Models and Fictions in Science

Peter Godfrey-Smith

Harvard University

1. Introduction

This paper discusses a philosophical problem generated by new accounts of one part of
scientific practice. I also discuss possible elements of a solution, but the emphasis is on the
difficulties.[1]

      Scientists, whose business is understanding the empirical world, often spend their
time considering things that are known not to be parts of that world. Standard examples are
ideal gases and frictionless planes. Examples also include infinitely large populations in
biology, neural networks which learn using biologically unrealistic rules, and the wholly
rational and self-interested agents of various social-scientific models.

      A natural first description of these things is as fictions, creatures of the imagination.
They do not exist, but at least many of them might have existed, and if they had, they would
have been concrete, physical things, located in space and time and engaging in causal
relations. Though imaginary, these things are often the common property of a community of
scientists. They can be investigated collaboratively. Surprising properties might be uncovered
by one investigator after being denied by another. In their status, though not their role, they
seem analogous to the fictions of literature.

      This natural-looking description is quite often resisted. I will give some reasons in
support of it, but the paper is not so much defence of this view as exploration of its
consequences. I also assume, without argument, a kind of scientific realism. By means of
theorizing, including the theoretical investigation of fictions of this kind, we learn about the

---

world, and not only about its observable structure. The paper investigates the consequences of those two ideas when combined.

One goal of the philosophy of science I take to be giving us a total picture of science of the following kind. One part of the picture is an account of scientific *practice,* broadly understood. This is an account of how scientists develop and investigate ideas, which representational tools they employ, and how choices for one view over another are made. A second part, more philosophically tendentious, is an account of what all this activity *achieves* – how it relates to the world at large, what kind of knowledge it makes possible. That second part of the picture is constrained by, as well as informing, our general metaphysics and epistemology. My argument will be that in the case of "model-based" science of the kind discussed here, sketching such an overall picture is particularly problematic. There is a picture suggested by a good description of scientific practice, but this picture sits badly with constraints we get from the other side. The fairer one is to the practice, the harder it is to tell the overall epistemological story.

2. Model-based science

The topic of the paper is "model-based science." This is understood as one style of science, not as what *all* science is, deep down. This is the style in which a paper may begin: "imagine a population of self-replicating molecules...," "assume a three-layer neural network learning by back-propagation...," or "consider a collection of agents playing one-shot prisoner's dilemmas at random...." What this seems to be is a style of theoretical work in which an imaginary system is introduced and investigated – an imaginary population, ecology, neural network, stock market, or society. The behavior of the imaginary system is explored, and this is used as the basis for an understanding of more complex real-world systems.

In giving a philosophical analysis of this work, a good starting-point is Giere's 1988 analysis of science. Giere presented this as an analysis of *all* theoretical science, but it is more useful when applied differently (Godfrey-Smith 2006a, Weisberg 2007). The main idea can be summarized with a diagram (Figure 1). Modifying Giere a little, this picture has it that the scientist uses words, mathematics, or some other representational medium to *specify* a *model system.* The model system can then be analyzed, described, and argued about. Once

understood, the model system can also be compared to various real-world *target systems*. The comparison is made on the basis of similarity relations between the two.
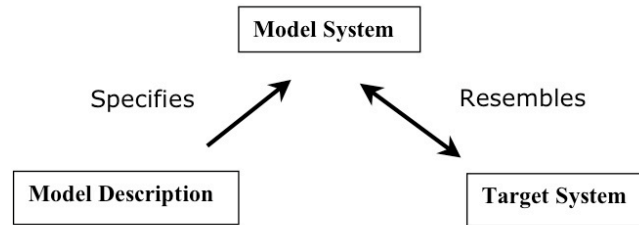

Figure 1. Modified from Giere (1988).

The appeal to similarity here is especially controversial. Goodman subjected such appeals to withering criticism in the mid 20th century (1972), and Giere's picture has been seen as suffering severely from these problems (Suárez 2003). Still, this picture gives us a good first description of one style, mode, or element of theoretical science, a style characterized by *indirect* representation of natural affairs. This indirectness is not just a role for imagination and hypothesis, or for epistemic caution, but involves a *deliberate detour* through fiction. Sometimes the reality of this category is as overt as a job description. Sometimes it looks more like one element or facet of an ever-shifting mix. Either way, it is not the only style of science or the only ingredient in the mix. Darwin's *Origin of Species* (1859) is not work of model-based science, and elsewhere (2006a) I have discussed examples from more recent biology. Model-based science has, to some extent, a history and sociology of its own; it differs in prominence across times and fields. It has become more prominent recently, and has also begun to achieve a kind of self-consciousness. It is often motivated by a combination of complexity in the target plus a motivation to employ particular kinds of exact methods (very different from Darwin's methods). It can also be motivated by the search for high generality.[2]

---

[2] The account given here is not a contribution to the "semantic view of theories," which aims to give a uniform analysis of *all* theorizing and draws on either the logician's notion of a model, or something closely related to it (Suppe 1977, Van Fraassen 1980, Lloyd 1988, French and Ladyman 1999). But elements of the view will also be assessed in a modified role at the end of the paper.

In the sketch above I focused on cases where the model system does not physically exist, and is investigated by scientists whose ultimate aim is understanding something else. Discussions of this work have also focused on cases where the model system is described mathematically. But model-based science in a broader sense includes at least three other kinds of work. One kind uses physical models. Here one real system is constructed and analyzed in order to understand another (Sterrett 2002 & 2006a, Weisberg forthcoming). Wind tunnels are a standard example. Engineers also still construct elaborate scale models of hydrodynamic systems such as river deltas and bays; after Hurricane Katrina the US Army Corps of Engineers built a 1/50 scale model of part of the New Orleans canal system to determine how it responds to waves and water levels.[3] A second category is work using computer simulations. A third involves the construction of informal "word models" that sketch possible processes and mechanisms, perhaps with a flow-chart or similar device. This work shades into what would usually be described as "thought-experiment" (Frigg forthcoming). Work on deliberately built physical models also shades into work using "model organisms" such as *E. coli* and yeast, and modern model organisms are often, in fact, partly human constructions (eg., "knock-out mice" in genetics). Whether or not there is a detour through "fiction," the same strategy of indirectness is visible in all these cases. A final feature of this work is that the word "model" itself is used diversely. So it is not a good idea to organize discussion around the question: "what are models?" Sometimes the term will be used for a model-description, or for a collection of them; sometimes it will be used for a model-system, or a collection of those. The style of science, along with its relatives, is fairly clear. The word "model" is used variously within and around the practice.

3. Fictions in modeling practice

We now start to look at an obvious question raised by the sketch above. What is the nature of the thing at the top of the triangle, in cases where the model system is not a physical object? In this section I address this question from the point of view of practice in model-based science; I look at what this thing is *taken* to be, how it is treated. Then we find that, at

---

[3] Information on the New Orleans model can be found at:
http://chl.erdc.usace.army.mil/chl.aspx?p=s&a=Projects;207. I am grateful to Susan Sterrett for introducing me to this example, and to both Sterrett and Michael Weisberg for information about the contemporary use of physical models.

least much of the time, model systems are treated as fictional systems that lack actual existence but are, in a sense, of the same kind as the target systems that the models are used to help us understand. Often the concern is with sets of cases rather than single cases, on both sides, but each model system itself is something that would be concrete if real; it would be an arrangement of physical entities, located in space and entering into causal relations. I don't claim it is always that way. There is a variety of styles within model-based science, and this includes a sort of "gradient of abstraction" with respect to the handling of models. Some of the variety will be discussed later in this section, but a handling of model systems as imagined concrete things, as fictional *concreta*, is a clear feature of much of this work. Model-based science often presents itself as an investigation of *imagined simpler analogues* of a target or targets.

This suggests an analogy between model systems and fictional objects introduced in novels and other forms of literature.[4] The world of a novel is something that does not actually exist, but would be concrete if real; it is apparently a *candidate* for physical existence. If real, it would be an organized system characterized by causal relations among its components. Such fictional objects, both in literature and science, can often be mathematically described, but that does not imply that the fiction is itself a mathematical object.

The issue of similarity between models and targets also appears unintimidating from this point of view. As far as practice is concerned, all of these seem to be roughly comparable.

(i) Comparison of two physical systems. The rate of poverty is similar in this society and that one. Bacterial genetics is like human genetics with respect to the basic "code" and the building blocks, though not with respect to the organization of genetic material. Here we also find the style of modeling in which a model system is physically built: the pattern of flooding in this warehouse-size system is similar to a pattern of flooding that occurred in New Orleans.

(ii) Comparison of two fictional systems. Tolkein's "Middle Earth" is fairly similar to the world of Malory's King Arthur tales (*Morte D'Arthur*) – similar technologies are available, there are a

---

4   For presentations of this idea see Godfrey-Smith (2006a), Frigg (forthcoming), Thomson-Jones (forthcoming).

lot of wanderers. Discrete-generations and continuously breeding models of evolutionary change give similar results in simple cases (Crow 1986).

(iii) Comparison of a fiction with a physical system. The events in Orwell's *Animal Farm* are similar to those in Russia in the first part of the 20th century. Evolution in a deterministic one-locus evolutionary model with heterozygote advantage is a lot like what goes on in some human populations in Africa, with respect to the sickle-cell anemia trait. According to Robert Axelrod (1984), outcomes seen in some iterated prisoner's dilemma simulations resemble behaviors found in the trenches in World War I.

In each case, what we seem to be doing is comparing properties associated with one system with properties associated with another. The phrase "associated with" is used here as a vague cover-all. In the case of comparing two physical systems, we are looking for the sharing of physical properties themselves.[5] In the cases involving fictions, how we describe the situation depends on an underlying theory of fictions – I will return to this below. We might say we are comparing features *attributed to* things in the fiction, or implied by features that are attributed. We might instead – or as well – say we are comparing properties the two fictional systems would have if they were actual. Or we might, with Lewis (1986), say something that treats the fictional system as a non-actual but real thing that *has* properties in the same sense seen in the case of the actual system.

Whichever way the details go, these comparisons are guided by mappings between properties of the kinds exhibited by ordinary objects – poverty maps to poverty, acidity to acidity, wandering to wandering, stability to stability. There is no need – at least so far – to step back to consider only the formal backbone that the systems may instantiate – the mix of reflexive, symmetric, and transitive relations that can be identified in the overthrow of a ruler or the retention of a gene. We *might* find reason to look for these more subtle formal similarities across systems, either actual or fictional, so that relation R_1 is mapped to R_2 in virtue of things other than the physical character of the relations. We might, via abstraction and isomorphism, map a political process to a mechanical one, or learning to natural selection. But that is a special kind of similarity assessment that goes beyond the routine one

---

[5]   Or, often, we are looking for the instantiation of nearby determinates of the same determinable (Thomson-Jones forthcoming).

seen when we say that *Animal Farm* is like Soviet Russia. The analogy with fictions also makes clear that there is nothing philosophically derailing in the context-sensitivity of similarity judgments. With Goodman and others, we can allow that in *some* sense, everything is similar to everything else. That does not stop the assessent of relevant similarities across fictions, and between fictions and actual systems, being informative and significant. When *Animal Farm* was published, no one thought it had no special relation to Russia because everything is similar to everything else, or because Russia was not ruled by farm animals.

This analysis also sheds light on the role of computer simulations of complex systems, a style of modeling which fits awkwardly into other philosophical descriptions. Many strange things are said, especially by scientists, about computer models of such things as planets, populations, and economies. They may say: "the population is inside the computer." This is not meant literally, of course, but it is seen as somehow along the right lines. I think it is a very strange thing to say, even granting much metaphorical license. Instead, what the computer does is act as an aid to the scientific imagination; its role is dependent on the imaginative activities discussed above. A modeler specifies some set-up, some significant possible configuration of organisms or atmospheric factors, and finds a way to use a computer to determine the consequences of such a set-up. Physical regularities in the operation of the computer are exploited to tell us what would happen to a system of such-and-such a kind, to tell us the consequences of the scenario we imagined. The computer is needed because our ability to specify such configurations outruns our ability to work out how they would behave. Sometimes a computer simulation might be developed and then someone may come along and show that an analytic treatment of the problem was possible after all. The object of analysis – the imagined scenario – can be the same, though the methods brought to bear on it may change.

The investigation of fictional systems is central to model-based science, but I do not want to argue that fictional *concreta* are always involved. I now discuss some qualifications and complications. First, work may become focused on mathematical structure itself, for whatever reason. At the other end of the scale, talk in a modeling context may be directed squarely on a real-world physical system, and take the form of approximate description *of* that target. Here it is clear that an empirical system is the object of the discussion, and what is distinctive is the fact that people are engaging in special forms of description of it

involving imagined modifications.[6] Levy (forthcoming) argues that these cases may be analyzed with Kendall Walton's concept of "prop-oriented make-believe" (Walton 1993). We engage in games of make-believe about an actual object, with the purpose of learning more about that very object. The aim may be to investigate which of its causal properties are "difference-makers" in some context and which are not.

The thing that makes a difference here is the fact that the only object of discussion is the real-world target. I agree that some fictionalizing in science can be understood like this. But some of it looks different. The role of some other, very significant, fictions in modeling might be illustrated with a "hub-and-spoke" analogy. In these cases, what scientists do is give an exact description of one case of the target phenomenon, which acts as a "hub" that anchors a large number of other cases. The "other" cases include *all* the actual-world ones; the hub is a fiction. The central models of both evolutionary change and population growth within modern biology work like this, for example (Roughgarden 1979, Maynard Smith 1998). When a scientist deals with an actual population, he or she combines knowledge of the hub case with many *ad hoc* tools relevant to the case at hand. (The evolutionist uses scraps of demography, scraps of ecology, etc., as needed.) Approximate knowledge of actual cases is achieved via exact knowledge of a hub case plus a shifting array of more empirical concepts and methods. This organization of theory is one in which idealized models do not go away once knowledge becomes highly developed. They retain an explanatory role as a consequence of their generality (Levins 1966, Weisberg 2004).

I turn to a third apparent complication. Within work guided by fictions in the present sense, the fictions are usually very schematic, with few properties discussed. Often what is specified is a large class of systems, unified by algebra. This initially seems a point of disanalogy with the literary case, but in fact it is not. Philosophical discussion of fiction often focuses on novels, especially 19th century novels, which fill out their stories with a lot of detail. But there are also more schematic, parable-like fictions. In both the literary and scientific cases, there is variation in the extent to which fictions are filled out. Richly realistic novels are akin to elaborate computer simulations. Spare and schematic fictions are akin to abstract analytic models. Parables are like narrative algebra.

---

[6] Klein (forthcoming) calls these "quasi-idealizations." He distinguishes these from from (genuinely) idealized models which cannot be seen as representations of particular real systems because so much that is causally important in those systems is altered or omitted.

Both scientific and literary cases might be represented formally by invoking large sets of fully-specific possibilities that are compatible with a given description. Phenomenologically, though, it does not usually feel like that. Instead, it seems to be a general fact about people that we can engage in a kind of schematic imagining in which some elements of a scenario are made definite and others are left open. These models and stories, in their psychological role, are a bit like animals that inhabit the "mid-water" zone in the oceans, 500 meters or so deep. Many of the animals here are largely transparent, but because of the demands of physiology are not entirely so, and their visible fragments may be scattered throughout them. They look like a sort of mixture of the schematic and the concrete, which we can somehow hold in our heads at once.

This section has aimed at giving an account of model systems as they appear within the practice of model-based science, without worrying about whether this account is defensible from a metaphysical point of view. Thomson-Jones calls this the "face-value practice." To use a phrase suggested by Deena Weisberg, a view of model systems as imagined concrete things which many scientists can simultaneously investigate is the "folk ontology" of model-based science, the ontology that is implicit in the practitioners' routine behaviors. Borrowing a label from philosophical discussions of literature, this folk ontology is a kind of "fictional realism" (Brock 2002, Everett 2005). And Giere's picture is the implicit model *of* the modeler.

The idea that theorizing may involve *some* sort of deployment of fiction is far from a new one (Fine 1993). But previous discussions have often been conducted within the project of giving a philosophical analysis of all theorizing. Fictions do not have the same role in all of science. They are a particular kind of tool, and their role changes over time and space. Since WWII, model-based science has probably become more prominent, and more recognizable as a distinct strategy rather than an ingredient in a blend. (This would make some sense of the earlier tendency to see fictionalizing as either everywhere or nowhere.) Thinking and talking of model systems as imaginary *concreta* may have become more noticeable, too. This is perhaps especially due to the role played by computers. Computers have turned attention away from analytical methods to some extent. They also make it possible to model more causal detail, and are powerful tools for visualization.

9

4. A problem, with analogies

As indicated in the Introduction, an account of the practice of model-based science is one part of an overall philosophical story. The other part is an account of how this sort of work relates us to the world at large, what it achieves or may enable us to achieve. This second part of the story is told using our best general picture of what the world is like. This includes constraints that may come from general metaphysics and epistemology, where I assume that those enterprises are carried out in a naturalistic spirit.

One way to fill out the overall picture would be to work within the folk ontology of the modelers. Then we might say: model systems are fictional things which have various similarity relations to real-world systems. We learn about the former, and use that knowledge to illuminate and adapt us to the latter. From a point of view external to the practice, however, there seem to be good reasons to resist this picture. It is hard to know what counts as a too-realistic treatment of fictional entities, what counts as attributing them too much object-hood. But we might reasonably start resisting at the point where the explanation treats the model system as a shadowy additional graspable thing, either of an abstract platonist kind or a modally-realist kind as seen in David Lewis (1996). The Giere picture itself functions here as a kind of model; it captures something very accurately – the indirectness – but its similarity to the real phenomenon in other respects needs close scrutiny. The overall story is likely to be a partially deflationary one with respect to the status of fictional systems.

Another way to put the problem is as follows. Treating a fictional model system as an *object* is almost inevitable within the practice. Object-talk here seems to involve little burden, and seems to capture what we need. In the case of physically constructed models this description is completely undeniable. Then an extra object is definitely part of the story – we aim to understand one thing by investigating another – and the relation between model and target is evidently a matter of similarity with respect to contextually relevant physical properties. Model-based science of the sort that uses fictions seems quite similar to that older style of modeling. Rather than actually constructing an extra system, we merely imagine one, and do so in such a way that we can work out how it would behave. This all seems straightforward when we take a point of view inside the practice, but from a more external perspective, the introduction of extra fictional "objects" in our explanation of how this works is surely a problem.

One way to look at the situation is as a new problem posed by modality for the philosophy of science – or a new perspective on "the" problem of modality in science. For an earlier philosophy of science the problem looked like this: what we have access to and can talk about is an array of empirical particulars. How can we possibly have knowledge about the ways in which these empirical particulars are tied together in laws? The problem in this paper, in contrast, starts from the fact that one part of science spends much energy engaging in the description of elaborate fictions. But somehow this helps us get a handle on other things, including empirical particulars. The problem is understanding how we go from apparent knowledge of what seem very dubious modal facts – about how fictional systems behave – to knowledge of how things work in the real world.

So far I have looked at a two-way comparison between scientific and literary fictions. The problem can be made more vivid by adding another comparison, yielding a three-way analogy.

The situation with model-based science is analogous to one found in the philosophy of mathematics. It is sometimes said that all practicing mathematicians are platonists, a comment associated with Paul Bernays (1935). This is not suppposed to be a claim about the mathematicians' overt self-conception or ideology, but about the implicit ontology guiding their work: mathematical research is conducted as investigation of real objects in an abstract realm. Let us suppose for purposes of illustration that this is true, that an implicitly platonist outlook is a feature of successful mathematical practice – in Weisberg's terms again, that platonism is the folk ontology of research mathematics.

As before, it is a different task to look at the practice from outside, and describe what the work actually achieves. When we do this, we apply our all-things-considered ontology. *We* need not be platonists. We might be – we might be made so, perhaps grudgingly, by the role of mathematics in science (Quine 1948, 1960).[7] But we might instead give a deflationary view (Field 1980 & 1989, Yablo 2005). That means that we have to give an account of the real achievements of the field that does not use the ontology embodied in the practice. Though mathematicians are wrong to be platonists, they succeed in discovering important things nonetheless.

---

[7] "In a contest for sheer systematic utility to science, the notion of physical object still leads the field. On this score alone, therefore, one might still put a premium on explanations that appeal to physical objects and not to abstract ones, even if abstract objects be grudgingly admitted too for their efficacy elsewhere in the theory." (Quine 1960, p. 238)

A three-way comparison can then be made: models in science, mathematical objects, literary fictions. In the mathematics case, the practice is platonist. Fictionalism may be reached under external philosophical pressure, and appears as a deflationary view. Merely invoking the idea of a fiction does not take us far, however, and an antipathy to psychologism may prevent the simplest ways of analyzing fictions being available in this case. A further feature of the mathematical case is the problem of the "unreasonable effectiveness" of mathematics in dealing with the natural world (Wigner 1967).

In the literary case, fictionalism in one sense is overt; certain things are presented as "mere creations of the mind." A philosophical problem arises because the way people talk about fictional entities makes it seem that they cannot have *too* much unreality. This problem is seen especially with respect to statements that are both "inside and outside" a fiction, statements that bridge a fictional world with the real one. These are sometimes categorized as "critical" (Brock 2002) or "transfictional" (Frigg forthcoming). Here is a standard example due to Terence Parsons: "Sherlock Holmes is more famous than any real detective." In this context, "fictionalism about fictional entities" (Brock 2002, Everett 2005) appears as a philosophical view that is deflationary of the small puff of reality that fictional objects might have been accorded because of the ways we talk about them.

In the literary case, there is no problem of "unreasonable effectiveness." If a Martian came down and tried to explain everything there was to explain about our practice of fictionalizing in this context, a description of the role of narratives in the coordination and extension of human imagining could be a sufficient explanation. There is some learning about the world from literary fictions, but this is a comparatively low-powered phenomenon that does not pose special problems. The effectiveness of literary fiction in dealing with the world is a reasonable effectiveness, not an unreasonable one.

The scientific case combines features from each of these. As in the literary case, we have entities that are presented overtly as fictions. But they are handled in a way that takes seriously a distinction between true and false claims about the fiction, and in a way that involves constant similarity comparisons between fictional and real-world objects. So practice seems to embody a kind of "fictional realism." The scientific case also has something in common with the mathematics case, and that is a problem of "unreasonable effectiveness." By means of modeling we learn a great deal about how things do and can

work in the world. A description of the coordination and elaboration of imaginings cannot be a complete explanation.

I will discuss one more aspect of the three-way comparison: the role of impossibilities. Literary narratives, scientific models, and mathematical theories can all specify impossible things. Here I mean impossible in a strong sense, involving logical as opposed to physical possibility.[8] In the literary case, this fact can be used to exert constraints on the philosophical debate (Everett 2005). "Realist" attitudes to fictional objects have problems with the fact that impossible objects and scenarios seem to have the same sort of literary role as possible ones; impossibilities can apparently be *normally-functioning* fictions. That motivates a deflationary treatment of fictional entities. If we were tempted to see fictions are shadowy but graspable additional real entities, this temptation is much reduced once we see that impossibilities have the same sort of role as possibilities.[9]

Things are different in the mathematical case. Here, I take it, there is constant pressure to avoid contradiction, and most would say that a contradictory mathematical object cannot exist. There are problem cases (Colyvan forthcoming), and the issue is not entirely clear-cut, but a good argument could be made that impossible mathematical objects have a different role from others. A platonist does not have immediate problems if they deny that impossible objects are among the graspable and knowable things that mathematics studies. So the mathematical case is different from the literary one, in which the absence of a sharp difference between the role of possible and impossible scenarios tells against some kinds of realism.

The scientific case is less clear. Some have argued that impossible scenarios are often used, and fruitfully, in science (Vaihinger 1924, Fine 1993, Thomson-Jones forthcoming). Here is a simple class of examples from the fields I have drawn on in this paper. Many biological model-descriptions have the following form: "consider an infinite population of Xs, which mates/interacts/forms pairs at random...." This is a common way of talking, but it is problematic. The infinity in question is presumably a countable infinity. Kolmogorov axioms for probability do not permit a uniform distribution over a countable infinity of possibilities. So if we imagine picking one individual and asking who it is going to mate with,

---

[8] Giere (forthcoming) discusses the problem but uses a weaker, physically grounded, sense of "impossible" than this, so some entities he discusses as impossible I treat here as possible.
[9] Brock (2002) argues that some kinds of fictional realism do not have this problem: only the more "concrete" kinds, not the ones that treat fictional objects as real but abstract.

we cannot say that all members of the population have an equal chance of being the mate, in a countably infinite population. That would amount to positing a uniform distribution over the population-minus-the-individual-chosen. The modeler could say instead: consider a very large finite population. Then there is no problem, but the very-large-finite assumption does not permit exact application of some of the assumptions that people often want to use.[10]

In response, it might be argued that in some kinds of modeling, a possible scenario is specified as a first step, and then the scenario is *approximately* described with mathematics. Far from *being* the model, the mathematics does not even exactly *apply* to it. Another possibility is to say that modeling often deals with a patchwork of individually coherent imaginary fragments, and descriptions switch between these faster than is obvious. Encountered impossibilities indicate that we did not notice a switch. Other responses are no doubt available. The overall question is whether impossibilities have a sharply different role from that of genuine possibilities. That is what is probably true in the mathematics case and false in the literary case. If impossibilities are "normally-functioning" fictions in science, this would push against more realist options. If impossibilities have a very different role, this is not a problem, and perhaps a plus, for the more realist side.

5. Possible solutions

In this section I discuss some possible solutions to these problems. I will look at some approaches that are fairly close to the picture suggested by the practice of modeling and others that are more deflationary. The options also have different connections to the analogies between mathematics, scientific modeling, and literary fiction discussed in the previous section.

The aim, as noted earlier, is to develop an account that integrates with a reasonable package of views within general epistemology and metaphysics. A complication arises from the fact that quite a lot of systematic metaphysics is best seen as model-building itself.[11] This is not overt within metaphysics, but model-building is something that people can do while having a very different ideology. Modeling is a response that our minds tend to make to

---

[10]   The problem would be avoided if infinitesimal values of probability were allowed. I am grateful to Alan Háyek for assistance with this example.
[11]   This idea is discussed in more detail in Godfrey-Smith (2006b).

certain kinds of problems, and I suggest that many of the famous systems of metaphysics are best seen in these terms – as models that give us some insight into how things can be organized, but which do not have the direct application to the world that many metaphysicians envisage. Here I have in mind especially work within ontology, work aimed at describing the basic make-up of the world, including treatments of properties, modality, and abstract objects. As that list shows, this includes a lot of work relevant to the problem of understanding models in science. That does not itself create a circularity, of the bad kind. An account of how model-like structures can be used to represent the world may be applicable to parts of metaphysics itself.

I will discuss three approaches. The first is to treat fictional model systems as abstract objects of some kind. The argument might go like this. The case of mathematics shows that we have good reason to think that some abstract objects are real. And whether or not fictions in general are abstract objects, models in science might be understood in this way. We describe and come to understand model systems in the same way seen in the case of mathematical objects, and in both cases we can use that knowledge to help us deal with real-world systems.

This view retains from the face-value picture of modeling the idea that the model system is an extra entity, an object of study in its own right. The part of that picture which is modified is the account of the similarity relations between models and targets. These relations are no longer seen as ordinary similarity comparisons, in which physical properties are mapped to physical properties, because abstract objects do not have properties of the right kind. Instead there is a more abstract mapping of some kind, preserving only the formal structure of the relations between objects on each side.

Such an approach is close to one version of the "semantic view of theories" – those earlier ideas reappear here in a new role. According to this version of the semantic view, scientific theories in general specify *structures*, in the model-theoretic sense (Suppe 1977, French and Ladyman 1999). These are sets of objects with relations defined over them. The relations are understood extensionally – just as ordered *n*-tuples of objects that stand in the relation. By means of various mapping operations, tighter and looser, we can use these structures as representations of actual-world systems.

Returning to the analogies in the previous section, this would be a reduction of the scientific problem to the mathematical one. Such a story would work most simply if a

platonist attitude to structures themselves was taken. Many writers within the semantic view seem to satisfy Bernays' observation about platonism at least as well as the mathematicians he had in mind. The resulting view inherits well-known problems encountered in the literature on the semantic view concerning the nature of the required mappings (Van Fraassen 1997, Bueno and Colyvan forthcoming). The view may not apply generally to model-based science, even if those problems are overcome. It only applies when the scientists' representations can be seen as specifying a determinate structure.

The second option draws on work by Frigg (forthcoming). Frigg makes use of an influential account of literary fictions, Walton's "make-believe" account. (1990). (The same work is used in a different way by Levy, as discussed earlier.) On this approach, a theory of fiction is a theory of the coordination of imaginative acts, and the elaboration of those acts through narrative. The account is designed to avoid taking fictional objects metaphysically seriously.

In the debate about literary fictions, as noted in the previous section, there is much discussion of "transfictional" statements, which appear to speak inside and outside a fiction at once. Frigg uses the example of the character Morris Zapp from David Lodge's novel *Changing Places*. Suppose (modifying Frigg's example) someone says "Zapp is no more conceited than most academics." This statement seems to bridge fictional and real worlds. As such, it can serve as a partial model for the problem of understanding the role of fictions in science. Frigg (diverging from Walton) suggests that we analyze at least a large class of such statements as comparisons of properties. Zapp does not *instantiate* properties such as conceit, as he does not exist, but he has been attributed them by the fiction, and that suffices to put conceit – and a specific level of it – "on the table" for discussion and comparison. We can then compare the attributed level of conceit with the level instantiated by various real people. This, Frigg says, will also work for fictions in science. If I say that some actual rabbit population behaves like a population in a model, I am taking the fiction to have put certain properties on the table which can be compared to the properties of a real rabbit population.

> [T]ransfictional statements about models should be read as prefixed with a clause stating what the relevant respects of the comparison are, and this allows us to rephrase comparative sentences as comparisons between properties rather than objects..... Hence, truth conditions for transfictional statements (in the context of scientific modelling) come down to truth conditions for comparative statements between properties, which

> are unproblematic in the current context (that is, the problems that attach to them have nothing to do with issues surrounding fictional discourse). (Frigg, forthcoming)

Frigg treats non-existent objects as problematic – something to be avoided. He thinks properties, however, can be used in an explanation with less anxiety. It is not that properties are entirely unmysterious, but the problems they bring must also be confronted in philosophical contexts that have nothing to do with fiction. I am not so sure that there is such a difference between the candidates. When dealing with fictional models, many of the properties that are being introduced will be uninstantiated ones. These do seem to raise special problems of the same kind as those seen with fictional objects. In many contexts it is perfectly straightforward to talk of uninstantiated properties including very elaborate ones. You know what someone is saying if they say there are dozens of better ways this room could have been arranged. But in those same low-key contexts, it is also perfectly straightforward to talk of fictional or non-existent objects. You also know what someone is saying if they say that Sherlock Holmes would have liked this problem, and it is shame he does not exist. It is not clear that giving an explanation of modeling in terms of uninstantiated properties is more down-to-earth than giving one in terms of non-existent objects.

On Frigg's approach, what is kept intact from the picture suggested by the practice of modeling is the nature of the comparisons between model and target. There is no need to deflate the straightforward and intuitive treatment of model/target similarities discussed in section 3. The model system as an *object*, as the bearer of properties, is deflated. When we learn about the world through modeling this is a matter of learning something about how various properties are related, even though the investigation may make use of properties that are not found empirically.

This approach also gives rise to a different way of tying together the three problems discussed in the previous section. In the philosophy of mathematics, *structuralism* (or *ante rem structuralism*) holds that mathematics is the study of structures or patterns in themselves, ways in which things can be arranged (Shapiro 1983, Resnick 1981). For example, arithmetic is "not understood as the study of a particular set consisting of *the* natural numbers, but rather as the study of the natural-number-structure, the structure or pattern of any system that has an infinite sequence of objects with an initial object and a successor relation" (Shapiro 1983, p. 534). Regarding the overall status of such an account, Shapiro says (p. 538):

On this account, the problem of the relationship between mathematics and reality is a special case of the problem of the instantiation of universals. Mathematics is to reality as universal is to instantiated particular. As above, the '"universal" here refers to a pattern or structure; the "particular" refers not to an individual object, but to a system of related objects. More specifically, then, mathematics is to reality as pattern is to patterned.

And from a footnote to that passage: "At worst, [Resnik's and my] work can be seen as reducing certain problems concerning mathematics to problems concerning patterns."

Rather than an assimilation of the scientific problem to the mathematical one, as in the semantic view of theories, we have a possible assimilation of all three problems to the problem of universals. This approach would take properties and patterns to be at the bottom of a collection of modalizing and fictionalizing faculties, thereby treating mathematics, fictional models, and literary works in a unified though ontologically contentious way.

A third option is more deflationary. Perhaps all talk of fictional systems in model-based science should be seen as no more than a psychological prop for organizing talk that achieves something quite different. Modelers have the habit of expressing claims as descriptions of how fictional systems behave, but what is achieved by this talk is the investigation of deductive relationships and conditionals of various kinds, along with the development of approximately true descriptions of real systems. When I say that this is what is "achieved" in talk of model systems, I do not mean that we can analyze what modelers *mean* in those terms. Perhaps the meaning of modelers' talk is tied up in an ultimately incoherent "fictional realist" ontology. Instead, modelers achieve something useful despite themselves, in the same way that a deflationist about mathematical objects thinks that mathematicians achieve useful things despite their misguided view of their subject matter.

To say that talk of model systems is a psychologically exotic way of investigating conditionals (and the like) is not itself to solve the problem. It is natural to think that the useable output we get from modeling is generally a conditional – a claim that *if* such-and-such a configuation existed, it would behave in a certain way. The configuations in question, however, are usually known *not* to exist, so the problem of explaining the empirical usefulness of this kind of knowledge reappears. The overall aim, though, is a kind of mild "error theory" about modeling, one that avoids treating a model system as any kind of

intermediary or *tertium quid*, a view that is deflationary about both model-systems as objects of study and talk of model/world similarity.

I don't know which of these three options is most promising. Mixtures may also be feasible. In closing, I will make one further connection to the problem of mathematical objects. Historians of mathematics sometimes write of a characteristic historical sequence in which mathematical concepts *begin* life as a tools for talking about empirical or at least antecedently accepted things, and then gradually become objects of study in their own right. They take on "a life of their own" (Derbyshire 2006). This applies to numbers themselves, and also to other entities such as functions, permutations, and sets. Similarly, a model-building tradition in some area of science may *begin* with attempts to describe actual goings-on. This leads to approximation and idealization – to deliberate simplification, and a "loose fit" between descriptions and actual events. It may then happen that this fictionalizing becomes more systematic, giving rise to a tradition in which fictional objects are studied as topics in their own right. Scientists in the field get used to discussing how such systems behave, get used to talking of what is true or false of them – get used to treating a fictional model system as an object in itself. An outsider might come along and re-interpret this talk as a collection of claims about deductive relations, conditionals, and so on, giving a deflationary account of model systems as objects. But inside the field, it does not seem like this; there it feels like the fictional model system has a life of its own.

<p style="text-align:center">*　　*　　*</p>

References
Axelrod, R. (1984). *The Evolution of Cooperation.* New York: Basic Books.

Bernays, P. (1935). "On Platonism in Mathematics," in P. Benacerraf and H. Putnam (eds.) *Philosophy of Mathematics: Selected Readings.* Cambridge: Cambridge University Press, 2nd edition, pp. 258-271.

Brock, S. (2002). "Fictionalism about Fictional Characters," *Nous* 36: 1-21.

Bueno, O. and M. Colyvan (forthcoming). "An Inferential Conception of the Application of Mathematics."

Colyvan, M. (forthcoming). "The Ontological Commitments of Inconsistent Theories." *Philosophical Studies.*

Crow, J. F. (1986). *Basic Concepts in Population, Quantitative, and Evolutionary Genetics*. New York: Freeman.

Darwin, C. (1959). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Derbyshire, J. (2006). *Unknown Quantity: A Real and Imaginary History of Algebra*. Washington DC: Joseph Henry Press

Everett, A. (2005). "Against Fictional Realism," *Journal of Philosophy* 102: 624-649

Field, H. (1980). *Science Without Numbers: A Defence of Nominalism*. Princeton: Princeton University Press.

Field, H. (1989). *Realism, Mathematics, and Modality*. Oxford: Blackwell.

Fine, A. (1993). "Fictionalism," *Midwest Studies in Philosophy* 18: 1-18.

French. S. and J. Ladyman, (1999). "Reinflating the Semantic Approach," *International Studies in the Philosophy of Science* 13: 103-121.

Frigg, R. (forthcoming). "Models and Fiction." To appear in *Synthese*.

Giere, R. (1988). *Explaining Science: A Cognitive Approach*. Chicago: Chicago University Press.

Giere, R. (forthcoming). "Why Scientific Models are not Works of Fiction." To appear in M. Suárez (ed.)*, Fictions in Science: Philosophical Essays on Modeling and Idealization.*

Godfrey-Smith, P. (2006a). "The Strategy of Model-Based Science," *Biology and Philosophy* 21: 725–740.

Godfrey-Smith (2006b). "Theories and Models in Metaphysics." *Harvard Review of Philosophy* 14: 4-19.

Goodman, N. (1972). "Seven Strictures on Similarity." In *Problems and Projects*. New York: Bobbs-Merrill, pp. 437-447.

Klein, C. (forthcoming). "Idealization is Simplification."

Levins, R. (1966). "The Strategy of Model-Building in Population Biology." *American Scientist* 54: 421–31

Levy, A. (forthcoming). "Idealization as Fiction."

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.

Lloyd, E. (1988). *The Structure and Confirmation of Evolutionary Theory*. Boulder: Greenwood Press.

Maynard Smith, J. (1998). *Evolutionary Genetics*. (2nd edition). Oxford: Oxford University Press.

Quine, W. V. O. (1948). "On What There is," reprinted in his *From a Logical Point of View*. Cambridge MA: Harvard University Press, 1953, pp. 1-19.

Quine. W. V. O. (1960). *Word and Object*. Cambridge MA: MIT Press.

Resnick, M. (1981). "Mathematics as a science of patterns: Ontology and reference", *Nous* 15: 529-550.

Roughgarden, J. (1979). *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. New York, NY: MacMillan.

Shapiro, S. (1983). "Mathematics and Reality." *Philosophy of Science* 50: 523-548.

Sterrett, S. G. (2002), "Physical Models and Fundamental Laws: Using One Piece of the World to Tell About Another", *Mind and Society* 5: 51-66.

Sterrett, S. G. (2006) 'Models of Machines and Models of Phenomena', *International Studies in the Philosophy of Science* 20: 69 - 80

Suárez, M. (2003). "Scientific Representation: Against Similarity and Isomorphism," *International Studies in the Philosophy of Science* 17: 225-244.

Suppe, F. (ed.) (1977) *The Structure of Scientific Theories*. 2nd ed. Chicago: University of Illinois Press

Thomson-Jones, M. (forthcoming). "Missing Systems and the Face-Value Practice." To appear in *Synthese*.

Vaihinger, H. (1924). *The Philosophy of 'As-If'*. (C. K. Ogden, trans.). New York: Harcourt Brace.

Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press.

Van Fraassen, B. (1997). "Structure and Perspective: Philosophical Perplexity and Paradox", in M. L. Dalla Chiara et al. (eds.) *Logic and Scientific Methods*. Dordrecht: Kluwer, 1997, pp. 511-530.

Walton, K. (1990). *Mimesis as Make-Believe*: *On the Foundations of the Representational Arts*. Cambridge MA: Harvard University Press.

Walton, K. (1993). "Metaphor and Prop Oriented Make-Believe," *European Journal of Philosophy* 1: 39-57.

Weisberg, M. (2004). "Qualitative Theory and Chemical Explanation." *Philosophy of Science* 71,

1071–1081.

Weisberg, M. (2007). "Who is a Modeler?", *British Journal for Philosophy of Science* 58, 207–233.

Weisberg, M. (forthcoming). "Models of Modeling."

Wigner, E. (1967). "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," in his *Symmetries and Reflections: Scientific Essays*. Bloomington: Indiana University Press, pp. 222–237.

Yablo, S. (2005). "The Myth of the Seven." In M. Kalderon (ed.), *Fictionalism in Metaphysics*. New York: Oxford University Press, pp. 88-115.