

# The Semantic View and the $\alpha$ -model

Richard Lawrence  
rwl@sas.upenn.edu  
235 S. 42nd St., Apt. 2B  
Philadelphia, PA 19104

April 9, 2007

## **Abstract**

In what follows, I discuss the semantic view of theories, specifically in the ‘state space’ formulation that has been developed by Bas van Fraassen, Frederick Suppe, and others and applied by Elisabeth Lloyd. I consider the claims that the state space view makes about how scientific theories are best understood. I then discuss a particular model from network theory, the  $\alpha$ -model developed by Duncan Watts, and try to apply the state space view to the  $\alpha$ -model theory. I argue that the way Watts uses the  $\alpha$ -model is not described very easily under the state space approach, because its parameters are idealized in a way that the state space approach does not account for, and because the relation between the  $\alpha$ -model and the world is not one of statistical fit between its models and data.

# 1 Introduction

Recent debate in philosophy of science has recognized that an important kind of scientific explanation is based on *models*. Just what a model is, though, and how it can describe or explain anything, has not been settled. The major advocates for understanding how models function in scientific theories have been proponents of the *semantic view of theories*, a philosophical position that holds that theories are to be understood directly in terms of the models they present, instead of in terms of their particular linguistic formulations.

The semantic view, as developed by Patrick Suppes, Bas van Fraassen, Frederick Suppe, and others, has evolved into two different approaches to analyzing scientific theories. The first, due to largely to Suppes' work, is a set-theoretic approach, where models are viewed as "possible realizations" of theories and data. In Suppes' view, theories and actual data are related via a hierarchy of such models. [5, 6] The second approach, developed by van Fraassen, Suppe, and others, views theories as structures that specify certain parameters for describing phenomena. These parameters are arranged as the axes of a "state space" or "phase space." A model, on this view, is a particular assignment of values to the parameters of the theory; it is therefore seen as a point in or "trajectory through" the state space. Observation data is plotted in the state space as well, so that the basis for comparing a theory to the world is a mathematical relation between data points and model curves in the space.

The two approaches share what I will call the *basic semantic view*. The basic semantic view holds that scientific theories are best understood as extralinguistic entities that may be variously described in language. Thus, a theory is a structure or set of structures which satisfies (in the logician's sense) its linguistic formulations. To give a theory directly, one must give these structures. So, as van Fraassen says, to present a theory in science is

to specify a family of structures, its *models*; and secondly, to specify certain parts of these models... as candidates for the direct representation of observable phenomena. [7, p. 64]

This basic tenet is what makes the view a *semantic* one: by holding that theories are something other than their linguistic formulations, proponents of the semantic view hope to escape the complicated discussions of syntax that plagued the logical positivists before them.

In what follows, I will be concerned with the additional claims beyond the basic semantic view that are held by advocates of the state space approach. I will then discuss a particular example from network theory, the  $\alpha$ -model of Duncan Watts, and the "small world problem" it was designed to help elucidate. The  $\alpha$ -model, I contend, is a challenging case for the state space approach, because it seems to be idealized in a way other than state space advocates think it should be, and because the way Watts compares it to the world is not through statistical fit of data.

## 2 The Semantic View of Theories

Under the semantic view, a theory is to be thought of as independent of how it is specified in language. We identify a theory by directly picking out the class of models or structures that satisfy its axioms. The notion that a model ‘satisfies’ a linguistic formulation of a theory is borrowed from logic: it is the notion that structures *make true* or *realize* the meanings of linguistic sentences and symbols. For example, the sentence “This structure is a dense linear ordering” is satisfied by  $\mathbb{R}$ , the structure of the real numbers, but not by  $\mathbb{N}$ , the structure of the natural numbers. If  $T$  is a theory of dense linear orderings, then  $\mathbb{R}$  is a model of  $T$  and  $\mathbb{N}$  is not.

Proponents of the semantic view extend this notion of “satisfaction” or “being a model of” beyond purely formal and mathematical languages to the language of science. For example, both a universe in which the center of gravity has a velocity of zero and a universe in which the center of gravity has a velocity  $v > 0$  (all other things being equal) are models of classical mechanics, since Newton’s axioms are satisfied in both. The domain or “intended scope” of the theory of classical mechanics is *everything that has mass*; so any physical system in which the laws of classical mechanics hold is a model satisfying the theory, whether it is as large as the galaxy or as small as an apple falling to the ground. Other scientific theories have different domains, and therefore different models. In biology, for example, populations of a predator and its prey whose densities fluctuate according to Volterra’s equations are a structure satisfying Volterra’s theory.

It’s important to notice that we are still employing the term “model” here in the logician’s sense, wherein a *structure* is a model of a *theory*. Models are concrete and particular, while theories, as linguistic entities, are abstract and may be interpreted differently in different domains. It’s not important that the objects or relations in a model actually exist; it’s only important that, if they do exist, then they satisfy the theory asserted about them. There may not be any real predator-prey populations or physical systems satisfying Volterra’s or Newton’s equations exactly, but this does not prevent us from describing a variety of situations in which either theory *would* be true. This is a very different use of the word “model” than is typically found in science. The scientific meaning of “model” is almost the opposite of its meaning in logic: biologists would say that Volterra’s equations are a mathematical model of real fluctuations in the population densities of predators and prey under certain conditions. In order not to confuse the two, some semantic view theorists have introduced different terminology for the scientific notion of “model.” Van Fraassen, for example, calls the scientific notion a “model-type.” He explains:

In the scientists’ use, ‘model’ denotes what I would call a model-type. Whenever certain parameters are left unspecified in the description of a structure, it would be more accurate to say... that we described a structure-type. Nevertheless, the usages of ‘model’ in meta-mathematics and in the sciences are not as far apart as has

sometimes been said. I will continue to use the word ‘model’ to refer to specific structures, in which all relevant parameters have specific values. [7, p. 44]

The idea here is that scientists often think of models as mathematical structures that capture what is common to many particular systems, while logicians think of those particular systems as models. So, while a physicist will say that

$$x = A \sin(\omega t) + B \cos(\omega t) \tag{1}$$

models the position  $x$  of a harmonic oscillator over time, a logician will think that the values of  $A$ ,  $B$ , and  $\omega$  must be specified before the equation actually represents a model of the theory of harmonic oscillators. Part of the reason for making this distinction is that proponents of the semantic view like van Fraassen believe that what makes a model *scientific*, and not simply mathematical, is that it (or some part of it) is capable of being empirically confirmed. Since (1) can only be confirmed after  $A$ ,  $B$ , and  $\omega$  have been given, it’s this specified version that is the true model of the theory.

So far, I have been characterizing what I earlier called the “basic semantic view,” the view that theories are to be understood as collections of models. Following van Fraassen, we can summarize the basic semantic view as making two claims about how a scientific theory  $T$  should be understood:

**Semantic claim:**  $T$  presents a family of structures, its models, which satisfy the relations asserted by  $T$  in a certain domain. Understanding  $T$  involves understanding these models, independently of how they are described in language.

**Empirical claim:**  $T$  specifies properties of its models as candidates for the direct representation of observable phenomena.

What does the state space view add to these claims? To answer this question, we should first give an account of what the state space view of theories looks like. The best summary I have seen is given by Frederick Suppe in *The Structure of Scientific Theories*. [4] He explains the view this way. First, it’s important to recognize that any scientific theory has a domain of phenomena known as its “intended scope”: that part of the world which the theory intends to describe and explain. It would be extremely difficult, and not very useful, for a theory to try to describe each of the phenomena in its intended scope separately; rather, a scientific theory “abstracts certain parameters from the phenomena and attempts to describe [all] the phenomena in terms of just these abstracted parameters.” [4, p. 223] This is what is typically meant, I take it, when philosophers of science say that theories ‘organize’ phenomena: they mean that theories describe many phenomena in some domain in terms of just a few concepts or parameters, which gives us a mechanism for saying how distinct phenomena are related (the same concepts apply to both) and how they are different (different concepts apply to them, or different values of the same parameters). So far, then, Suppe is presenting a fairly standard view.

What he says next, however, is slightly more controversial:

In effect the theory assumes that only the selected parameters exert an influence on the phenomena and thus that these parameters are uninfluenced by any other parameters in the phenomena. As such the theory assumes that the phenomena are *isolated* systems under the influence of just the selected parameters. [4, p. 223]

I suspect that this assertion might draw criticism from scientists, particularly those working in fields where the phenomena are explicitly recognized as being influenced by more parameters than the theory uses or can discover. This is true in ecology, for example, where the natural systems under study are *necessarily* unisolated from the environments in which they occur. Still, it seems reasonable to assume that, for pragmatic reasons, scientists often make some kind of simplifying assumptions when constructing a theory. These might come in the form of an assumed-isolated system, or they might be something weaker, like an “other things being equal” clause, or the agreement to lump unknown parameters into an error-term in their equations. At any rate, it’s probably true that no theory is perfect, and that one reason many theories are imperfect is that enumerating and understanding all the parameters influencing their intended domains is a long and difficult task.

The conclusion that the state space view draws from the fact that theories use an incomplete set of parameters to describe the phenomena in their intended domains is that the systems described by theories are *idealizations*. Suppe and others (somewhat confusingly) refer to these idealized models as “physical systems.” I will call them “physical models” in order to avoid confusing them with *actual* physical systems, and to highlight the fact that they are just the structures that the semantic view thinks are so important to understand. An ideal harmonic oscillator, whose position is exactly described by equation (1) once the parameters have been specified, is an example of the physical models that Suppe has in mind. The physical models of a theory are described completely in terms of the theory’s selected parameters for describing phenomena in its domain. Because these parameters are never enough to completely characterize actual, unisolated phenomena in the theory’s domain, the physical models are idealizations of those phenomena; they are, on Suppe’s view, what the phenomena *would have been* if those phenomena were free of the influence of outside parameters.<sup>1</sup> [4, p. 224]

Once scientists have enumerated the parameters of a theory and constructed the physical model-types they believe to characterize what the phenomena would be in isolation from all other parameters, the state space approach is ready to begin its analysis of the theory. A *state* is a specification of a value for each of

---

<sup>1</sup>The notion of ‘idealization’ has not received as much attention in the literature as perhaps it should. It seems likely that systems can be idealized during the construction of models in a variety of different ways, depending on the intentions of the scientist building the model. To admit the influence of unknown parameters is one kind of idealization, for example, but the exclusion of mathematical terms containing only known parameters for the purpose of mathematical simplicity is another. Since it is my goal here to present the state space approach to scientific theories, and not to provide a taxonomy of the kinds of idealization, I have chosen to simply present Suppe’s view and leave the notion otherwise unanalyzed.

the theory’s parameters.<sup>2</sup> The states allowed by the theory are just those that fit into its physical model-types, i.e., its mathematical laws. The theory of ideal gases, for example, has the model-type:

$$PV = nRT \tag{2}$$

Thus, states in this theory are ordered tuples of the form  $\langle g_P, g_V, g_n, g_T \rangle$ : an ideal gas  $g$ ’s pressure, volume, number of moles, and absolute temperature. ( $R$  is a constant.) Tuples not satisfying equation (2)—for example, those in which  $g_P g_V > g_n R g_T$ —are not allowed states in the theory.

If a theory has  $n$  parameters, we can construct an  $n$ -dimensional *state space* consisting of all the possible (both allowable and disallowed) tuples under some ordering. That is, we let the dimensions of the state space be the parameters of the theory, and let the points in the space be the possible states arranged along each dimension according to the ordering. In the four-dimensional ideal gas space, for example, the point  $\langle 2_P, 2_V, 1_n, 300_T \rangle$  is ‘further out’ along each dimension than the point  $\langle 1_P, 1_V, 0.5_n, 150_T \rangle$ . (Depending on the units used, of course, neither of these points is necessarily an allowed state in the theory of ideal gases.) Most often, the ordering of each dimension is the usual ordering of the real numbers, but it need not be; so long as the possible values for each parameter can be ordered in *some* way, they can be arranged in a state space.

Since the laws of a theory impose constraints on which states are allowable among all the possible states, and the physical models of the theory are those structures that satisfy its laws, the physical models cover all of the allowed states in the state space; the other points are not allowed by the theory, so the theory has no models to occupy them. Quite often (though not necessarily), the physical models of a theory are dynamic, and have a value for each parameter through some length of time. We thus envision physical models as *trajectories through the state space*, that is, as curves parameterized by time drawn in  $n$  dimensions. If the models are static, every model is identified with exactly one point in the space.

We can summarize the discussion of the state space view so far with the following additional claims about a theory  $T$ :

**Idealization claim:**  $T$  describes phenomena in its intended domain in terms of an incomplete set of parameters, but its models are described completely in terms of those parameters; therefore, the models are idealized representations of actual phenomena.

**State space claim:**  $T$  can be represented as a set of allowed states in the  $n$ -dimensional state space  $S(T)$  whose dimensions are the parameters of the theory. The models of  $T$  cover the allowed states in  $S(T)$ , and none of

---

<sup>2</sup>I should note here that the terms “state” and “parameter” are technical terms employed by semantic view theorists in just the way I have defined them above. Scientists may have different notions of what these terms mean, and though the different uses are supposed to be related, any departure here from those meanings should be taken as a choice made by semantic view theorists, not as a misunderstanding of what the terms mean in, e.g., physics.

the disallowed states. If the models are dynamic, they can be seen as time-parameterized curves through the state space.

There is one final issue for which I would like to summarize the state space view's perspective: the question of how theories and the world are to be compared. Recall that the state space approach, and the semantic view in general, are motivated by a desire to overcome the difficulties associated with syntactic approaches to the analysis of scientific theories. On the standard positivist account, for example, theoretical claims are reducible to claims in a 'pure observation language' through the application of the theory's internal logical calculus and the theory's 'bridge laws,' which define basic theoretical concepts in terms of purely observational concepts. The canonical example is that of claims in psychology about pain, which is a 'theoretical concept' because it allegedly cannot be observed directly. Rather, we infer that a psychological subject is in pain because, for example, we observe it screaming, writhing in a way that suggests agony, trying to escape the stimulus, etc. Hence, we can in principle reduce claims about pain to claims about the observation concepts 'screaming,' 'writhing,' 'escaping,' and so forth via a bridge law that states " $x$  is in pain if and only if  $x$  is screaming, or  $x$  is writhing, or  $x$  is trying to escape. . ."

The problems with this view are well-catalogued, and I do not intend to reiterate them here. For my purposes, it is enough to note that on the positivist approach, the way that theoretical claims reduced to observation concepts are compared to the world must essentially be taken as primitive: after a sufficient reduction takes place, the observation concepts in a claim are supposed to be simple enough that we can simply *tell* whether or not they obtain by observing the world. The whole positivist project rests on the idea that scientific theories, claims, and concepts are meaningful because they are *verifiable*, but no one has been able to satisfactorily explain what 'verification' is, and how to resolve disputes about whether or not observations verify scientific claims. There is thus a gap in the positivists' account of the relation between theories and the world, and it is a gap that the semantic view is interested in closing.

Under the state space approach to the semantic view, the basis for comparing theories to the world is some mathematical relation between the theory's models and observation data in the state space. The idea is that, given the parameters and models of a theory, we can compare the theory to the world by making observations of the natural systems the models are supposed to describe, and seeing how well the models fit the data. 'Fit' is the mathematical relation in the state space that closes the gap between theories and the world. There are different ways to define the mathematical fit of data to a model curve that are appropriate in different situations. Thus, fit is a flexible enough notion to capture the different types and strengths of verification. We can say that data fit a model more or less well; we don't have to make a binary choice as to whether or not they verify a scientific theory.

Van Fraassen and Lloyd both use the term "isomorphism" to describe the mathematical relation of fit. Lloyd describes the model-world relation this way:

Empirical claims are made about relations between models and nat-

ural systems; a natural system is described by a model when the model is isomorphic in certain respects to the natural system. [2, p. 72]

Similarly, van Fraassen says that for a theory to be empirically adequate<sup>3</sup> is for it to have some model such that “all actual appearances are identifiable with (isomorphic to)” a part of that model, where “isomorphism is of course total identity of structure.” [7, pp. 43–45]

In a strict sense, an isomorphism is a bijective correspondence between two sets that preserves relations on those sets. We might therefore interpret the state space view as holding that actual data points must be in bijective correspondence with a set of points belonging to one model of a theory for the theory to have adequately described the phenomena. Moreover, if there is any relation between the data points, such as succession in time, those relations must be preserved when the points are mapped onto the appropriate model: earlier data samples must map onto earlier points on the model-curve in the state space, and so forth.

This would be a very strong claim for the state space view to make—so strong, in fact, that it would probably be untenable. The relation of isomorphism, taken in this strict sense, is both too strong and too weak to be a good candidate for a relation between models and data. It’s too strong because it would be too difficult to obtain, most of the time: it’s extremely unlikely in most situations that a reasonable number of observations will be in perfect correspondence with a set of points on one of the model-curves in the state space. Each additional observation increases the chance that some data point’s relations to the other points will not be preserved by any map onto a curve in the space. What’s more, this interpretation of “isomorphism” ignores one of the main tenets of the state space view: that models are idealizations, so they won’t fit any data set *perfectly*. It’s for this reason that we brought in the notion of fit in the first place.

On the other hand, the strict interpretation of isomorphism is too weak because isomorphism is a symmetric relation: if  $A$  is isomorphic to  $B$ , then  $B$  is isomorphic to  $A$ . This is a problem because we think the relation of ‘modeling’ is *asymmetric*: if a structure  $A$  is a model of phenomenon  $B$ , it’s not also supposed to be the case that “ $B$  models  $A$ ,” at least not in the usual sense. Moreover, two models of a theory can be isomorphic to each other without our being able to see that they are models of any real phenomena at all, just as two pictures of the same object can be isomorphic without either being a representation of the other. (Rather, they are both representations of the object itself.) Hence, strict isomorphism is too weak a notion to pick out the relation between models and

---

<sup>3</sup>Empirical adequacy is van Fraassen’s anti-realist notion of what science aims for its theories to be. It is weaker than the realist notion of correspondence truth, but stronger than the standard notion of being ‘well-confirmed’ by past observations. Since van Fraassen claims that the semantic view is neutral on the question of realism, I take it that the notion of isomorphism he gives here is meant to apply to more than just empirical adequacy; it is meant to apply to all sorts of formulations of why we believe a theory is a good one. See [7] for more.



data.<sup>4</sup>

It's possible that the state space view has the resources to deal with one or both of these objections, but these objections are really directed at a straw man. I think both van Fraassen and Lloyd mean "isomorphism" in a weaker sense than the strict one I sketched out above, despite the bold claims with which they introduce the notion. In other parts of their work, both seem content to leave the business of relating models and data to traditional statistical methods. Van Fraassen, for example, says that "the measurement of how well a probabilistic model fits the data gathered... is a subject already of extensive study in statistics" [7, p. 194]; presumably, this is true of deterministic models as well. Lloyd similarly asserts that "fit can be evaluated by determining the fit of one curve (the model trajectory or coexistence conditions) to another (taken from the natural system); ordinary statistical techniques of evaluating curve-fitting are used."<sup>5</sup> [2, p. 147] All of this seems more in line with the state space view's general slogan that "philosophy of science should use mathematics, not meta-mathematics" than the view that the model-data relation is to be borrowed from logic.

I think the most judicious interpretation of the state space approach, then, is that the notion of 'fit' between models and data in the state space is basically the same as that of the statistical fit of a collection of data to a curve. This is already a broad and flexible notion; there are many different types of statistical fit, and different tests for determining the degree of a particular kind of fit. Moreover, it is a notion that is easily extended: if in some circumstances statistical fit does not apply or is inadequate, different types of fit that are similar in spirit can be defined. For example, when the models of a theory are not dynamic and hence don't correspond to curved trajectories in the state space but only to points, we might think of "goodness of fit" between a collection of a collection of data-points and a model-point as being inversely related to the volume of the  $n$ -dimensional solid required to enclose both the model-point and the data-points in the state space

The final claim of the state space approach, then, can be summarized as follows:

**Model-data relation claim:** The basis for comparing a theory  $T$  to the world is a mathematical relation between the models of  $T$  and the data from observations of phenomena in  $T$ 's domain. This mathematical relation consists in the statistical fit of the data to  $T$ 's models, thought of as curves in its state space, or in some similar relation appropriate to  $T$ .

---

<sup>4</sup>This objection is due to Roman Frigg. For more on this topic, see [1].

<sup>5</sup>Lloyd also gestures toward Suppes' set-theoretic approach as a way of relating models and data. It may be, she thinks, that a series of different "fits" are required to compare the models of a theory to data—for example, when there is no observable analog in the data for a concept used to define the theoretical models [2, p. 146]. Since I am not dealing with Suppes' approach here, I leave the exploration of this idea to the reader, but it is important to note that the two approaches to the semantic view are not entirely divergent, and that one may make up for the other's faults.

As we will see, this claim and the others will figure into the state space view’s account of the theory that I now turn to: the Watts  $\alpha$ -model.

### 3 The $\alpha$ -Model and Network Theory

I shall now discuss a theory that I believe demonstrates the usefulness and appeal of the basic semantic view, but isn’t satisfactorily described by the state space view. This theory is known as the “ $\alpha$ -model,” and was presented by Duncan Watts in *Small Worlds: The Dynamics of Networks Between Order and Randomness*. [8] The  $\alpha$ -model is an algorithm for generating undirected graphs with properties similar to those of real social networks. It was proposed as an initial (hence “ $\alpha$ ”) approach to answering questions about what Watts calls the “small world phenomenon.”<sup>6</sup>

The small world phenomenon is the generalized version of the colloquial observation that “everyone is connected within six degrees of separation.” The idea is that every person belongs to a social network, the network of acquaintanceships. These networks overlap, forming a global web in which it is possible to choose any two people at random and find a series of acquaintances that connect them. If we choose persons  $A$  and  $D$ , for example, and  $A$  knows  $B$ ,  $B$  knows  $C$ , and  $C$  knows  $D$ , then  $A$  and  $D$  are somehow “connected,” even though they may not know each other, or even be aware of each other’s existence. The “six degree of separation” claim is that any two people, chosen at random in the global acquaintanceship network, can be connected in this way with six or fewer intermediate acquaintances in the chain. In this sense, the world of human social relations is small, even though it has billions of members.

This claim about six degrees of separation, if it is true, is a surprising result to most people. The reason is that acquaintanceships are highly *localized* in that most of any one person’s acquaintances are confined to a small geographic area, and in that many of those acquaintances are also acquaintances of each other. Given that every person can only know a relatively tiny fraction of all the people in the world, and that many of the links from one’s acquaintances won’t reach outside one’s own social circle, the idea that there are fewer than six links required to reach *anyone* else from a given starting point seems unlikely.

There are also many other types of networks in the world, of course: computer networks, neural networks, the network of genetic lineage. The question arises as to whether any of these might also be “small worlds,” in the sense just described: whether, though they have large numbers of members, any member can be reached from any other in a small number of steps along the network’s connections. The questions of whether small worlds exist, under what condi-

---

<sup>6</sup>Though Watts thinks of the  $\alpha$ -model as a mathematical model, I will throughout the rest of this paper think of and refer to it as a “theory.” The reason for this is that  $\alpha$ -model consists of a set of rules for picking out what advocates of the semantic view would call its models: a class of undirected graphs. Though I will not use the term “ $\alpha$ -theory” in order to remain consistent with Watts, it would be more appropriate to think of the graphs it generates, and not the  $\alpha$ -model itself, as the models of the theory.

tions they arise, and how to describe their properties, are the focus of Watts' book. He formulates the problem this way:

Assuming that a network can be represented by nothing more than the connections existing between its members and treating all such connections as equal and symmetric, a broad class of networks can be defined, ranging from highly ordered to highly random. The question then is *Does the Small-World Phenomenon arise at some point in the transition from order to disorder, and if so, what is responsible for it?* [8, p. 24]

Watts is particularly interested in the “transition from order to disorder” because it seems, *prima facie*, that the topologies of many real, natural networks like the acquaintanceship network lie somewhere between those of highly-structured artificial networks and those of random networks, because connections are neither centrally planned nor formed totally independently of those already in place. The  $\alpha$ -model was designed to take this feature of natural networks into account, and to investigate the small world phenomenon in a rigorous and mathematical way. Before I can describe the details of the model, however, a short exposition of concepts from graph theory is in order.

### 3.1 A network theory primer

The most fundamental concept required for understanding the  $\alpha$ -model is the concept of an *undirected graph*. An undirected graph is a mathematical structure that consists of a set of points (“vertices”) and a two-place symmetric *edge relation* describing the connections (“edges”) between those points. Formally, we denote a graph by  $G(E, V)$ , where  $V$  is the set  $\{v_i\}$  of vertices in  $G$ , and the edge relation  $E$  is the set of pairs  $\{\langle v_i, v_j \rangle : v_i \text{ is connected to } v_j\}$ . This is the sort of structure that Watts has in mind when he says that a network can be represented by “nothing more than the connections existing between its members.” To represent a network with an undirected graph, you need only know which vertices are connected to which others; you don't need to know anything about what those connections are like, what sort of objects the vertices are, or any other information that might be relevant to describing a network in other contexts.

Given an undirected graph  $G$ , a *path* between two points  $v_1$  and  $v_2$  in  $G$  is a set of edges that connect vertices between  $v_1$  and  $v_2$ ; it's a set of edges you could “walk along” to arrive at  $v_2$  from  $v_1$  if you could only step along the edges of the graph. For any two vertices in  $G$ , there is a *shortest path* of  $l$  edges between them. For example, if  $v_1$  is connected to  $v_2$ ,  $v_2$  is connected to  $v_3$ ,  $v_3$  is connected to  $v_4$ , and  $v_4$  is connected to  $v_5$ , there is a path of length  $l = 4$  between  $v_1$  and  $v_5$ , so the shortest path between  $v_1$  and  $v_5$  is not longer than 4 edges. The path from a point to itself has length 0. If no path exists between  $v_1$  and  $v_2$ , we say the path between them has *infinite length*. If there are any paths of infinite length in  $G$ , it is *disconnected*; if all paths are finite, then  $G$  is *connected*,

meaning that any point can be reached from any other by walking along the edges of the graph. (Note that we are assuming here that the set of vertices  $V$  is finite. This is a reasonable assumption for graphs intended to represent real networks, but it is by no means mathematically necessary. In infinite graphs, however, the correlation between finite path lengths and connectivity breaks down.)

Given these definitions, we can define two macro-level properties of finite undirected graphs that are of special interest in the  $\alpha$ -model. The first of these is the *characteristic path length*: the characteristic path length  $L(G)$  of a graph is the median of the means of the shortest path lengths from each vertex  $v_i$  to all the others. That is, for each vertex  $v_i$  we calculate the length of the shortest path  $d(v_i, v_j)$  between  $v_i$  and every other vertex  $v_j$ . We calculate the mean length of these paths  $D_{v_i}$ , which in a graph of  $n$  vertices is:

$$D_{v_i} = \frac{\sum_{i \neq j} d(v_i, v_j)}{n - 1}$$

Then we calculate the median of these  $D_{v_i}$  to obtain the characteristic path length  $L(G)$ :

$$L(G) = \text{median}(\{D_{v_i} : v_i \in V\}) \quad (3)$$

Intuitively, the characteristic path length is the average number of steps required to walk along the edges of the graph from a typical vertex to another vertex chosen at random; it measures how ‘close’ the vertices of the graph are in the absence of an underlying spatial metric.

The second macro property of graphs that we are interested in with respect to the  $\alpha$ -model is the *clustering coefficient*  $\gamma(G)$  of a graph. To define this property, we first define the *neighborhood* of a vertex  $v$  to be the set of points  $\Gamma(v)$  in  $G$  that are connected to  $v$ :  $\Gamma(v) = \{v_j \in V : \langle v, v_j \rangle \in E\}$ . Note that by convention,  $v \notin \Gamma(v)$ . The *degree*  $k_v$  of a vertex  $v$  is the number of neighbors it has:  $k_v = |\Gamma(v)|$ . The *clustering around a vertex*  $\gamma_v$  is the total number of edges between  $v$ ’s neighbors divided by the total number of edges that are possible amongst those neighbors:

$$\gamma_v = \frac{\text{total number of edges in } \Gamma(v)}{\text{total possible edges in } \Gamma(v)} = \frac{|\{\langle v_i, v_j \rangle : v_i, v_j \in \Gamma(v)\}|}{\binom{k_v}{2}}$$

The idea is that the clustering around a vertex  $v$  captures the notion of how closely-knit  $v$ ’s ‘social circle’ is, if we imagine that the edge relation in  $G$  represents friendship between two vertices. If  $\gamma_v = 1$ , then all of  $v$ ’s neighbors are friends with each other in addition to being friends with  $v$ . If  $\gamma_v = 0$ , on the other hand, then none of  $v$ ’s neighbors are friends with each other; in the absence of  $v$ , the paths between them have lengths of at least 2.

Given this notion of clustering around a vertex, we can finally define the clustering coefficient  $\gamma(G)$  of the graph as the average over the clustering around all vertices. If  $G$  has  $n$  vertices, then

$$\gamma(G) = \frac{\sum_{v \in V} \gamma_v}{n} \quad (4)$$

The clustering coefficient of the graph captures for the whole graph what the clustering around a vertex captures on a local level: how ‘tightly-knit’ or locally-clustered the graph is on average. When  $\gamma(G) = 1$ ,  $G$  consists entirely of one or more *completely connected components*: a series of subgraphs in which every vertex is connected to every other, but no vertex is connected to any others in other subgraphs. When  $\gamma(G) = 0$ , by contrast, no neighbor of a vertex  $v$  is connected to any other vertex in  $\Gamma(v)$ . Such a graph is usually very *sparse*; that is, the total number of edges in the graph is much smaller than the total number of edges that are possible:  $|E| \ll \binom{n}{2}$ . It is possible, however, for a relatively sparse graph to still have a high clustering coefficient: if  $G$  consists of many completely connected but disjoint subgraphs, for example, then  $\gamma(G)$  will be equal to 1, but  $G$  will be sparse because each vertex is connected to only a small fraction of the total number of vertices it *could* be connected to.

When a graph  $G$  has a high clustering coefficient and a low characteristic path length simultaneously, and it is also relatively sparse,  $G$  is intuitively a ‘small world’: though connections in the graph are highly localized because  $\gamma(G)$  is high, the ‘degree of separation’ between any two vertices is low because  $L(G)$  is low.  $G$  should also be sparse, because it’s trivial that very dense graphs can have high  $\gamma(G)$  and low  $L(G)$ : it’s no surprise that the world is small if everyone knows almost everyone else. It is therefore the questions of when and how these properties co-occur that the Watts  $\alpha$ -model attempts to answer.

### 3.2 The $\alpha$ -model and its parameters

How can networks that are highly clustered, as well as relatively sparse, exhibit short characteristic path lengths? For Watts, the question is motivated by what we know about random graphs and about real social networks. Random graphs (that is, undirected graphs of the type described above, generated by forming edges on a set of  $n$  vertices by choosing from all the possible edges  $\langle v_i, v_j \rangle$  uniformly at random) have a low characteristic path length, but seem to be poor models of real social structure, since, in particular, it’s no more likely for two vertices in a random graph to be connected to each other if they have a common neighbor than if they do not. In reality, by contrast, many of my friends know each other, as do many of theirs, and so on: the clustering coefficient of the acquaintanceship network should be much higher than the typical clustering coefficient of a random graph. On the other hand, some empirical work<sup>7</sup> suggests that the characteristic path length of the acquaintanceship graph is relatively short. Thus, it seems that real social networks like the acquaintance graph have properties in common with both random graphs and highly-clustered graphs with lots of local structure, but are identical with neither.

The Watts  $\alpha$ -model is an algorithm for generating graphs that can have both a high clustering coefficient and a low characteristic path length. It builds a graph between a chosen number of vertices  $n$  in a stepwise fashion, adding a

---

<sup>7</sup>Watts credits Stanley Milgram with beginning an empirical investigation of the small-world phenomenon in the 1960s. [8, p. 18] See, for example, [3].

single edge to the graph at each iteration of the algorithm, and terminating after there are enough edges in the graph that its average degree  $k$  reaches some pre-defined value. The resulting graphs have values of  $\gamma$  and  $L$  that are determined by the parameters of the algorithm.

The algorithm proceeds by first visiting each vertex  $i$  in the graph in turn and calculating its “propensity”  $R_{i,j}$  to connect to every other vertex  $j$  in the following way:

$$R_{i,j} = \begin{cases} 1 & \text{if } m_{i,j} \geq k \\ \left[\frac{m_{i,j}}{k}\right]^\alpha & \text{if } k > m_{i,j} > 0 \\ p & \text{if } m_{i,j} = 0 \end{cases} \quad (5)$$

The parameters  $k, p$  and  $\alpha$  and the variable  $m_{i,j}$  in this equation have the following meanings:

- $k$  is the average degree of the graph (that is, the mean of  $k_v$  over all vertices  $v$ ), which, like  $n$ , is specified in advance
- $m_{i,j}$  is the number of neighbors that  $i$  and  $j$  already share on the current iteration of the algorithm
- $p$  is a baseline random probability that  $i$  and  $j$  will connect, even if they have no neighbors in common;  $p \ll \binom{n}{2}^{-1}$
- $\alpha$  is a “tunable parameter” defined on  $[0, \infty)$

Intuitively,  $R_{i,j}$  formalizes the notion that connections form in social networks between two vertices  $i$  and  $j$  with different probabilities, depending on how many acquaintances  $i$  and  $j$  already share.  $i$  and  $j$  have some random, baseline probability  $p$  of meeting, even if they don’t have any common friends. At the other extreme, if  $i$  and  $j$  share more acquaintances than most people know, they have too many common friends *not* to be connected to each other. How connections form between vertices between these two cases is controlled by  $\alpha$ .

Once all the  $R_{i,j}$  have been calculated, they are normalized to the unit interval: each vertex  $j$  is assigned some half-open interval on  $[0, 1)$ , the width of which is proportional to the fraction with which  $i$  has a propensity to connect to  $j$  among all other vertices:

$$\text{width of } j\text{'s interval} \propto \frac{R_{i,j}}{\sum_{i \neq v} R_{i,v}}$$

Note that  $j$ ’s interval is disjoint to the intervals assigned to every  $v \neq j$ . A random number is then generated on  $[0, 1)$ , which must fall into one of these intervals—say, the interval for vertex  $v$ . Then an edge is created between  $i$  and  $v$ . This process repeats until the chosen value of  $k$  is realized by the graph. [8, pp. 46–47]

By fixing  $n$  and  $k$  in advance, we can guarantee that the resulting graph will be sparse; thus, we are interested in the effect of the other parameters on the properties of clustering and characteristic path length.

The probability  $p$  that a vertex  $i$  will connect to a vertex  $j$  with which it shares no neighbors is what allows the algorithm to get up and running. Starting from a set of vertices with no edges between them,  $p$  guarantees that every  $R_{i,j}$  is non-zero, so that the intervals of  $[0, 1)$  corresponding to the vertices  $j$  actually have a width. Without this condition, no edges would ever be formed. However,  $p$  contributes to the ‘random character’ of the resulting graph: the larger  $p$  is, the better chance there is of  $i$  connecting to a vertex  $j_1$  with which it shares no neighbors as compared to the chance that  $i$  will connect to a vertex  $j_2$  with which it shares one or more neighbors. For this reason,  $p$  is kept very small, so that the amount of randomness in the graph generated is almost completely determined by the central parameter,  $\alpha$ .

The parameter  $\alpha$  of this algorithm is used to vary the amount of randomness in the resulting graph. When  $\alpha = 0$  (a circumstance Watts refers to as the “Caveman world”), new connections are formed at each iteration almost entirely on the basis of the existing edges, since  $R_{i,j} = 1$  for all vertices  $i$  and  $j$  that share at least one neighbor, and  $R_{i,j}$  is small (equal to  $p$ ) for vertices that have no common neighbors. The resulting graph is highly clustered, consisting mostly of isolated ‘caves’ of vertices that are all connected to each other, but not to other vertices outside the group. On the other hand, as  $\alpha$  approaches infinity (a circumstance Watts calls the “Solaria world”), new edges form almost entirely at random, since the  $R_{i,j}$  approach 0 for vertices that share at least one but less than  $k$  neighbors. Hence, the other cases (where  $m_{i,j} \geq k$  or  $m_{i,j} = 0$ ) become relatively more important in determining the structure of the graph. Since it is extremely rare for two vertices to share more than  $k$  neighbors, most edges form with probability  $p$ , and the resulting graph turns out to be mostly random in its structure.

Another important parameter in determining the outcome of this construction algorithm is the choice of what Watts calls a “substrate,” an existing set of edges between the vertices that serve as an input to the algorithm. The substrate can have a strong effect on the structure of the output graph. If the substrate is empty, for example, then running the algorithm when  $\alpha = 0$  often results in the disconnected graph of isolated ‘caves’ mentioned above. The characteristic path length  $L$  of such a graph is infinite by definition. To avoid this problem, Watts chose to use a ring substrate (i.e., a graph in which every vertex has exactly two neighbors), so that the resulting graph would always be connected and therefore have a finite value of  $L$ , even at low values of  $\alpha$ . This aids in the comparison of graphs across a range of values for  $\alpha$ , but it’s a significant fix-up that reduces the size of the class of graphs that the algorithm could construct, so Watts spends a fair amount of time justifying his choice. The use of other substrates have varying effects depending on their topologies; Watts chose a ring substrate mostly because it had the *smallest* effect on the properties of clustering and characteristic path length. [8, pp. 58–66]

### 3.3 Conclusions from the $\alpha$ -model

Does the  $\alpha$ -model produce graphs that have the desired properties of both high clustering and low characteristic path length? To make a long story short: it does. By running the algorithm many times at specified values of  $n$ ,  $k$  and  $p$  and averaging the properties of  $\gamma(G)$  and  $L(G)$  over the resulting graphs, Watts found that as  $\alpha$  increases, both the path length and clustering of the graphs increase briefly, then drop off sharply. The path length, however, consistently peaks and then falls *before* the clustering coefficient, so that there is a class of graphs in which  $\gamma(G)$  remains high while  $L(G)$  is very low. (For graphs of 1,000 vertices with  $k = 10$ , for example, this happens when  $\alpha$  is between 5 and 10.) [8, pp. 52–58]

The  $\alpha$ -model therefore demonstrates that, in the abstract, small world graphs do exist. This is the most important conclusion that can be derived from analysis of the outputs of the model: the “Small World Phenomenon,” as Watts calls it, does arise for a certain class of graphs among all the graphs that could be generated by the  $\alpha$ -model’s construction algorithm, and it arises “in the transition from order to disorder” that occurs with increasing  $\alpha$ . As a first step in the investigation of the problems of defining, describing, and explaining small world networks, the  $\alpha$ -model is a great success.

There are also some secondary conclusions to be derived from the  $\alpha$ -model. Watts discusses some empirical results in connection with the ‘relational graph model,’ a construction algorithm based on the  $\alpha$ -model and its successor, the  $\beta$ -model. He compares graphs generated by the relational graph model to three real networks for which complete data about the connections is available: the “Kevin Bacon graph,” in which the vertices are the set of Hollywood actors, and the edges consist in the relation of having acted in some movie together; the “Western states power grid graph,” in which the vertices are power stations and the edges are major power transmission lines in the western United States; and the “*C. Elegans* graph,” in which the vertices are neurons and the edges are synaptic connections in the nervous system of the famous roundworm *Caenorhabditis elegans*. Without exhaustively reviewing the results here, I can give Watts’ conclusion from them: each of these networks could be considered a small world, since they are relatively sparse, and have relatively high values of  $\gamma$  and low values of  $L$ ; random graphs generated on the same number of vertices as each of these networks also have low  $L$ , but fail to have high  $\gamma$ ; and graphs generated by the relational graph model tend to fare better in modeling both  $\gamma$  and  $L$  in these networks than the random graph model, though both fare pretty poorly in modeling clustering in the *C. Elegans* graph. It therefore seems that the relational graph model (and, by extension, the  $\alpha$ -model) fits the available empirical data better than other mathematical models available. [8, pp. 139–161]

It is important to note, though, that Watts never directly compares graphs generated by the  $\alpha$ -model to real networks, at least not in his presentation of the data in *Small Worlds*. Of course, this does not make the  $\alpha$ -model any less of a scientific theory: it was proposed as a preliminary means of investigating



an empirical problem, it is based on plausible assumptions about real social networks, it is mathematically rigorous, and it distinguishes as important two properties of networks than can be measured in both the models of the theory and in empirical data. By anyone’s standards, the  $\alpha$ -model surpasses mere mathematical formalism or pseudo-science. It’s just that, as a scientific theory, it was never intended to represent phenomena in the way philosophers of science often imagine theories do. I shall now turn, therefore, to seeing whether the state space view of scientific theories can adequately describe the kind of theorizing the  $\alpha$ -model presents.

## 4 A Challenge to the State Space View

To show that the  $\alpha$ -model presents a challenge to the state space view, I will defend the following claims in turn:

1. The  $\alpha$ -model presents a family of structures.
2. The  $\alpha$ -model specifies certain properties of these models as candidates for “the direct representation of observable phenomena.”
3. The  $\alpha$ -model does involve various idealizations and simplifying assumptions, but not necessarily of the kind that the state space view envisions.
4. The relations between the models of the  $\alpha$ -model and the world are not easily described by the notion of ‘fit’ in a state space.

If the first two of these claims are true, it means that the  $\alpha$ -model meets the requirements of the basic semantic view for being a scientific theory, and so is a candidate for further description under the state space approach. If the second two are true, however, it means that the state space approach will not tell an adequate story about how the  $\alpha$ -model functions as a theory; in this sense, the  $\alpha$ -model presents a challenge to the state space view.

### 4.1 The $\alpha$ -model and the basic semantic view

It should be clear that the  $\alpha$ -model does indeed present a family of structures. The structures it presents are a class of semi-random undirected graphs. Watts’ use of the term “ $\alpha$ -model” to refer to this whole class of graphs (or, more precisely, the rules by which those graphs are generated) is consistent with Van Fraassen’s claim that, in scientific use, “model” typically means “model-type.” The particular graphs generated by the  $\alpha$ -model, for a specified number of vertices  $n$ , and specified values of  $k$ ,  $p$  and  $\alpha$ , are the actual (instantiated) models of the theory, on the semantic view. It’s also clear that this class has quite a lot of members—at least enough so that Watts could do the statistics required to draw his conclusions—but not *every* undirected graph belongs to it. It would be extremely unlikely, for example, that an iteration of the  $\alpha$ -model algorithm would produce an undirected cycle. Thus, the family of structures

presented by the  $\alpha$ -model is a non-trivial one: it has multiple members, so it's properly called a "family," but it's not so inclusive as to be uninteresting; an undirected graph must really have certain properties in common with other members of the class in order to be among the models presented by the  $\alpha$ -model.

In this respect, the  $\alpha$ -model is an almost ideal example of the semantic view's claim that theories are best understood as collections of models. Watts himself notes that the construction algorithm he gives is only one possible way of picking out the class of graphs that he intends [8, p. 46]; it's really those graphs, and not the way they are constructed, that are the heart of the theory. Moreover, the models are very simple, in the sense that they are completely characterized by simple mathematical structures (the set of vertices and the edge relation on them), and they are even of the sort that logicians are familiar with. I therefore do not think any semantic view theorist should dispute claim (1).

The second claim is equally uncontroversial. The  $\alpha$ -model specifies properties exhibited by its models as candidates for the direct representation of observable phenomena: namely, the properties of path length and of clustering. Indeed, Watts' stated purpose in constructing the model was to find a simple algorithm that would pick out a class of graphs in which these properties coexisted. The concepts by which these properties are defined are also 'observable' in real networks, on any reasonable definition of that term: given enough data about a network, it is easy to compute the number of vertices it has, its average degree, the length of a path between two vertices, and so forth. Even the positivists would be satisfied that the  $\alpha$ -model does not employ any irreducibly theoretical concepts.

I take it that the basic semantic view's empirical claim, as I sketched it out above, is mostly meant to exclude purely mathematical theories from the domain of theories it is attempting to characterize. True scientific theories must be prepared to say something about the observable world. The empirical claim says only that the models of a scientific theory should have properties that *could* represent phenomena; it says nothing about how *well* the models must do so, and it says nothing about the intentions of the scientist who constructs or uses the theory in his investigation of an empirical problem. The  $\alpha$ -model therefore satisfies both the semantic claim and the empirical claim, even though Watts may not have intended that it describe any real networks very faithfully, so it's a scientific theory that deserves further examination under the state space approach.

## 4.2 The $\alpha$ -model and the state space approach

How well does the  $\alpha$ -model fit the picture of scientific theories that the state space approach favors? To answer this question, we have to see whether it satisfies the idealization claim, the state space claim, and the model-data relation claim.

I claimed in (3) above that the idealizations of the  $\alpha$ -model are not of the sort that the state space view envisions. Recall that the models of a theory  $T$  are described completely by some set of parameters that scientists 'abstract' from

the phenomena in the intended scope of the theory. The models of  $T$  are said to be idealizations because the unisolated systems that scientists are attempting to describe and explain with  $T$  are under the influence of other, unknown factors.

The parameters of the  $\alpha$ -model are: the number of vertices  $n$  and average degree  $k$  that must hold in the output graph; the baseline probability  $p$  that two vertices will connect on an iteration of the construction algorithm when they share no neighbors;  $\alpha$  itself; and the substrate graph. It should be clear that  $n$  and  $k$  are the sort of parameters that are abstracted from the phenomena in the way that the state space view envisions: every network must have values for these parameters, no matter how it was generated. Probabilities are also standard parameters in scientific models, though they may or may not be abstracted in any obvious way from existing data, so I shall assume that the state space view has no problem accounting for  $p$ .

The other parameters, however, are more of a problem.  $\alpha$  is explicitly given as a kind of mathematical ‘knob’ by which the output of the construction algorithm is controlled; it is a formal construction without an intended correlate in real networks. Likewise, the substrate graph that is fed into the construction algorithm has an important effect on the structure of the graph that comes out; but the inclusion of a substrate is also more a formal property of the theory (one that had to be justified) than a parameter extracted from real data. In general, it is not possible to examine a network and decide what sort of substrate it must have been built on, for it may not have been built on any substrate at all. Already, then, the state space view’s account of idealization seems not to capture what happens in the  $\alpha$ -model.

What about the claim that the models that come out of the construction algorithm must be idealizations because the parameter list is incomplete? It’s true that, with respect to the properties of clustering and characteristic path length, other parameters might influence the values of these parameters in real networks in a way not captured by the  $\alpha$ -model.  $n$ ,  $k$ , and  $p$  are certainly not enough to determine the values of  $\gamma$  and  $L$  for a given graph, and we know that  $\alpha$  and the substrate are not parameters that have any real correlates, so something else not mentioned in the  $\alpha$ -model must determine  $\gamma$  and  $L$  in the domain of small world networks. This certainly looks like the sort of idealization the state space view has in mind, then.

I would point out, though, that the state space view of idealization carries with it the idea that the parameters of a theory *would* be sufficient to characterize the phenomena in the theory’s domain if those phenomena were free of the influence of unknown parameters. For Watts’ theory, this is probably not true: there is no guarantee that even all the possible combinations of  $n$ ,  $k$ ,  $p$ , and  $\alpha$  would generate the variety of  $\gamma$  and  $L$  found in real networks. (I exclude substrate as a parameter here because, if one is allowed to pick the substrate from all the possible undirected graphs, a substrate could be selected that already had all the desired properties, making the representation of all  $\gamma$  and  $L$  a trivial matter.) Moreover, I doubt that Watts would claim that his theory was ‘idealized’ in this way: if his goal had been a completely faithful representation of what the properties of real networks would be in the absence

of other parameters, he would not have made an unrealistic parameter like  $\alpha$  the central parameter of his theory.

I do not deny, therefore, that the state space view has something interesting to say about idealization in the  $\alpha$ -model, but I think the true story is more complex than even the somewhat strong notion of idealization presented by Suppe suggests. The  $\alpha$ -model is idealized not just in the sense that its parameter list is incomplete, but also in the sense that some of its parameters do not correspond in any straightforward way to properties of phenomena in its intended domain. I maintain, then, that claim (3) above is a reasonable one.

And what of the fourth claim, that the notion of ‘fit’ in a state space between the structures of the  $\alpha$ -model and observation data does not capture the ways in which the theory is compared to the world? Suppose we constructed a state space for the  $\alpha$ -model. What would this space look like? The most natural way to build it would be to include dimensions for all the numerical parameters of the theory (that is,  $n$ ,  $k$ ,  $p$ , and  $\alpha$ ) and dimensions for the two properties we are interested in,  $L$  and  $\gamma$ .

Since the  $\alpha$ -model was not presented as a theory intended to model the *topologies* of real networks (that is, the exact set of edges on a specified number of vertices), we do not need to include any dimensions in the space that would encode graph topology, even though it would be *possible* to compare the topologies of real networks to the topologies of  $\alpha$ -model graphs. Watts was really only interested in seeing how the macro-properties of clustering coefficient and path length were affected by varying  $\alpha$ ; so long as our space includes those three dimensions, then, it is complete.

We can now say that each graph  $G$  generated by the construction algorithm of the  $\alpha$ -model corresponds to exactly one point in this space, which we can name using the vector  $\langle n_g, k_g, p_g, \alpha_g, \gamma_g, L_g \rangle$ . Note that this graph need not be unique: another, different graph  $H$  generated by the construction algorithm could have the same values along each dimension. Since the models of the  $\alpha$ -model theory are not dynamic, we don’t extend them to include a series of successive states in the state space—they are not ‘trajectories’ but points, so the traditional notion of statistical fit of data to curves in this space won’t work as a comparison of the  $\alpha$ -model and the world. Moreover, because a point  $\langle n_i, k_i, \alpha_i, \gamma_i, L_i \rangle$  does not uniquely determine an  $\alpha$ -model graph, it would be strange to think that the fit of some collection of data to a *point* in the state space is a comparison of one model to the world. The notion of ‘fit’ between real networks and the models of the  $\alpha$ -model theory must lie elsewhere.

As I said earlier, this is not necessarily a problem for the state space view. There are many ways that we can define fit, and many ways we can extend it. I do want to emphasize, though, that an important part of the state space view’s concept of fit is that data be compared to a single, specific model of the theory. Data can fit different models of the same theory more or less well, and goodness of fit might be the very relation that helps us pick out which model most faithfully represents the particular phenomena under study; but we fit the data to one model at a time. It doesn’t make sense in general to say that data fit two models of the theory *simultaneously*, or that data fit the disjoint union

of two models: we can draw lines describing how far a given data point is from one model curve or the other in the state space, but we can't generally draw a single line that describes "how far away it is from both."

I admit that this is a rather subtle conclusion to draw from the state space view, and that van Fraassen, Suppe, or Lloyd might easily deny the claim if presented with it directly. I think it is implicit in their discussion of how models are related to data, though, and there's no reason it shouldn't be: traditional statistical techniques are supposed to describe that relation, and those techniques are generally based on calculating the fit of a data set to one model at a time. It's just this feature of the notion of fit, though, that I think causes problems with the  $\alpha$ -model: the most important ways in which the  $\alpha$ -model relates to the world are not ways in which data about a single real network is compared to an  $\alpha$ -model graph.

The best way to see this is to look at how Watts uses the  $\alpha$ -model in his investigation of whether or not small worlds can exist in real social networks. His process, as I reconstruct it, went something like this: first, he chose values for  $n$ ,  $k$  and  $p$ . Then he used the  $\alpha$ -model's construction algorithm to generate a (reasonably large, for statistical purposes) number of graphs using a ring substrate and a particular value of  $\alpha$ . For each of the resulting graphs, he calculated the clustering coefficient and characteristic path length, and he averaged these values over the whole collection to obtain a single point  $\langle \alpha_i, \gamma_i, L_i \rangle$ . He repeated this process at successive values of  $\alpha$  until the average values of  $\gamma$  and  $L$  stabilized.

By plotting the points  $\langle \alpha_i, \gamma_i, L_i \rangle$ , Watts was able to conclude that some sparse graphs do have a high clustering coefficient and a low characteristic path length. This led him to create new algorithms (the  $\beta$ -model and the relational graph model) that explored the small world phenomenon in different ways. He eventually obtained data about some real networks, and attempted to fit that data to the relational graph model, which was in part derived from the  $\alpha$ -model. He discovered that the relational graph model generated graphs whose properties tended to fit the data from the real networks better than the properties of random graphs.

If this reconstruction, which follows the arc of Watts' book, is anything like what really happened, it seems Watts was never very much concerned with how  $\alpha$ -model graphs fit any real data. The relation between the models of the  $\alpha$ -model and the world is mostly one of 'showing small worlds to be possible': that is, the collection of graphs generated by the  $\alpha$ -model showed that social networks *could* simultaneously exhibit high  $\gamma$  and low  $L$ , given a few plausible assumptions.

If we wanted to represent this relation using a state space, we would probably select an 'empirical subspace' of the total state space where the values of  $n$ ,  $k$ , and  $p$  are fixed, as Watts did. We would show that we could generate a large number of graphs whose values for  $\alpha$ ,  $\gamma$ , and  $L$  were all bounded by some solid in that space, and that solid contained a region where  $\gamma$  was high and  $L$  was low. We would argue that the existence of this solid showed that *if* we collected data about real networks, and represented those networks as undirected graphs,

those networks *could* be small worlds, because the data might fall into this region of the state space; there is more reason to think that it could than that it couldn't, given the properties of the output graphs from the  $\alpha$ -model.

I am belaboring this point in order to make it plain that something other than 'fit' must be the relation between the  $\alpha$ -model graphs and the world. We are not comparing real data, or potentially real data, to any particular model in the  $\alpha$ -model state space. Rather, we are using the range of values of  $\gamma$  and  $L$  in the whole *collection* of models to argue that real data could fall in this range as well. We could not make this argument without reference to the whole collection; since each particular model is a graph generated by a semi-random procedure, we would have no way of using a single model to demonstrate the real possibility of small worlds, because any particular model could be an unlikely case or a mathematical fluke. If the state space view's notion of 'fit' was the only way that scientists could compare theories to phenomena in their intended scope, the  $\alpha$ -model would not have been able to show the important results that it did. That, I think, is reason to believe in the soundness of claim (4) above.

## 5 Conclusions

I have argued that the  $\alpha$ -model is an empirical theory that presents a set of models whose properties are candidates for empirical confirmation, in the sense that they might be exhibited by real networks. This means that the  $\alpha$ -model is a structure that proponents of the state space approach believe they can provide a description for in terms of the way it idealizes its representation of phenomena and the way observation data fit its models. I have shown that the idealization inherent in the  $\alpha$ -model is not entirely described by the idea that 'models are determined completely by an incomplete set of parameters': not only are the parameters of the  $\alpha$ -model incomplete (they were never intended to be anything else); at least two of them are formal devices that have almost no interpretation in real phenomena, and could not have been abstracted from any observations. I have also shown that the  $\alpha$ -model does have a state space, or at least we can construct one for it, and that each of its models occupies a point in this space. Nevertheless, it seems that the most important way in which the  $\alpha$ -model relates to the world is not adequately characterized by the state space view's notion that the model-data relation is that of statistical fit. As a collection, its models demonstrated that it would not be fruitless to try and gather data about the small world phenomenon; but no data set was ever fit to one of those models, even under a broadened conception of fit.

I therefore claim that the state space approach cannot describe the  $\alpha$ -model theory without extending its picture of how scientific theories work. Objections to this claim would probably come in one of two forms: either that I have a misinformed view about what the state space approach to theory analysis entails, or that I have been unimaginative in my description of the  $\alpha$ -model as I applied that approach, so that my description was inadequate. I am of course open to corrections of the first sort. I do not claim to have given a complete

summary of everything held by the semantic view of theories, or by the state space view in particular. I have, however, tried to give the views of multiple authors when they varied in their opinions, and to characterize the semantic view in the most judicious way possible, so I think I should be relatively safe on this front.

It's certainly true that the challenges to the state space approach that I raised in the previous section could be resolved by modifying the notions of idealization and of fit. The notion of idealization given by Suppe in his summary of the view is too limiting because it insists that the parameters of a theory be abstracted from phenomena, which makes no room for theories that have a parameter like  $\alpha$ . This is easily fixed by enumerating the other types of idealization that operate in scientific theories. As for fit, it would also be possible to enumerate other types (such as 'fit' between a *collection* of models and data, which would be required to describe the  $\alpha$ -model's relation to data), but I think this would be to mask the real issue. Theories bear other relations to natural systems—and to scientists, and to other theories, and so on—than the relation of statistical fit between their models and data: I might compare a theory and the world by showing that a theory says certain phenomena are impossible, for example, and looking to see if that prediction holds. These other relations are missing from the state space view, as I think the  $\alpha$ -model demonstrates.

Of course, my criticisms are not worth much if they are not general. The  $\alpha$ -model is in several respects an atypical theory, and if the state space view only has trouble describing a few atypical theories, it won't have much trouble at all. I think, though, that there are probably many analogous theories across the spectrum of science. The reason is that the  $\alpha$ -model is what might be called a 'preliminary theory' or a 'theory of a problem.' Its purpose was to answer the questions, "Can small world networks be defined by some set of measurable properties?" and "If so, is it possible that real networks could have those properties?" By answering those questions affirmatively, the  $\alpha$ -model laid the grounds for further empirical work.

It's because the  $\alpha$ -model was a theory of a problem in this sense, aimed at showing the *possibility* of further work rather than representing phenomena in a faithful and direct way, that it fell outside the state space view's usual notion of a scientific theory. But questions about whether it's possible to define and investigate an empirical problem in some domain are probably as common as the domains themselves, and 'theories of problems' intended to answer those questions might be just as common. Scientists have limited resources, and they need to answer such questions before they plunge whole-heartedly into gathering exhaustive data and building models. In physics, for example, scientists formulate theories about the properties of new particles before they try to design the accelerators that will detect them. Likewise, in the science of networks, the  $\alpha$ -model provides a reason to look further into the small world phenomenon. To deny the importance of the preliminary investigations that answer questions about the possibility of further empirical work in either of these fields would be a mistake on the part of any philosophy of science; a good philosophy, by contrast, should account for them.

## References

- [1] Roman Frigg. Scientific representation and the semantic view of theories. *Theoria*, 55:49–65, 2006.
- [2] Elisabeth A. Lloyd. *The Structure and Confirmation of Evolutionary Theory*. Princeton University Press, 1988.
- [3] Stanley Milgram. The small world problem. *Psychology Today*, 2:60–67, 1967.
- [4] Frederick Suppe, editor. *The Structure of Scientific Theories*. University of Illinois Press, 1974.
- [5] Patrick Suppes. A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese*, 12:287–301, 1960.
- [6] Patrick Suppes. *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, chapter Models of Data, pages 252–261. 1962.
- [7] Bas C. van Fraassen. *The Scientific Image*. Oxford University Press, 1980.
- [8] Duncan J. Watts. *Small Worlds*. Princeton University Press, 1999.