets, but are simply the resulting "structured sets." In other words, one way to get a class of models (in van Fraassen's sense) is to take a first-order theory T and construct its class $\mathrm{Mod}(T)$ of models. But once we have arrived at $\mathrm{Mod}(T)$ we should throw away the ladder: we should forget that we used the language L in order to define the class of models. More generally, any other class $\mathcal{M}$ of mathematical structures will also count as a theory — we don't even need a first-order language L to begin with.

But what sorts of things are allowed to be in the class of models? What is a *mathematical structure*? The first-order case provides a paradigmatic example of a mathematical structure: an n-tuple of the form $\langle S, R_1, \ldots, R_{n-1} \rangle$ where the R_i are relations on the set S. Granted, for a structure such as $\langle S, R_1, \ldots, R_{n-1} \rangle$, we can easily find a language L for which it is an L-structure. But there are more complicated cases of mathematical structures — such as topological spaces — that cannot be derived in this way from a first-order language.

At present, semanticists seem to prefer the account of mathematical structures given in Nicholas Bourbaki's *Theory of Sets* (see Da Costa and French (2003)). But the argument of this paper will not depend on any nuances about the notion of a mathematical structure. For my argument to go through, I only need the semanticist to grant a weak sufficient condition on theory-hood: the class $\mathrm{Mod}(T)$ of models of a first-order theory T is (the mathematical part of a) theory in their sense.[2]

[2]A point of clarification is in order: obviously, semanticists do not reduce theories to a mere class of

3 Identity Crisis for Theories

I first show that the semantic view gives an incorrect account of the identity of theories. Its failure is complete: it identifies theories that are distinct, and it distinguishes theories that are identical (or at least equivalent by the strictest of standards).

According to the semantic view, a theory is a class of models. So, if I hand you two classes of models, say $\mathcal{M}$ and $\mathcal{M}'$, under what conditions should you say that they are the *same* theory? The anglo-american semanticists (e.g. van Fraassen) have never ventured a mathematically precise definition of the form:

(**X**) $\mathcal{M}$ is the same theory as $\mathcal{M}'$ iff ...

Thus, while criticizing the logical positivists for an implausible view of theories, the semanticists present us with a view so imprecise that it cannot be criticized. I call "no fair." The semantic view cannot be an "alternative" to the received view if it does not even offer an account of the identity of theories.

The objective of this section is to show that if theories are classes of models, then there is no plausible way to fill out (**X**). I consider three obvious proposals, in ascending order by strength:

(**E**: Equinumerous) $\mathcal{M}$ is the same theory as $\mathcal{M}'$ iff $\mathcal{M} \cong \mathcal{M}'$, i.e. there is a bijection $F : \mathcal{M} \longrightarrow \mathcal{M}'$.[3]

(**P**: Pointwise isomorphism of models) $\mathcal{M}$ is the same theory as $\mathcal{M}'$ just in case there is a bijection $F : \mathcal{M} \longrightarrow \mathcal{M}'$ such that each model $m \in \mathcal{M}$ is isomorphic to its paired model $F(m) \in \mathcal{M}'$.

(**I**: Identity) $\mathcal{M}$ is the same theory as $\mathcal{M}'$ just in case $\mathcal{M} = \mathcal{M}'$.

By seeing how these three proposals fail, it will become clear that there is really no way to formulate good identity criteria for theories when they are considered as classes of models.

3.1 The semantic view identifies distinct theories

We begin by considering criterion **E** which says that theories $\mathcal{M}$ and $\mathcal{M}'$ are equivalent iff their sets of models $\mathcal{M}$ and $\mathcal{M}'$ are isomorphic, qua sets. That criterion is intuitively far too weak: it will equate too many theories. Let's drive this point home with a a simple example from propositional logic. In what follows, we use T or T' to denote theories of first-order logic, where their individual languages (not assumed the same) are implicitly understood. When we need to be explicit, we write $L(X)$ for the language of theory X.

---

models. As explicated by Giere, Suppe, and van Fraassen, a theory is a class of models *plus* a theoretical hypothesis. But my attack has nothing to do with this second component of the semantic view of theories. I mean only to show that the first component is a mistake, i.e. a class of models is *not* the correct mathematical component of a theory.

[3]Let's ignore for the time being the problems with the set/class distinction. Let's suppose instead that the semantic view identifies theories with *sets* of models.

Example (Propositional Theories)**.** Let $L(T)$ be a propositional language with a countable infinity of 0-place predicate symbols (i.e. propositional constants) $p_1, p_2, \ldots$. We work throughout with classical logic, so $L(T)$ is equipped with connectives $\wedge, \vee, \longrightarrow, \neg$. Let T be the empty theory in $L(T)$, i.e. the theory whose only consequences are tautologies. Let $L(T')$ add to $L(T)$ a new propositional constant q, and let T' be given by the infinite set of axioms $\{q \vdash p_i : i \in \mathbb{N}\}$.

Fact. Let $\mathcal{M}$ be the set of models of T, and let $\mathcal{M}'$ be the set of models of T'. Then $\mathcal{M}$ has the same cardinality as $\mathcal{M}'$.

Proof. Obviously T has $2^{\aleph_0}$ models, i.e. truth-valuations. For T', let v be a truth-valuation. On the one hand, if $v(q) = 1$ then $v(p_i) = 1$ for all i. On the other hand, $v(q) = 0$ is consistent with any assignment of truth-values to the p_i. Thus T' has $2^{\aleph_0}$ models. □

Thus, criterion **E** tells us that T and T' are the same theory. But these theories are intuitively inequivalent; and our intuition here is backed by the standard account of definitional equivalence of (syntactically formulated) theories.

Definition. Let T and T' be theories. Let $F : L(T) \longrightarrow L(T')$ be a map of the underlying languages that takes variables to variables, and n-ary predicate symbols to wffs. F can then be canonically extended to map terms of $L(T)$ to terms of $L(T')$, and formulae of $L(T)$ to formulae of $L(T')$. We say that F is an *interpretation* of T in T' just in case for each axiom $\varphi \vdash \psi$ of S, $F(\varphi) \vdash F(\psi)$ is a theorem of T'.[4]

Of course, if there is no interpretation of T into T', then the two theories cannot be definitionally equivalent.

Definition. Let T and T' be theories, and let $F : T \longrightarrow T'$ and $G : T \longrightarrow T'$ be interpretations. We say that G is a *weak inverse* of F just in case for each wff φ of $L(T)$, $GF(\varphi)$ is T-provably equivalent to φ, and for each wff ψ of $L(T')$, $FG(\psi)$ is T'-provably equivalent to ψ. If there is a weakly invertible interpretation $F : T \longrightarrow T'$, then T and T' are said to be *definitionally equivalent.*

Fact. The theories T and T' are not definitionally equivalent.

Proof. Suppose for reductio ad absurdum that $F : T \longrightarrow T'$ and $G : T' \longrightarrow T$ give a definitional equivalence. Then Gq is a T-atom under the implication relation. Indeed, if $r \vdash Gq$ then $Fr \vdash FGq \simeq q$. Since q is an atom relative to T' provability, either $Fr \simeq \bot$ or $Fr \simeq q$. In the former case, $r \simeq GFr \simeq \bot$; in the latter case $r \simeq GFr \simeq Gq$. Thus, Gq is an atom relative to T provability, which is a contradiction.[5] □

[4]For variations on this definition, see (Hodges, 1993, 219ff) and (Szczerba, 1977, 133). We allow predicate symbols to be mapped to formulas — thus allowing, for example, interpretations that take a predicate to an open sentence.

[5]It is perhaps easier to see what is going on here if one looks at the Stone Space of the corresponding Lindenbaum algebras. The Stone space for T is the Cantor space C. The Stone space for T' is $C \sqcup \{*\}$. These spaces have the same cardinality, but are not homeomorphic.

To summarize this example: there is a standard criterion of equivalence of syntactically formulated theories, namely definitional equivalence. By this criterion, the theories T and T' are inequivalent. But the semantic view of theories reduces T and T' to their respective sets of models, $\text{Mod}(T)$ and $\text{Mod}(T')$, and these two sets are isomorphic, i.e. equinumerous. Moreover, the semanticist cannot distinguish $\text{Mod}(T)$ from $\text{Mod}(T')$ on the grounds that the former consists of mappings from the language $L(T)$ and the latter consists of mappings from the language $L(T')$. Indeed, the semanticist has precluded reference to language in individuating theories. Therefore the semantic view identifies theories that should be treated as distinct.

Example (From Propositional to Predicate)**.** The semanticist might not know how to respond to the previous example: when he thinks of a "model," his paradigm example is an L-structure where L is a predicate language. Since the previous example uses 0-place predicates (i.e. proposition symbols), one might worry that it is not typical. However, we can easily modify the example to overcome this worry.

Let $L(T)$ be the language with a countable infinity of 1-place predicate symbols $P_1, P_2, P_3, \dots,$ and let T have a single axiom $\exists_{=1}x(x = x)$ (there is exactly one thing). Let $L(T')$ be the language with a countable infinity of 1-place predicate symbols $Q_0, Q_1, Q_2, \dots,$ and let T' have axioms $\exists_{=1}x(x = x)$ as well as $Q_0x \vdash_x Q_ix$ for each $i \in \mathbb{N}$.

Clearly T and T' have the same number of models; so they are equivalent according to criterion **E**. What is more, every model of T is isomorphic to a model of T' and vice versa. Indeed, a model of T has a domain with one object that has a countable infinity of monadic properties, and model of T' also has a domain with one object that has a countable infinity of monadic properties. Therefore, T and T' are equivalent according to criterion **P**.

And yet, T and T' are intuitively inequivalent. We might reason as follows: the first theory tells us nothing about the relations between the predicates; but the second theory stipulates a non-trivial relation between one of the predicates and the rest of them. Again, our intuition is backed up by the syntactic account of equivalence: the theories T and T' are *not* definitionally equivalent. Therefore both **E** and **P** identify theories that they should not.

Example (Categorical Theories)**.** For this example, we recall that there is a pair of first-order theories T and T', each of which is κ-categorical for all infinite κ, but which are *not* definitionally equivalent to each other. (Many such examples can be found, for example, in the work of Boris Zil'ber on totally categorical theories (Zil'ber, 1993). In fact, Zil'ber has classified these theories in terms of geometric invariants.) By categoricity, for each cardinal κ, both T and T' have a unique models (up to isomorphism) with domain of size κ. Thus, there is an invertible mapping that pairs the size-κ model of T with the size-κ model of T'. Hence, the equinumerosity criterion **E** would entail that T and T' are equivalent theories. What is more, for each cardinal number κ, T has a unique model of cardinality κ and T' has a unique model of cardinality κ. Therefore, for each model m of T there is a model m' of T' and a bijection (isomorphism of sets) $i : m \longrightarrow m'$. Thus, every model of T is isomorphic to a model of T', and by criterion **P** they are the same theory — even though they fail to be definitionally equivalent.[6]

3.2 The semantic view distinguishes identical theories

We have just seen that the semantic view would equate theories that are intuitively distinct. We will now see that the semantic view also makes the opposite mistake: it would distinguish theories that are intuitively equivalent.

Begin with the most stringent criterion **I**, which says that no two *distinct* sets of models can represent the same theory. Such a criterion is prima facie too strict, because it would not allow for the same theory to admit alternative axiomatizations in distinct languages L and L'. Consider the following example:

Example (Autosets vs. Groups)**.** An *autoset* is a set with a transitive action on itself. The theory of autosets can be formulated in a language $L(T)$ with a single binary function symbol $\circ$, for which we use infix notation. Let T be given by the following three axioms:

$$\vdash_{x,y,z} (x \circ y) \circ z = x \circ (y \circ z) \qquad \vdash_{x,y} \exists z(x \circ z = y) \qquad \vdash_{x,y} \exists z(z \circ x = y).$$

A model of T is an autoset.

The theory of groups can be formulated in a language $L(T')$ with a binary function symbol $\circ$, a unary function symbol i, and a constant symbol e. Let T' consist of the standard group theory axioms: associativity, identity, and inverses.

By the lights of criterion **I**, the theories T and T' are distinct After all, a model of T is a pair $\langle S, \circ \rangle$ and a model of T' is a quadruple $\langle G, \circ, i, e \rangle$. Two is not equal to four, so the class of autosets is not identical to the class of groups.

But a simple exercise in abstract algebra shows that the theory of autosets is definitionally equivalent to the theory of groups. In particular, T entails that the predicate

$$Px \;\equiv\; \exists y(y \circ x = y = x \circ y),$$

is uniquely satisfiable, hence T defines a constant symbol e. Similarly, T entails that the relation

$$Rxy \;\equiv\; x \circ y = e,$$

is functional, and hence T defines a function symbol i. In other words, although an autoset is not a group, each autoset carries *definable* group-theoretic structure (an identity element and an inverse function). But the very notion of definability is not available via a purely semantic approach: the notion of definability presupposes reference to the language in which the theories were formulated.

Thus, criterion **I** fails to allow intuitive cases of theoretical equivalence — cases which are supported by the syntactic notion of definitional equivalence. Does the liberalized criterion

[6] One might point out here that the map $i : m \longrightarrow m'$ is a bijection but not necessarily an isomorphism of models. But there is *no* definition of isomorphism between models of different theories. See the next section, especially the final paragraph.

P fare any better? If two theories are equivalent, does it follow that each model of the first is isomorphic to a model of the second? The following two examples provide a negative answer, thereby establishing that even criterion **P** is too strict.

Example (Boolean Algebras)**.** Let $\mathcal{B}$ be the class of complete atomic Boolean algebras (CABAs), i.e. an element B of $\mathcal{B}$ is a Boolean algebra such that each subset $S \subseteq B$ has a least upper bound $\bigvee(S)$, and such that each element $b \in B$ is a join $b = \bigvee b_i$, where the b_i are atoms in B. Now let $\mathcal{S}$ be the class of sets.

What does the semantic view say about the relation between the theories $\mathcal{B}$ and $\mathcal{S}$? Obviously $\mathcal{B} \neq \mathcal{S}$, and so criterion **I** entails that these theories are inequivalent. Furthermore, an arbitrary set S cannot be equipped with operations that make it a Boolean algebra; e.g. there is no Boolean algebra whose underlying set has cardinality 3. Thus, there are structures in the class $\mathcal{S}$ that are not isomorphic to any structure in the class $\mathcal{B}$. Hence, criterion **P** entails that these two theories are inequivalent.

I claim, however, that the "theory of sets" is equivalent to the "theory of complete atomic Boolean algebras." Indeed, to each set S, we can associate a CABA, namely its powerset $F(S)$ with the operations of union, intersection, and complement. Furthermore, the set $G(F(S))$ of atoms of $F(S)$ is naturally isomorphic (as a set) to S. In the opposite direction, to each complete atomic Boolean algebra B we can assign a set, namely the set $G(B)$ of its atoms; and it follows that B is isomorphic (as a Boolean algebra) to $F(G(B))$. To summarize, there are a pair of mappings $F : \mathcal{S} \longrightarrow \mathcal{B}$ and $G : \mathcal{B} \longrightarrow \mathcal{S}$ that are inverse to each other, up to isomorphism.

The previous example might not have convinced the semanticist to change his ways. He might be willing to bite the bullet and say that the theory of sets is not equivalent to the theory of CABAs. The problem is that we haven't given an independent reason for thinking that these are equivalent theories. We fix this problem with the next example in which we again display two theories T and T' whose models are not individually isomorphic, but where now T and T' can be proven to be definitionally equivalent.

Any interpretation $F : T \longrightarrow T'$ between theories gives rise naturally to a map $F^* : \mathrm{Mod}(T') \longrightarrow \mathrm{Mod}(T)$ between their models. To see what is going on here, consider two prominent classes of examples. First, let $L(T')$ result from adding a new relation symbol to $L(T)$, but let $T' = T$ and let $F : T \longrightarrow T'$ be the obvious "embedding" of $L(T)$ into $L(T')$. Then F^* takes a model of T' and "forgets" what that model assigned to the new relation symbol. Second, let $L(T') = L(T)$, but let T' result from adding some new axioms to T, and let $F : T \longrightarrow T'$ be the interpretation of T into T' that results from the identity map on $L(T) = L(T')$. Then F^* takes a model of T' and shows us that it is also a model of T.

Thus, interpretations induce model maps and, in particular, definitional equivalences induce model maps.

Proposition. *A definitional equivalence of theories does not necessarily entail that these theories have isomorphic models. In particular, there are first-order theories T and T', and a definitional equivalence $F : T \longrightarrow T'$. Furthermore, for any definitional equivalence $F : T \longrightarrow T'$, there is a model m' of T' such that the cardinality of m' is not equal to the cardinality of $F^*(m')$.*

Proof. Let T be the empty theory formulated in a language with a single binary predicate R. Let T' be the empty theory formulated in a language with a single ternary predicate S. Myers (1997) proves that there is a definitional equivalence consisting of maps $F : T \longrightarrow T'$ and $G : T' \longrightarrow T$.

Now we prove that there is no definitional equivalence $F : T \longrightarrow T'$ such that $\mathrm{Card}(n) = \mathrm{Card}(F^*(n))$ for all models n of T'. For this, we only need the simple fact that definitional equivalences are *conservative* with respect to isomorphisms between models; that is, if $F^*(n) \equiv F^*(n')$ then $n \equiv n'$. (This follows from the fact that F has a pseudo-inverse G, and G^* preserves isomorphisms. That is, if $F^*(n) \equiv F^*(n')$ then $n \equiv G^*F^*(n) \equiv G^*F^*(n') \equiv n'$.) Now let A be the set of isomorphism classes of models n of T' such that $\mathrm{Card}(n) = 2$. Let B the set of isomorphism classes of models m of T such that $\mathrm{Card}(m) = 2$. Clearly B is a finite set that is larger than A. By conservativeness, $F^*(B)$ is larger than A, hence there is a $n \in B$ such that $F^*(n) \notin A$. But then $\mathrm{Card}(n) = 2$ and $\mathrm{Card}(F^*(n)) \neq 2$. □

From this proposition, we draw a crucial interpretive corollary:

> **Theoretical Equivalence is Global:** *An equivalence between two classes of models is not necessarily induced pointwise by isomorphisms of individual models.*
>
> That is, two classes of models $\mathcal{M}$ and $\mathcal{M}'$ might be equivalent even when *there is no sense in which individual models in $\mathcal{M}$ are isomorphic to individual models in $\mathcal{M}'$.* (Here the phrase "no sense" is validated by the fact that the paired models can have domains of different cardinality. Two models with different cardinalities cannot be isomorphic in any sense of the word.)

Before proceeding, we draw two further philosophical corollaries.

First, the global nature of equivalence shows the incorrectness of the "model isomorphism criterion for theoretical equivalence." Recall that the model isomorphism criterion would rule two theories inequivalent if the models of the one theory are not isomorphic to the models of the other theory. (I claimed that such a criterion is at work in recent arguments for the inequivalence of Hamiltonian and Lagrangian mechanics.) But we have seen that definitionally equivalent theories need not have pairwise isomorphic models. Therefore, pointing out that two theories have non-isomorphic models does not settle the question of whether those theories are equivalent.

Second, the globality of theoretical equivalence spells trouble for structural realism — at least those versions that cash representation out in terms of isomorphism or similarity. According to these versions of structural realism, a theory is true just in case it accurately represents the structure of the world, or more precisely:

> A theory is true just in case it has a model m that is isomorphic to the world w.

But which formulation of the theory should we choose? Suppose that the theory could be formulated either by the class $\mathcal{M}$ or by the class $\mathcal{M}'$ of models, but that (as in the case above) the models in $\mathcal{M}$ are not isomorphic to the models in $\mathcal{M}'$. Then which formulation of the theory should we use to evaluating the isomorphism claim? If the world is isomorphic to a model in $\mathcal{M}$, then it is *not* isomorphic to a model in $\mathcal{M}'$.

Of course, a standard realist response to this problem would be to assert privilege for a certain formulation of the theory. Although there might be a mathematical equivalence between the classes $\mathcal{M}$ and $\mathcal{M}'$, the realist will take one of the classes as dividing nature at the joints. But such a response will hardly be attractive to a structural realist, who would not ascribe ontological import to differences of formulation.[7]

4 Relations between Theories

We have already shown that the semantic view fails miserably at individuating theories: it conflates distinct theories, and it is blind to some equivalences between theories. But one might hope that these are only failures in theory, and that in practice, the semantic view gets things right. What I mean here by, "in practice," is the use to which philosophers of science put the semantic view of theories. Philosophers of science have used the semantic view to support their views of the observable/unobservable distinction, and of intertheoretic reduction, among other things. One might hope that the failures of the semantic view noted above do not taint these more consequential discussions, or the conclusions drawn therefrom. But I have bad news: the semantic view also gives wrong answers about when one theory is a subtheory of another, and about when one theory is reducible to another. All in all, conclusions drawn from the semantic view of theories are completely unreliable.

Let us look closely now at the famous motivating example given by van Fraassen in *The Scientific Image* (van Fraassen, 1980). Consider the following geometric axioms:

$A1$ For any two lines, there is at most one point that lies on both.

$A2$ For any two points, there is exactly one line that lies on both.

$A3$ On every line there lie at least two points.

$A4$ There are only finitely many points.

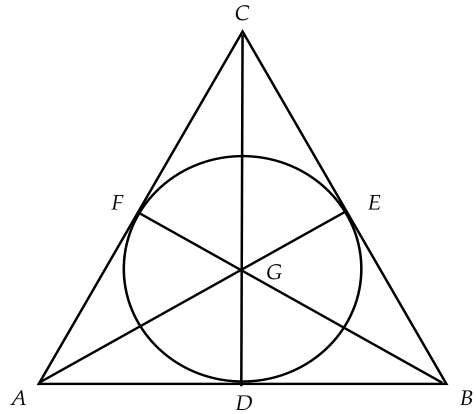
$A5$ On any line there lie infinitely many points.

Van Fraassen then defines three theories: the core theory T_0 has axioms $A1$, $A2$ and $A3$; theory T_1 results from adding $A4$ to the core theory, and theory T_2 results from adding $A5$ to the core theory. Figure 1 shows that both T_1 and T_2 are consistent: the diagram m_1 consisting of just the seven points $A, \dots, G$ is a model of T_1, and the entire drawing m_2 is a model of T_2.

According to van Fraassen, a semantic approach gives a superior account of the relationship between these theories than does a syntactic approach. In particular, he claims first that a syntactic view can see only that T_1 and T_2 are inconsistent: "logic tells us that [T_1

[7] Thanks to Kyle Stanford for this point.

Figure 1: Seven Point Geometry



and T_2] are inconsistent with each other, and there is an end to it.' (van Fraassen, 1980, 43). In contrast, van Fraassen claims that a semantic view sees interesting relationships between T_1 and T_2: in particular, each model of T_1 is *embeddable* in a model of T_2.

> "... that seven-point structure can be *embedded* in a Euclidean structure ... This points to a much more interesting relationship between the theories T_1 and T_2 than inconsistency: every model of T_1 can be embedded in (identified with a substructure of) a model of T_2. This sort of relationship, which is peculiarly semantic, is clearly very important for the comparison and evaluation of theories, and is not accessible to the syntactic approach." (van Fraassen, 1980, p43–44)

Thus, a semantic view is supposed to show its superiority as a means for analyzing relations between theories.

In the years since van Fraassen first used "embeddability" to formulate constructive empiricism, several philosophers have been at pains to argue that embeddability — and other interesting relations between theories — can also be explicated via syntactic means; see, for example, (Turney, 1990). If that's so, then the syntactic approach can do just as much as the semantic approach. But I wish to take a harder line: I claim that the semantic approach *cannot* explicate the relation of embedding between theories.

What does it mean to say that the seven point model m_1 is *embeddable* in the Euclidean model m_2? What is the definition of "embedding" that is being used? Obviously, an embedding cannot be just any function, e.g. the function that maps everything to a single point is not an embedding. Similarly, an embedding cannot simply be a one-to-one map; because such maps could also mess up geometrical relations.

The claim that m_1 can be embedded into m_2 is true in context, namely the context of the background theory T_0. In particular, if we think of m_1 and m_2 as being represented by drawings on transparencies, then there is a *rigid motion* that carries m_1 on top of m_2. But "rigid motion" is a theory-laden concept: it denotes a transformation that preserves the relations definable in the core theory T_0. Generalizing from this example, we derive the following take-away point:

> *Theory-dependence of embedding:* The notion of a "permissible embedding" of one structure/model into another structure/model depends on some background theory. In particular, "m is embeddable into m'" is a relation between models m and m' of a *single* theory.

An obvious corollary of the theory-dependence of embedding is that "embeddable" is *not* a relation that holds between models of two different theories; and so this notion cannot immediately be used to explicate concepts such as "empirical adequacy of a theory" or "reducibility of one theory to another."

On a conciliatory note, I do grant that there is an interesting relation between van Fraassen's theories T_1 and T_2 — but the relation probably shouldn't be called "embeddability", since that term already has a technical use in model theory, as a relation between models of a *single* theory. Rather, T_1 and T_2 are both, by definition, specializations of the theory T_0. That is, they result from T_0 by adding some axioms. Whenever a theory T' is a specialization of T, then there is obviously a *syntactic* interpretation map $F : T \longrightarrow T'$, namely the identity map. In the case at hand, we thus have two interpretations

$$\Pi_1 : T_0 \longrightarrow T_1, \qquad \Pi_2 : T_0 \longrightarrow T_2,$$

and these yield model maps

$$\Pi_1^* : \mathrm{Mod}(T_1) \longrightarrow \mathrm{Mod}(T_0), \qquad \Pi_2^* : \mathrm{Mod}(T_2) \longrightarrow \mathrm{Mod}(T_0).$$

Furthermore, since "line" can be defined in terms of pairs of distinct points, for each model m_1 of T_1, there is a model m_2 of T_2 such that $\Pi_1^*(m_1)$ is embeddable (relative to the theory T_0) into $\Pi_2^*(m_2)$.

In short, the models of T_1 can be compared with the models of T_2 because they can both be thought of as models of the common core theory T_0, which comes equipped with a notion of an embedding between its models. But without the syntactically specified theory T_0, we wouldn't know how to compare models of T_1 with models of T_2.

To further clarify issues here, it might help to look at a simpler example that shares the relevant features of van Fraassen's example. Consider the following two theories:

$$E_2 = \text{there are exactly two things.}$$
$$E_3 = \text{there are exactly three things.}$$

Following van Fraassen's line of reasoning, we might say: On the one hand, there is no interesting *syntactic* relation between E_2 and E_3; they are simply inconsistent. On the other hand, each model of E_2 can be embedded in a model of E_3, an important fact that is visible only from a *semantic* perspective. Is this a good analysis of what is going on here?

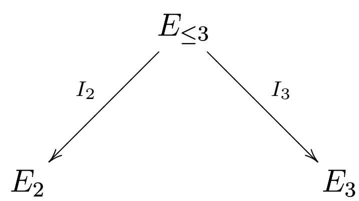
Let's unpack the example. For each $i \in \mathbb{N}$, define the first-order sentences $E_{\leq i}$ (there are at most i things), $E_{\geq i}$ (there are at least i things), and E_i (there are exactly i things). Then for all $i, j \in \mathbb{N}$ with $i \leq j$,

$$E_i \iff E_{\leq i} \wedge E_{\geq i}, \qquad E_{\leq i} \wedge E_{\leq j} \iff E_{\leq i}.$$

Note also that $E_{\geq i}$ is pure existential, i.e. a string of existential quantifiers applied to a quantifier-free sentence. In particular, E_3 results from $E_{\leq 3}$ by adding a single existential axiom. From these facts we note the obvious further fact that both E_2 and E_3 result from adding axioms to $E_{\leq 3}$:

$$E_2 \iff E_{\leq 3} \wedge E_2, \qquad\qquad E_3 \iff E_{\leq 3} \wedge E_{\geq 3}.$$

as depicted in the diagram of interpretations:



where I_2 and I_3 are the identity interpretations. Thus, we conclude:

> There is an interesting *syntactic* relation between E_2 and E_3, namely, they are specializations of a common theory $E_{\leq 3}$; moreover, E_3 results from adding a pure existential axiom to $E_{\leq 3}$.

I claim further that any interesting semantic relation between E_2 and E_3 is nothing but a mirror image of this basic syntactic relation.

Incidentally, the considerations of this section point to yet another fatal defect in the "pointwise isomorphism" criterion **P** for theoretical equivalence. Recall that criterion **P** says that two classes of models $\mathcal{M}$ and $\mathcal{M}'$ should be deemed equivalent just in case there is a bijection $F : \mathcal{M} \longrightarrow \mathcal{M}'$ such that each model $m \in \mathcal{M}$ is isomorphic to its paired model $m' = F(m) \in \mathcal{M}'$. We have just seen, however, that the relation "m is embeddable in m'" is theory-dependent. For the same reasons, the relation "m is isomorphic to m'" is also theory-dependent, since what counts as an "isomorphism" will depend on the theory in question. Thus, criterion **P** cannot possibly give an adequate account of the identity of theories.

5 Theories versus Formulations

We turn to a final purported advantage of the semantic view of theories: the semantic view is supposed to be language independent.

Any non-trivial first-order theory admits alternative formulations. First, within a single language L, a given theory can be axiomatized in distinct ways, say with axiom set T or axiom set T'. Of course, such a superficial difference can be remedied by taking a theory to be a set of sentences that is closed under the consequence relation; thus $Cn(T) = Cn(T')$ is the same theory. A more seriously difficult is posed by theories T and T' formulated in different languages, i.e. $L(T) \neq L(T')$.

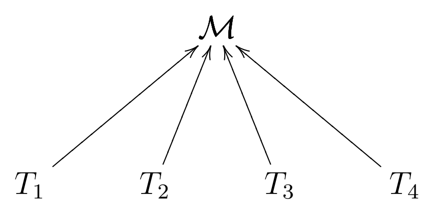
Frustration with trying to give conditions for equivalence between theories in different languages may be responsible for the semanticists search for "invariant" formulations of theories. According to Suppe, "...theories are not collections of propositions or statements, but rather are extra-linguistic entities which may be described or characterized by a number of different linguistic formulations" (Suppe, 1977, 221). Similarly, van Fraassen indicates that the class of models is the invariant that lies behind different formulations: "...while a theory may have many different formulations, its set of models is what is important" (van Fraassen, 2008, 309). Even more strongly, van Fraassen and Muller state:

> "In the semantic approach, we pride ourselves on not being so languagebound as one was during the hegemony of the syntactic view. Here a theory is not identified with or through its formulation in a specific language, nor with a class of formulations in specific languages, but through or by a class of models." (Muller and van Fraassen, 2008, 201)

Finally, van Fraassen attributes the failure of the syntactic view of theories to its attachment to formulations rather than to the underlying invariant:

> "In any tragedy, we suspect that some crucial mistake was made at the very beginning. The mistake, I think, was to confuse a theory with the formulation of a theory in a particular language." (van Fraassen, 1989, 221)

Thus, the picture given by semanticists is of a many-to-one relationship between formulations of a theory in a particular language (syntax) and a single class of models (semantics). In a picture:

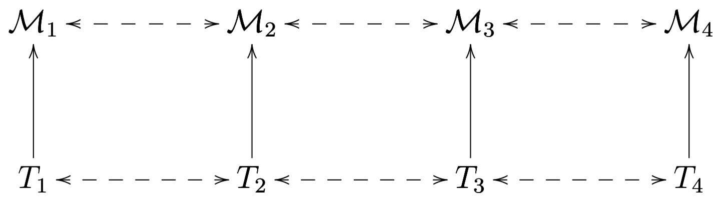


where $T_1, T_2, \ldots$ are theory formulations, and $\mathcal{M}$ is the 'invariant' class of models. Thus, the semanticists think of the relation between syntactic axiomatizations and classes of models as many-to-one, and analogous to the relation between coordinates and underlying geometric objects, or to the relation between sentences and propositions.

The picture of the class of models as an 'invariant' carries some initial plausibility — witness, e.g., the case of different axiomatizations of group theory, or different axiomatizations of vector space theory. Why would we call two different syntactic theories different formulations of the *same* theory unless they had the same class of models? But that is misleading: in interesting cases of alternative formulations, not only are the formulations different, but so are the very classes of models.

But there is a correct picture lurking in the neighborhood: when we say, correctly but imprecisely, that two theories T and T' have the "same" models, we mean that the models of T are somehow interconvertible with the models of T'. For example, every group can be converted into an autoset by "forgetting" its inverse operation and its identity element; similarly, every autoset can be converted into a group by defining an identity element and an inverse function. In fact, model theorists have a name for this sort of interconvertibility: it is called "mutual definability." However, the notion of definability requires reference to language, and so is not available on a pure semantic view of theories.

As we have now detailed at great length, there are equivalent theories (e.g. different axiomatizations of group theory) that have distinct classes of models. Thus, as opposed to the many-to-one picture, a more accurate picture of the relation between syntactic structures and semantic structures (for a single theory) is the following:



Here the dotted lines are supposed to indicate some sort of *equivalence*, a notion which should be discussed at greater length. On the bottom (syntactic) row, we already have many good examples of equivalence, such as different axiomatizations of group theory. And for the top (semantic) row, we also have some fairly simple, but uncontroversial examples of equivalence, e.g. models of group theory versus models of autoset theory.

6 Esquisse d'un Programme

The semantic view of theories is plagued by many ills. But can it be cured? Should we try to cure it? Some might say that the problems here are caused by *over-technicalizing* the concept of a scientific theory, i.e. with trying to provide a formal analysis of the concept. Such seems to be the view of Gabriele Contessa:

> "Philosophers of science are increasingly realizing that the differences between the syntactic and the semantic view are less significant than semanticists would have it and that, ultimately, neither is a suitable framework within which to think about scientific theories and models. The crucial divide in philosophy of science, I think, is not the one between advocates of the syntactic view and advocates of the semantic view, but the one between those who think that philosophy of science needs a formal framework or other and those who think otherwise." (Contessa, 2006, 376)

I agree, and disagree. I agree that the debate between syntactic and semantic views is less significant than was advertised by the semanticists. However, Contessa's implication is that we have to make an either–or choice between a "formal framework" for philosophy of science and some alternative. But what would "informal philosophy of science" look like? Should the informal philosopher of science eschew all use of mathematical notation or concepts? But how then should the informal philosopher of science discuss quantum mechanics or general relativity or string theory?

Indeed, there is another crucial divide that lies even deeper than the one indicated by Contessa: the divide between those who want to give a unified framework for all the sciences, and those who do not aspire for such a framework. For those who do not aspire for a unified framework, it would be legitimate to employ a formal framework for those sciences that themselves employ a formal framework (e.g. mathematical physics), and a less formal framework for those sciences that themselves are less formalized (e.g. evolutionary biology).

Patrick Suppes famously said that, "philosophy of science should use mathematics, and not meta-mathematics" (see van Fraassen, 1980, 65). But meta-mathematics is part of mathematics! And there is no clear distinction to be drawn between the two approaches. Furthermore, for some sciences, there is no distinction to be made between discussing a scientific theory "in its own language", or we might say "on its own terms," and discussing a scientific theory "in formal language." Philosophers of science should not be afraid of using all the tools that scientists use, including mathematical logic!

Indeed, the defects in the semantic view that I have identified are not due to over-mathematization *per se*; rather, these defects are due to inadequate mathematization. More precisely, the semantic view was not wrong to treat theories as collections of models; rather, it was wrong to treat theories as *nothing more than* collections of models. Beginning with a syntactically formulated theory T, we can construct its class $\mathrm{Mod}(T)$ of models. But we have more information than just the collection of models: in particular, we have information about *relations* between these models. For example, any sentence φ induces a relation on $\mathrm{Mod}(T)$, namely the relation "m assigns the same truth value as m' to φ." There are other such relations, but none of these relations can be seen if we reduce a theory to a bare set of models.

This point has long been known to mathematicians and logicians; indeed, this point is a straightforward corollary of Stone's duality theorem for Boolean algebras.

Given a propositional theory T, consider its set $\mathrm{Mod}(T)$ of models. Can we recover T from $\mathrm{Mod}(T)$? Does the set $\mathrm{Mod}(T)$ contain as much information as the syntactic object T? Obviously not: as we have seen, there are distinct theories T and T' whose sets of models $\mathrm{Mod}(T)$ and $\mathrm{Mod}(T')$ are indistinguishable qua bare sets. How then should $\mathrm{Mod}(T)$ and $\mathrm{Mod}(T')$ be distinguished from each other? That was the question that Marshall Stone took up in the 1930s; and Stone's answer was that $\mathrm{Mod}(T)$ and $\mathrm{Mod}(T')$ have natural topological structure in terms of which they differ. In particular, define a topology on $\mathrm{Mod}(T)$ by saying that a sequence (m_i) of models converges to a model m just in case for each sentence p, the truth value $m_i(p)$ converges to the truth value $m(p)$. Then the theory T can be recovered (up to definitional equivalence) by extracting the compact open subsets of the topological space $\mathrm{Mod}(T)$. In other words, the topological space $\mathrm{Mod}(T)$ *does* contain all the information as the syntactic object T.

Fine, you might say: for the trivial case of propositional theories, we could rehabilitate the semantic view of theories by taking a theory to be a *structured set* of models, namely a topological space of models. But this strategy will not obviously work in the general case — because Stone's theorem only works for propositional theories.

But here there is good news to report: generalizations of Stone's duality theorem have been proven by Michael Makkai (1993), and more recently by Steve Awodey and Henrik Forssell (2008; 2010). The technical details of these results are far too complex to summarize here. Suffice it to say, however, that the question of what structure is naturally possessed by a class of models is highly non-trivial, and calls for some serious (meta)mathematical research. Furthermore, the outcome of these investigations holds interest for anyone who wishes to understand the identity criteria for (formalized) theories, and the relations that can hold between (formalized) theories — in particular, for all philosophers of the exact sciences. Despite philosophy of science's recent trend towards de-formalization and imprecision, (meta)mathematicians continue to provide us with invaluable tools for discussing philosophical issues with clarity and rigor. We have only ourselves to blame if we do not take advantage of these tools.

Acknowledgments: The final version was improved by comments from Jeremy Butterfield, Bas van Fraassen, Gideon Rosen, Jim Weatherall and an anonymous referee for *Philosophy of Science*.

References

Awodey, S. and H. Forssell (2010). First-order logical duality. http://arxiv.org/abs/1008.3145.

Beatty, J. (1979). *Traditional and semantic accounts of evolutionary theory.* Ph. D. thesis, University of Indiana.

Beatty, J. (1980). What's wrong with the received view of evolutionary theory? In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1980, pp. 397–426. Philosophy of Science Association.

Bickle, J. (1993). Connectionism, eliminativism, and the semantic view of theories. *Erkenntnis 39*(3), 359–382.

Cat, J. (2007). The unity of science. *The Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/entries/scientific-unity/

Contessa, G. (2006). Scientific models, partial structures and the new received view of theories. *Studies in History and Philosophy of Science 37*(2), 370–377.

Curiel, E. (2009). Classical mechanics is Lagrangian; it is not Hamiltonian; the semantics of physical theory is not semantical. Unpublished manuscript: London School of Economics.

Da Costa, N. and S. French (2003). *Science and partial truth: A unitary approach to models and scientific reasoning*. Oxford University Press, USA.

Forssell, H. (2008). *First-order logical duality*. Ph. D. thesis, Carnegie Mellon University.

Frigg, R. (2006). Scientific representation and the semantic view of theories. *Theoria 55*, 37–53.

Hardcastle, V. (1994). Philosophy of psychology meets the semantic view. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1994, pp. 24–34. Philosophy of Science Association.

Hodges, W. (1993). *Model theory*. Cambridge University Press.

Klein, C. (2011). Multiple realizability and the semantic view of theories. *Philosophical Studies*, forthcoming.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science 29*(3), 409–424.

Lloyd, E. A. (1984). *A semantic approach to the structure of evolutionary theory*. Ph. D. thesis, Princeton University.

Lloyd, E. A. (1994). *The structure and confirmation of evolutionary theory*. Princeton University Press.

Makkai, M. (1993). *Duality and definability in first order logic*. American Mathematical Society.

Muller, F. and B. van Fraassen (2008). How to talk about unobservables. *Analysis 68*(299), 197–205.

Myers, D. (1997). An interpretive isomorphism between binary and ternary relations. In J. Mycielski et al. (Eds.), *Structures in logic and computer science: a selection of essays in honor of A. Ehrenfeucht*, pp. 84–105. Springer.

North, J. (2009). The "structure" of physics: a case study. *Journal of Philosophy 106*, 57–88.

Suppe, F. (1977). *The structure of scientific theories*. University of Illinois Press.

Szczerba, L. (1977). Interpretability of elementary theories. In R. Butts and J. Hintikka (Eds.), *Logic, Foundations of Mathematics and Computability Theory*, pp. 129–145. Dordrecht: D. Reidel Publ. Co.

Thompson, P. (1983). The structure of evolutionary theory: A semantic approach. *Studies in History and Philosophy of Science 14*(3), 215–229.

Thompson, P. (1989). *The structure of biological theories.* State University of New York Press.

Thompson, P. (2007). Formalisation of evolutionary biology. In M. Matthen and C. Stephens (Eds.), *Handbook of the Philosophy of Biology*, pp. 485–523. New York: North Holland.

Turney, P. (1990). Embeddability, syntax, and semantics. *Journal of Philosophical Logic 19*, 429–451.

van Fraassen, B. (1980). *The scientific image.* Oxford University Press.

van Fraassen, B. (1989). *Laws and symmetry.* Oxford University Press.

van Fraassen, B. (2008). *Scientific representation: paradoxes of perspective.* Oxford University Press.

Zil'ber, B. (1993). *Uncountably categorical theories.* Providence, RI: American Mathematical Society.