**Language, Meaning, and Information:**
**A Case Study on the Path from Philosophy to Science**
**Scott Soames**

Near the beginning of the final lecture of *The Philosophy of Logical Atomism*, in 1918, Bertrand Russell articulates a view of the relationship between philosophy and science for which there is much to be said.  He says:

> "I believe the only difference between science and philosophy is that science is what you more or less know and philosophy is what you do not know.  Philosophy is that part of science which at present people choose to have opinions about, but which they have no knowledge about. Therefore every advance in knowledge robs philosophy of some problems which formerly it had… [and] a number of problems which had belonged to philosophy will have ceased to belong to philosophy and will belong to science." (P. 154)

In short philosophy is the way we approach problems that are presently too elusive to be investigated scientifically. The goal is to frame questions, explore possible solutions, and forge conceptual tools needed to advance to a more definitive stage of investigation. My topic is the case study of this form of philosophical progress provided by the study of language, meaning, and information in the past 132 years.

The story begins with the development of modern symbolic logic instigated by Gottlob Frege and Bertrand Russell at the end of the 19th and beginning of the 20th centuries.  Initially, their goal was to answer two questions in the philosophy of mathematics: *What is the source of mathematical knowledge?* and *What are numbers?* They answered that *logic* is the source of mathematical knowledge, that

*zero* is the set of concepts true of nothing, that *one* is the set of concepts that are true something, and of only that thing, that *two* is the set of concepts true of some distinct x and y, and nothing else – and so on. Since the concept *being non-self-identical is true of nothing* it is a member of the number zero; since the concept *being this year's Selfridge lecturer* is true of me and only me, it is a member of the number one; since the concept *being my child* is true of Greg and Brian Soames, and only them, it is a member of the number two. Other integers follow in train. Since numbers are sets of concepts, the successor of a number n is the set of concepts F such that for some object x of which F is true, the concept *being an F which is not identical to x* is a member of the number n. With zero and successor under our belts, we define the class of natural numbers as the smallest class that contains zero, and that, whenever it contains something, always contains the successor of that thing.  With all this in place, multiplication is defined as repeated addition, addition is defined in terms of counting, and counting is seen as repeated moving from a number to its successor.  In this way all of arithmetic was derived from what Frege and Russell took to be pure logic. When, in similar fashion, classical results of higher mathematics were derived from arithmetic, it was thought that all classical mathematics could be so generated.  *In short, logic was seen the foundation of all mathematical knowledge*.

That was Frege and Russell's breathtaking dream. The reality was more complex. The first step in the project was the development of the modern *predicate calculus*, which combined traditional truth-functional logic – familiar from the Stoics onward – with a powerful new account of quantification (expressed by the words 'all' and 'some') supplanting the more limited syllogistic logic dating back to Aristotle. The key move was to trade-in the subject/predicate distinction of syllogistic logic for an expanded version of the function/argument distinction from mathematics. Applied to quantification, this meant treating the claims expressed by ⌜Something is F⌝ and ⌜Everything is G⌝ as predicating *being instantiated* of the concept expressed by F and *being universally instantiated* of the concept expressed by G. The result was an enormous increase in expressive power.

Although Frege's first-order version of this system was sound and complete – in the sense of proving all and only genuine logical truths – the concepts needed to precisely state and prove this meta-logical result were still 50 years away. In itself, this didn't interfere with the reduction of mathematics to logic. More serious was the fact that at this early stage modern logic was intertwined with what is now called "naïve set theory" – according to which for every well-formed condition on objects stateable in one's language there is a set – perhaps empty, perhaps not – of all and only the things satisfying that condition. To think of this as a principle of

logic is to think that *talk* about an individual's *being so-and-so* is interchangeable with talk about it being in the set, or class, of things that are so-and-so.

When Russell's paradox demonstrated the contradiction at the heart of naïve set theory, the consequences for the project of reducing mathematics to logic were profound. It immediately became clear that the principles required to generate sets without falling into contradiction are less obvious, and subject to more genuine doubt, than the arithmetical principles that Frege and Russell hoped to derive from them. This undercut the strong initial epistemological motivation for reducing mathematics to logic. Partly for this reason, the ultimate demarcation between logic and set theory that worked itself out over the next few decades was one in which set theory came to be viewed by most as itself a kind of elementary mathematical theory, rather than a part of logic. Although reductions of arithmetic and classical mathematics to set theory could still be done, with illuminating results for the foundations of mathematics, the philosophical payoff was not as great as Frege and Russell first imagined. The philosophy of mathematics had made real progress, but its fundamental problems had not been solved.

This *philosophical* shortcoming was compensated by the birth of new *scientific* investigations. Powerful systems of logic had been developed, together with new deductive disciplines – proof theory and model theory – to study them. A system of logic in the modern sense consists of a formally defined language, plus a

proof procedure, often in the form of set of axioms and rules of inference. A proof is a finite sequence of lines each of which is an axiom or a formula obtainable from earlier lines by inference rules. Whether or not something counts as a proof is decidable merely by inspecting the formula on each line, and determining whether it is an axiom, and, if it isn't, whether it bears the structural relation to earlier lines required by the rules. Since these are trivially decidable questions, it can always be decided whether something counts as a proof, thus forestalling the need to prove that something is a proof.

In a purely logical system, the aim is to prove all and only the *logical truths* as theorems, and to be able to derive from any statement all and only its *logical consequences*. *Logical truth* and *logical consequence* are defined semantically. To think of these as semantic notions is, of course, to think of them as having something to do with *meaning*. Although this wasn't exactly how the founder of model theory, Alfred Tarski, initially conceived them, it is how his work was interpreted by Rudolf Carnap and many who followed. The key idea, on this interpretation, is that *we can study the meaning of sentences by studying what would make them true.* This is done by constructing abstract models of the world and checking to see which sentences are true in which models. When a sentence is true in all models it is a logical truth; when the truth of one sentence in a model always guarantees the truth of another, the second is a *logical consequence* of the

first; when two sentences are always true together or false together they are *logically equivalent*, which is the logician's model for having the same meaning.

By the end of World War II the model and proof theories of the Frege-Russell predicate calculi were well understood, and philosophers were already building new projects on this foundation. The first new aim was to construct logical systems in which an operator – variously interpreted as *it is logically true that*, *it is analytically true that*, or *it is necessarily true that* – is added to the predicate calculus so that prefixing it to a standard logical truth produces a truth. Since the operator can be iterated, and so applied to its own output, the resulting logic is more complex than that. But apart from problems answering the question *What is the logical, semantic, or metaphysical notion to be captured?* the technical ideas were pretty clear. Since these modal operators express operations defined in terms of truth at *model-like* elements, *logical models* for modal languages must contain such elements, often dubbed *possible world-states*. These are the items – thought of as *ways the world could have been* – relative to which terms have referents and sentences are true or false.

Let S be a sentence that is interpreted by a model M. To say that *sentence S is true at world-state W of model M* is to claim that *what S says, when interpreted by M, would be true if the world were in state W*. In other words, S is true at W iff the way that S represents the world as being is a way the world would be, if the

world were in state W. With this understanding of what it is for S to be true *at a world-state*, one may (i) the claim ⌈It is necessary that S⌉ to be true iff S is true at every world-state, (ii) take the claim ⌈It could have been true that S⌉ to be true iff S is true at some world-state, and (iii) take the claim, ⌈If it had been the case that S, then it would have been the case that R⌉ to be true at W iff among the world-states at which S is true, some states at which R is true are more similar to W than are any world-states at which R is false.

As before, *logical truth* and *logical consequence* are defined in terms of truth in a model; this time the basic notion is truth at *a world* in a model. Logical equivalence is then sameness of truth value at all worlds of all models, which remains the logician's approximation of sameness of meaning. Proof and model theory for modal systems – to which Saul Kripke made famous contributions in the 1950s – are complicated versions of similar theories for propositional and predicate calculi. Although it remains controversial which axiomatic systems of modal logic are correct, this is due in part to the fact that there are different modal notions to be formalized.

This new logical project brought with it a new philosophical focus. Frege's and Russell's interest in logic had been rooted in its role in answering their questions about mathematics. Since there are no distinctions between necessary and

contingent truths of mathematics, the new interest in logics of necessity and possibility came from different sources. One source was the ubiquity of modal reasoning in philosophy itself, as well as its potential use in empirical science. Another, more important, source came from the Fregean conception of language that informed the development of logic from the start. On that conception, *for S to be meaningful is for it to represent the world as being a certain way, which is to impose conditions the world must satisfy, if it is to be the way S represents it*. Since these are the truth conditions of S, being meaningful involves having truth conditions. Hence, it was thought, the systematic study of meaning requires a framework for specifying the truth conditions of sentences on the basis of their grammatical structure and the representational contents of their parts. This had already been achieved, to a limited extent, by the model theory of the predicate calculus. With modal logic, the truth conditions of sentences provided by an intended model plus the theory of truth at a world of a model were, *for the first time*, strong enough to be realistic approximations of the *meanings* of sentences. To learn *what the world would have to be like* to conform to the way a sentence represents it to be *is* to learn something approximating its meaning.

The significance of this development for the study of language can hardly be overestimated. What started with Frege and Russell as the development of logical tools to help answer philosophical questions about mathematics had first

transformed into the development of set theory plus meta-theories of the predicate calculi. Next it expanded to incorporate richer logical systems, which in turn, provided the basis for a systematic study of linguistically encoded information, that could, in principle, be applied to all languages. Having reached this stage, we have both a putative answer to the question *What is the meaning of a sentence S?* and a systematic way of studying meaning. The putative answer is that the meaning of S is its truth conditions, which are modeled by the set of those states w such that if the world were in w, then it would be as S represents it to be. The systematic method for studying meaning is to derive these truth conditions compositionally, from the interpretations of words occurring in the sentences. With this, one has a theoretical framework that can be applied to languages in general.

This is where the philosophically inspired study of linguistically encoded information stood in 1960. Since then, philosophers, philosophical logicians and theoretical linguists have expanded the framework to cover large fragments of human languages. The research program – to which Saul Kripke, Richard Montague, David Kaplan, Robert Stalnaker, David Lewis, Hans Kamp, Barbara Partee, Angelica Kratzer, Irene Heim, and others have contributed – starts with the predicate calculi and is enriched piece by piece, as more constructions found in languages like English, Japanese, and Hindi are added.

I have mentioned modal operators and counterfactual conditionals (*if it had been the case that ___, then it would have been the case that ___*). Operators involving time and tense can be treated along similar lines. Generalized quantifiers have been added, as have adverbs of quantification, and propositional attitude verbs such as *believe*, *expect,* and *know*. There are also accounts of adverbial modifiers, intensional transitives, indexicals, and demonstratives. At each stage, a language fragment for which we already have a truth-theoretic semantics is expanded to include more features found in natural language. As the research program advances, and more such features are incorporated, the fragments of which we have a good truth-theoretic grasp become more powerful and more fully natural-language like. Extending results so far achieved, one can imagine a time at which vastly enriched descendants of the original logical languages of Frege and Russell approach, or even match, the expressive power of natural language – allowing us to understand the principles by which information is encoded.

This program is now the dominant semantic approach in theoretical and empirical linguistics. If all that remained were to fill in the gaps and flesh out the details, philosophers would have done most of what was needed to transform initial philosophical questions into more tractable scientific ones. However, we *haven't* reached that point. Rather, we fall short in several ways. One of these involves the relationship between the information semantically encoded by a sentence, and the

assertions the sentence is used to make or the beliefs it is used to express. Up to now in my talk I have oversimplified that relationship by tacitly assuming that, with some exceptions, the semantically encoded information is identical, or nearly identical, what is asserted and believed by an utterance of the sentence. Although that, indeed, had long been the standard assumption, in recent years that it has increasingly been called into question as more and more philosophers and linguists have come to think that that meaning, or semantic content, alone doesn't always determine what is asserted, but rather interacts, in ways we are just beginning to understand, with relevant contextual information to produce assertive content.

In addition, it has often been assumed that since meaning is semantically encoded information, understanding a sentence is knowing of the sentence that it carries that information. However, this may also too simple. It is plausible to suppose that to understand a word, phrase or sentence is to be able to use it in expected ways in communicative interactions with members of one's linguistic community, which may involve graded recognitional and inferential ability that goes beyond our knowledge of a certain content that it is the content of the word, phrase, or sentence. Finally, we have identified semantic content with information that represents the world as being a certain way, but we haven't yet given a plausible story about what a piece of information is. This, in my opinion, is our most urgent task.

The model of the information semantically encoded by a sentence I have so far assumed identifies it with the set of possible world-states in which the sentence is true. On this view, a sentence represents things as being one way or another, and so has truth conditions, *because* the information it encodes – i.e. the proposition it expresses – represents things in that way, and has truth conditions. This is what is identified with the set of world-states in which the sentence is true. But are propositions *really* sets of possible world-states – i.e. sets of maximal properties the world might have had? There are two powerful reasons to think they are not.

First, if they were then all necessary propositions would be identical with the set of all possible world-states, and so with one another. So there would be only one necessary truth. Moreover anyone who believed a proposition p would believe every necessary consequence of it (provided that we take it for granted that whenever one believes the proposition expressed by ⌈P&Q⌉, one believes the proposition expressed by P and the proposition expressed by Q), from which it follows that anyone believing a necessary falsehood, would thereby believe every proposition. Since no one could do that the view has the consequence than no one can believe the impossible. These results are obviously incorrect.

Second, sets of possible world-states aren't the right sorts of things to have truth conditions in the first place. Consider, for example, the set containing world-states 1, 2, and 3. Is it true or false? The question is bizarre no matter which three

states we choose. We could answer the question, if we could figure what things the set of world-states represents as being one way rather than another. But we can't do that, because the set doesn't represent anything at all. If we wanted, we could use it to represent *the actual state of the world as being in the set* – and so to make the claim that no maximal property of the world *outside* the set is instantiated. But we could equally well use the set to represent the actual state of the world as *not being in the set* – and so to make the claim that no maximal property *inside* the set is instantiated. Independent of interpretation by us, the set of world-states doesn't represent anything, doesn't make any claim, and so doesn't have truth conditions.

What about the *function* that assigns world-states 1-3 *truth* and all others falsity? This doesn't seem to be a piece of information either. Suppose we replace truth and falsity with USC and Lehigh. What does the function that assigns some world-states USC and others Lehigh represent? Without interpretation by us, it doesn't represent anything. Why, then, should the original function assigning truth and falsity be representational? *What, after all, are truth and falsity but properties we grasp primarily through their application to propositions? But, if propositions are needed to illuminate truth and falsity, then truth and falsity can't be presupposed as building blocks from which propositions are constructed.*

The illusion that propositions can be identified with functions from world-states to truth values may be fed by illicitly assuming an antecedent conception of

propositions. One who does this implicitly associates each assignment of a truth value to a world-state w with *the proposition that predicates that world-state of the universe*, and so is true iff w is the state the universe really is in. A function from world-states to truth values can then be associated with the disjunction of propositions correlated with its assignments of *truth* to such states. It could then be taken to say that the universe is either this way, that way, or the other – which would be fine, *if one already had independent accounts of propositions as things that predicate properties*, in this case *a world-state*, of other things, in this case *the universe*. But that is precisely what is lacking when one *identifies* propositions with functions from world-states to truth values. This shows that *possible world-state semantics* lacks an adequate account of propositions, and so is incomplete.

But, as I will explain in few moments, the real problem is far worse; in order to provide any genuine information about meaning at all possible world-state semantics must presuppose the very thing it lacks. The way out of this impasse is to provide a genuinely acceptable conception of what propositions really are. This is the next great stage of philosophical inquiry, if we are ever going to succeed in turning the study of language, meaning, and information into a scientific inquiry.

Here is a strategy for getting the job done. We start with the observation agents *represent objects as being a certain way* when they perceive them or think of them as being that way – e.g. when they perceive this briefcase, call it

'B', as being brown, or think of it as brown. . Next, we consider *what the agent does* – namely *represent B as brown*, which is to predicate the property *being brown* of B. This cognitive act can itself be said to *represent B as being brown* in essentially the same sense in which some acts are said to be intelligent, stupid, thoughtful, or kind. Just as, very roughly, for an act to be intelligent, stupid thoughtful, or kind is for it to be one the performance of which marks an agent as behaving intelligently, stupidly, thoughtfully, or kindly, so for a cognitive act to represent B as brown is for it to be one the performance of which marks an agent as representing B as brown. We, as agents, use this sense of representation to isolate individual aspects of the thought and perception, of ourselves and others, in order to assess them for accuracy. When o is such that to perceive or think of o as brown is to represent it accurately, it is enormously useful and very natural to identify an entity – a particular sort of perceiving or thinking – plus a property that entity has when this sort of perceiving or thinking is accurate. The entity is a proposition, which is the cognitive act of representing o as brown. The property is truth, which the act has iff to perform it is for an agent to represent o as o really is.

For an agent to entertain a proposition is, in the simplest case, for the agent to predicate a property of something – which is simply for the agent to perceive the object as having the property, or to think of it as having that property. For example

to *entertain* the proposition that B *is brown,* is to see it *as brown*, or to think or imagine it as *being brown*. In all of these cases, one predicates *brownness* of the object – either perceptually or cognitively. The proposition itself is *the minimal cognitive or perceptual act* in which one predicates *being brown* of this thing, and so represents it as brown. For this proposition to be true is just for this thing to be as the proposition represents it. Since the briefcase is in fact brown, the proposition is true.

In this simple example, we are introduced to what truth is, and what propositions are. Propositions are individual cognitive acts, or sequences of such, in which an agent represents things as being certain ways. A proposition is true when things are the way it represents them to be. We learn more about truth when we realize that p and the *proposition that p is true* are necessary and apriori consequences of one another, and that any warrant for asserting, denying, believing, or doubting one is warrant for taking the same attitude to the other. Propositions can also be expressed by sentences. When I say, "This is brown," gesturing at my briefcase, *my use of the sentence* is true because the proposition I use it to express is true. Thus we extend the notion of truth to sentences by virtue of the propositions they are used to express. Since my perception and thought also represent the briefcase as brown, they are accurate, or veridical.

There is much more to say about this conception of propositions in my contributions to *New Thinking about Propositions*, by Jeff King, Jeff Speaks, and me, forthcoming from Oxford University Press. Today, I will end by sketching how it provides what must be added to familiar versions of possible world-state semantic theories, in order to arrive at genuine theories of meaning. The basic idea is that individual words stand for objects, properties, and functions. Phrases are assigned various combinations of these, and sentences are associated with still larger combinations, from which we can read off the sequences of cognitive acts that are the propositions the sentences express. Given such a sequence, we can specify how it represents things as being. This is combined with the principle that *a proposition p is true at a world state w iff were w to be instantiated things would be as p represents them to be*. From this we generate the truth conditions of propositions, from which the truth conditions of sentences expressing them are inherited.

That, in a nutshell, is how real propositions can be added to possible world-state semantic theories. In fact, it is *only* by adding such propositions that possible world-state semantics can give us *any information* at all about the meanings of sentences. Since this point has generally *not* been noted, I will close by explaining it. As I emphasized earlier, possible world-state semantics is premised on the idea that the truth conditions of a sentence are intimately related to its meaning. This is

reflected in our apriori knowledge that *a sentence that means (or expresses the proposition) that so-and-so is true iff so-and-so*. This knowledge allows us to derive information about meaning from statements about truth conditions. For example, we can derive ⌈'S' doesn't mean that R⌉ from ⌈'S' is true iff Q⌉, when Q and R are known to differ in truth value. This result is strengthened when truth conditions are relativized to world-states. However, doing so requires propositions.

How do we extract information about meaning from the claim *the Spanish sentence 'El Presidente habla ingles' is true at world-state w iff at w the President speaks English*? The first step is to ask what it means to say that x *speaks English* at w. The answer, we are told, is that say this is to say that *if w were instantiated, then x would speak English*. We next ask, what is it to say that a sentence is true at w? We know that for a sentence to be true is for the proposition it expresses to be true, so we ask, *What is it to say that a proposition p is true at w?* The answer, we are assured, is that it is to say that *if w were instantiated, then p would be true*. For example, to say that the proposition that the President speaks English is true at w is to say that if w were instantiated then the proposition that President speaks English would be true, and hence, the President would speak English.

So far so good. However, at this point we face a complication. In possible world-state semantics we *can't* quite say that for a Spanish sentence S to be true at w is for it to be such that *if w were instantiated, then S would be a true sentence of*

*Spanish (i.e. one which expresses a true proposition).* The reason we can't say that is that in possible-world-state semantics, S can be *true at w* even if S means nothing at w, or means something different at w from what it actually means. This shows that the *dyadic truth predicate of possible world-state semantics* is a technical substitute for our ordinary notion of truth. Using our ordinary notion, we say that S is true at w iff *at w,* S expresses a proposition that is true. But, since what S could have meant (or expressed) is no help in illuminating what S actually means (expresses), the possible worlds-state semanticist needs some other story. Although the strict possible world-state semanticist is at a loss, those of us who have real propositions at our disposal are not. To gain information about meaning from a truth-conditional semantics of this sort, we must understand its claim *that S is true at w iff at w, the President speaks English* as stating that *the proposition that S is used by us to express at the actual world-state is true at w* – i.e. iff at w, the President speaks English.

In short, in order to provide genuine information about meaning, the dyadic truth predicate of possible-world-state semantics must be understood as parasitic on the prior notions: *the proposition actually expressed by a sentence* and the *ordinary, unrelativized property truth of propositions.* It is by taking these for granted that we extract useful information about meaning from the truth conditions provided by such a theory. When such a theory gives us ⌈∀w 'S' is true-at-w iff at w, Q⌉ we

derive ⌈'S' doesn't mean that R⌉ when Q and R aren't necessarily equivalent. We do this by tacitly assuming ⌈If 'S' means (or expresses the proposition) that P, then necessarily the proposition 'S' actually expresses is true iff P⌉. Although this doesn't identify what S does mean, it does so up to necessary equivalence, thereby providing information about meaning that restricts the range of acceptable alternatives. What has not formerly been appreciated, but what must now be, is that *without the prior notions of truth and propositions here employed, even this limited information about meaning extracted from the semantic theory would be lost.*

This is the heart of the problem with the analysis of propositions as sets of world-states, or functions from world-states to truth values. If world-states are taken to be unexplained primitives, with the goal of using them to provide reductive analyses of properties, propositions, and meaning, then the method just given for extracting claims about meaning from possible-world-state semantic theories is unavailable. *Without prior accounts of propositions, truth, and the connection between meaning and truth, the theorems of such a semantic theory won't carry any information about meaning.* Thus, the great progress in the scientific study of meaning in human language that has sprung from work in the possible-world-state framework has been premised on our ability to make good on a tacit promise. The promise is to identify what truth, meaning, and, above all,

propositions – or pieces of information – really are. This is the task that we are

now, after studying this question for more than a century and a quarter, finally in a

position to accomplish.