Models as information carrying artifacts

Otto Lappi

Cognitive Science Unit
Department of Psychology
University of Helsinki

Anna-Mari Rusanen

Philosophy of Science Group
Department of Philosophy
University of Helsinki

1. Introduction

In science, models are used in many different ways: to test empirical hypotheses, to help in theory formation, to visualize data, and so on. Scientists construct and study the behavior of models, and compare this to observed behavior of a target system. *Modeling* involves indirect analysis of the real world via mediation of the models (Weisberg, 2007a).

Our premise is that for this to be possible models must *carry information* about their targets. When models are viewed as information carrying entities, this property can be used as a foundation for a representational theory of models. This account presents a way of avoiding the need to refer to modelers' intentions (or their mental states) as constitutive of the semantics of scientific representations. Moreover, we will show that an information theory based account of scientific representations can provide a naturalistic account of models which can deal the problems of asymmetry, relevance and circularity that afflict currently popular proposals based on user intentions.

Traditionally, there has been a strong tendency towards a clear-cut division of labor between philosophers of science and philosophers of mind. We believe that there are some important philosophical insights about representation that are relevant for both camps. For instance, the similarity, isomorphism and/or resemblance-based theories have difficulties in accounting for the asymmetry of representations (Cummins, 1989; Fodor, 1992). This problem has prompted the development of information-based semantics in the naturalistic philosophy of mind (Eliasmith, 2005; Usher, 2001) but is also familiar to philosophers of science (e.g. the discussion concerning scientific representations based on Suárez, 2003).

From the information semantic perspective, models as scientific representations can be considered a special case of a larger problem of naturalistic representation. In this paper we will look at what we think is the most promising avenue of developing this *information theoretic account of representational models*.

2. Models as representations

While it is widely recognized that models play a significant role in science, there remains a disagreement over how, and even whether, models *represent* their targets. One reason for this is that some philosophers find the whole concept of representation dubious, and attempts to sharpen the definition of representational relationships ambiguous, circular, or unsatisfactory for other reasons. For many, these suspicions have been reason enough to suggest giving up the attempts to say anything substantive about scientific models as "representations".

Some (e.g. Knuuttila, 2005) have suggested that it would be profitable to view models not primarily as representations, but as "epistemic artifacts" with other tasks (such as predicting experimental results, making explicit theoretical ideas, or unifying scientific knowledge). Others, such as Suárez (2004), have on the other hand argued for a "minimalist account of representation", according to which we should not try to define scientific representation naturalistically, on the basis of the properties of the models and their relation to the world, but instead on the pragmatics of the ways that competent users use the models as representations. The common wisdom seems to be that there can be no workable user independent account of the representational character of models (see e.g. Teller, 2001).

However, we think it is premature to give up the idea of the intrinsic representational character of models. It *is* possible to construct an autonomous theory of this representational character (i.e. one not derived from a prior intentional characterization of the models' users).

First we need to make clear what kinds of models we are here concerned with. The target of our discussion is models that are *scientific representations constructed in order to inform us about some aspects of nature*[1]. Models are public, man-made artifacts. They are not abstract entities (Giere, 1988), nor thoughts or other mental representations (Mäki, 2009). Models can still be abstract – e.g. mathematical or computational models - or concrete, such as Watson & Crick's physical scale model of the DNA molecule. The fully abstract ("metalogical") sense of models as set-theoretic structures satisfying a set of axioms is not included in the target of our analysis. Also, symbolic representation of some purely conceptual (mathematical or computational) structure is not included in our definition of "model": instead we call these purely symbolic structures only *templates* for models[2]. In what follows we will demonstrate that by applying the information theoretic account of representation into the discussion about models as scientific representations one may construe a coherent and illuminating naturalistic view of scientific representations.

[1] See also Jones (unpublished) for a useful taxonomy of models as *truth making maps*, *truth making structures*, *mathematical models*, *propositional models* and *physical models*. Jones' first two notions of models include an abstract notion of a model as providing an interpretation for a certain set of sentences so that those sentences come out true, with no reference to empirical data or natural phenomena - this is not the sense of model we are concerned here. Models in the third, fourth or fifth sense can do the job of representing an actual or merely possible system. It is models of these representational sorts that are our target of interest.

[2] It is important for the view developed here that a model can be decomposed into an abstract *template* (the "mathematical skeleton" of the model) and a *domain* which carries the information, and the empirical commitments, of the model and is the locus of its truth (or falsity). There are other examples of *decompositional* view of models; for instance Mäki has developed and defended a decompositional account of models (2008, 2009a, 2009b). Cf. also Humphreys (2002, 2004). Sometimes in the literature what we call the template of the model is already considered a "model". We want to argue that embedding of the template in a domain is an *essential* feature of models. There is one more sense of the term model in science: a system that is simple and can be more easily investigated can stand in for a larger class of systems (for example the fruit fly can be used as a model of inheritance and genetic regulation of development, or the mouse can be used as a model for human cancer or responses to anti-inflammatory drugs). While the model organisms are clearly not manmade artifacts, we believe this sense of "modeling" can be subsumed under our account. Suppose one wanted to use the mouse as a model for human cancer and human physiological responses to various treatments. One could not merely get a hold of cancer prone strains of mice, subject them to various carcinogenic conditions and treatments, and observe the outcomes. Instead, painstaking care, and a great deal of empirical research must be undertaken to establish correspondences between the mouse and the human "in the relevant respects". This analogy is part of the *domain* of the model. Many established or unstated assumptions about biophysics and the instrumentation are taken for granted, and certain peculiarities of mouse (or human) physiology might cause the scientist to treat some variables as "not correlated", meaning that observations of their values would not occasion inference from one organism to the other (these variables would effectively be part of the *template*).

2.1. Characterizing the conditions of the representational relation between models and their targets

The relation required to establish representationality can be schematically put as follows:

A *represents* B *if and only if* C.

Different philosophical construals of the condition C lead to different conceptions of the representational relationship. Many of the current alternatives are variations of the following:

(1) The derived intentionality conception: A *represents* B *if and only if it is so interpreted by intentional agents*.

One way to approach the representational relationship is to say that the models represent whatever the scientists themselves postulate the model to represents, or *intend* the model to represent. For instance, Teller (2001, p.397) writes "I take the stand that, in principle, anything can be a model, and that what makes a thing a model is the fact that it is regarded or used as a representation of something by the model users. Thus in saying what a model is the weight is shifted to the problem of understanding the nature of [mental] representation." Suárez (2004) offers an account of representation which is based on the view that models represent their target systems in virtue of their capacity to lead a "competent and informed user to a consideration of the target". The idea is that the intentionality of the mental systems of the models' users creates thus the semantic relationship between a model and the world.

It may be appealing to view scientific models as only based on "derived intentionality" (Searle, 1992). This will make accounts of scientific representations dependent on prior intentional characterization of the users, and on empirical facts about how scientists interpret their models.

However, this is a problematic view on several counts. First, empirically, the interpretational practices of scientists are complicated and not at all well understood. Second, the issues become unnecessarily complicated if all the "pragmatic constraints" of the modelers' interpretational activities are taken as constitutive of model semantics. The issue immediately arises: *which* practices in modeling are constitutive of the semantics of models as representations? It is here that the information semantic viewpoint can be useful. Merely postulating a representational relation between a model and an intended target does not necessarily make it so. Intentionality does not magically create a representational state of affairs between a model and its target. The modeling practices of scientists must involve *more* than good intentions, or merely talk. It must involve establishing an informational connection between the model and its target.

As a significant body of literature in the philosophy of mind indicates (REFS), it is possible to build an account of representation that naturalizes representation *directly* based on this informational connection. This is importantly different from the solution that naturalistically oriented philosophers of science have tended take. Philosophers of science usually see representation to be some kind of *similarity* relation, some degree structural "fit", between the model and some aspect of the world. This relationship of "fit" between the world and the model has been characterized as "isomorphism", "partial isomorphism", or "resemblance" (French, 2002; Maki XXXX, XXXX), and so on.

This *structuralist conception of models* (Frigg, 2006) conceives condition C as follows:

(2a) The similarity conception: *A represents B if and only if A is similar to B*.

The similarity conception is problematic on two grounds. First, similarity is always similarity in some *relevant* respect, threatening to import user intentions back into the frame, insofar as we would use the users' intentions to *identify* the "relevant" respects. Also, similarity is a vague notion in need of more rigorous characterization. Now, let´s take a look at (2b), which is a more precise version of (2a):

(2b) The isomorphism conception*: A represents B if and only if the structure exemplified by A is isomorphic to the structure exemplified by B*.

The isomorphism conception clarifies the notion of similarity, but still leaves open some important problems. A representation does not need, and it cannot be perfectly "similar", "isomorphic" or "identical" with the target system in all respects. The motivation for the whole debate is the observation that scientific models are typically known to be inaccurate. Almost *any* target system is too complex, and abstraction of details is used to reduce the degree of complexity, and counterfactual assumptions are put in place in order to create an idealized, but workable model. All this makes the idea of deriving representation from resemblance problematic[3].

It is an inescapable feature of scientific model building that it is not usually (or perhaps ever) possible to construct a fully comprehensive model of how a target system works in all its detail. What is important is that the model should be "similar enough" or "sufficiently isomorphic" to the target in the *relevant* respects. Now the problem is: how to characterize "sufficiency" of similarity and "relevance" of different aspects in the model? We need some way to identify the relevant structures in the model and in the world, whose isomorphism we are then to consider[4].

Without such constraints on what aspect of the model and of the world count as the relevant relata, it might often be the case that a model we intend to represent B in fact is more similar, isomorphic, to some completely arbitrary collection of entities X. Unless we have some constraints on what can count as "an object" for modeling, we could always construct *arbitrary mappings* between parts of the model and whatever our fancy dictates. Then the model will in fact be a model of this arbitrary X (which it is of course guaranteed to be isomorphic to and hence true of)[5]. This seems unacceptable. We will call this the *problem of relevance*.

---

[3] One reason for some philosophers' rejecting the structuralist account of models is that models are typically *abstract* (lacking features *known to be present* in the intended target), *idealized* (incorporating assumptions that are *counterfactual*, i.e. *known to be false* about the intended target), and *simplified* (representing only a few dependencies from among a multitude). This has led some philosophers to ask whether the view of models as representational makes any sense: if models are known to be false, how can models represent if they don´t represent faithfully, but instead miss out or distort many known features of the real world? If models misrepresent, or represent inaccurately, how can they offer knowledge? Insofar as idealization is taken to require the assertion of falsehood (e.g. Jones, 2005), idealization makes the models false descriptions of their target systems. However, it is important to make a distinction between the conditions for A to be a representation of B, and the conditions for A to be an accurate or a *true* representation of B. After all, A can only be false about B if A is *about* B (a similar approach can be found for example in Callender & Cohen, 2005).

[4] The received way is by referring to the intentions of the modelers. This would amount to making (2a) and (2b) effectively a variant of (1).

[5] A (re)interpretation of Putnam's (1988, pp.121-125) arguments about computation (possibly due to influence by Searle, 1992, ch.9) says that by stipulation anything can be *interpreted* to represent anything. We believe that genuine, information carrying representations differ from mere stipulations, since they allow us to have information

Also, if truth as well as reference is defined in terms of similarity the model *cannot* possibly radically misrepresent or be completely false about its target, as the target is, *by definition*, whatever the model is also true of. We will call this the *circularity problem*. Also, similarity alone does not seem to fix the reference of models correctly. If, say, a model represents anything and everything in the world that is *similar* to the model, then, for example, multiple copies of a model (e.g. scale models of DNA in various museums and exhibits) would thereby represent each other. Structuralistic accounts have been disputed on the logical grounds that isomorphism is a symmetric and reflexive relation, but representation is not (for instance, Suàrez, 2003). This problem is known as the *asymmetry* or *directionality problem*. Putting it briefly, an isomorphism or similarity relation between any systems – *a fortiori* a model and its target system - must be symmetrical, reflexive and transitive (Cummins, 1989; Fodor, 1990; Suárez, 2003). The representation relation as commonly understood is none of these things because of its directionality (targets do not represent their models, models represent their targets, and if the target system B of model A is itself a model of some S, the model A does not thereby represent S). But, as Suarèz also admits, this is problematic mostly - and perhaps only - for cases that ground the representation relation on similarity and isomorphism - not naturalizing the representation relation *per se* (Suárez, 2003).

Because of these problems, and especially because of the problem of relevance, many philosophers have found it necessary to invoke the intentions, or intentional "use" that of the modelers. Consider, for example, Mäki (2009b): "Agent A uses object M (the model) as a representative of target system R for purpose P; addressing audience E; at least potentially prompting genuine issues of resemblance between M and R to arise; describing M and drawing inferences about M and R in terms of one or more model descriptions D; and applies commentary C to identify the above elements and to align them with one another… I join those, such as Giere, who have emphasized the importance of purposes and intentionality in the notion of model as representation. The relationship so conceived has the form: A uses M to represent R for purpose P (Giere 2006, 60). So for an object to represent at all, an agent's intentionality is required."[6]

However, if one does not want to commit herself to these intentional, pragmatics-based accounts, there are other ways for the naturalist to pursue. In the philosophy of mind, information based accounts have largely superseded the isomorphism view in the naturalistic analysis of representation (see e.g. Cummins, 1989), where the most debated philosophical topic has revolved around the question of naturalizing the semantics of representation (Dretske, 1981; Millikan, 1989; Fodor, 1992; Usher, 2001).

A naturalist *could address the problems of* (1) *and* (2) *in one move* by using information-theory to characterize constraint C. Consider first,

(3) The statistical information conception: *A represents B if and only if A carries information about B.*

about the intrinsic properties of target systems that we would not be able to have on the basis of arbitrary representations. In this sense an information theoretic account is not only a descriptive, but also a normative theory for representations: It gives the criterion for distinguishing a "genuine" representation from arbitrary mappings.

[6] Note that the choice of capital letters is different from the notation in this paper.

The notion of information used here is *statistical information* or “Shannon information” (Shannon, 1948)[7]. Causal-informational theories of semantics (Dretske 1981) hold that the content of a mental representation is grounded in the information it carries about what does (Devitt 1996) or would (Fodor 1987, 1990) cause it to occur. The representational connection between A and B is provided by a causal-informational relationship between the representation and the things in the world. The statistical theories of information are used to make a use of the powerful concepts of probability theory, which provides exact statistical concepts with which to define the reference of representations (Usher, 2001). There are many advantages to this approach compared to qualitative discussions of “pragmatics” of modeling, among them increased conceptual precision and the opportunity to define semantics of scientific representations directly, without reference to prior intentionality of the users’ intentions. There are several variants of this account[8].

But (3), as it stands, is still too weak since A may carry information about a lot of things. In the context of scientific representations, it is crucial that A carries information about "relevant" or "interesting" aspects of B. The problem is how to define this information constraint further. One is really interested just in its carrying information about B. There are several attempts to solve this problem of relevance[9].

Consider next,

(4a) The reliable information conception: *A represents B if and only if there is a reliable information processing mechanism M that supports the information connection between A and B.*

The character of M is discussed in epistemology under reliabilism (Goldman, 1986). In the philosophy of mind in Fodor's (1992) information semantics and Ryder’s (2004) account of mental representation in terms of the mind/brain as a Model Making Mechanisms have similar features. In the present context, the model building process implements this reliable connection. This

---

[7]It is important to realize that there are at least two notions of "information" which need to be kept distinct. One is a notion of statistical information, the other we call structural information. Supplied with suitable conditions for "interpretation", both can be and have been used as a basis for a notion of "semantic information". Statistical information is "physical" in the sense that we can think about physical phenomena carrying information about other physical phenomena in this sense. This is the one most often used in what is often called information theory, based on Shannon's Mathematical Theory of Communication. It is also the notion of information semantics, following Dretske's (1981), would reduce semantics to. The other notion of information, structural information is based on the idea of coding propositions into symbol structures. The idea is that a material symbol system can assume different configurations which have formal structure (“syntax”). The information content of these structures can be defined quantitatively (e.g. algorithmic information), and we can assign interpretations to structures (e.g. based on their isomorhisms with "models" for these structures.). This notion of information derives from symbolic logic (Leibniz, Boole, Frege, Turing).

[8] See for example, Dretske, 1981; Fodor, 1992; Millikan, 1993.

[9] Marius Usher´s statistical reference theory is one very sophisticated example of those theories (Usher, 2001). The basic idea of it is that when a representation is tokened, the information it carries is about the class of items it carries the most information about, and not about what caused it in a singular case. Usher offers a very technical argument that uses the notion of “mutual information” for dealing this problem. According to Usher A represents B if A carries information about B and for any C that A carries information about, this information is lower than for B. (See Usher, 2001 for details).

process includes conceptual development, hypothesis formation, experimental methods, data analysis and hypothesis testing procedures. In this sense models are representational, not because of the intentions of the model builders *per se*, but because of the information theoretic properties of the model building process. This model process may of course be *directed* by the intentions of the modelers, but the key difference to the derived intentionality accounts is that these intentions do not enter into the definition of the relation itself.

Finally, it is of importance that the information relation be supported by an iterative, self correcting, process of interaction with the phenomenon (data gathering and model fitting). This is required to ensure non-accidental convergence between the world and the structure or behavior of (some parts of) the model. This idea can be formulated as follows:

(4b) Iterative incorporation of information conception: *A represents B if and only if there is a reliable iterative data gathering and hypothesis testing method M that supports the connection between some aspects F of A, and some aspects X of B.*

The representational character of scientific models are a product of, and defined by, the iterative model building process, in which information about a phenomenon (coming through from the empirical data) is incorporated into the model[10]. The "relevant" aspects that are represented then become just those aspects that the model building process ends up *tracking* – regardless of whether or not these properties are the ones the model builders hope, intend or believe the model to be tracking. We suggest that this is the most profitable way to understand the representational character of scientific models in a naturalistic way.

A scientific model may be decomposed into two kinds of elements: a *template*, an abstract part that is not representational, and a *domain* which is representational and provides the "interpretation" of the model. This distinction is based on Humphreys' (2002, 2004) view of scientific knowledge, understood in terms of computational models, but we suggest it can be extended to other kinds of models, too. The idealized core, or *template,* is a formal structure that can be considered in isolation of its use in modeling particular target phenomena. It is a set of constraints, a model schema, around which a model is built. (It may be borrowed from a different – perhaps quite unrelated – field, or derived from pure mathematics). The template must be complemented with a *domain*, in order to come up with a model of a target. The domain is that part of the model that is subjected to empirical testing, and therefore provides the informational connection to the world, and therefore, on the present account, gives the model its representational character. (The very same template may therefore be connected to domains that relate to completely different targets, i.e. the same basic structure may turn out to be useful in models that are semantically quite unrelated).

One may think of the process of model construction in the following way: A model is first constructed by setting up a template, T (with the idea of borrowing organizing ideas from existing theory, or perhaps coming up with a mathematical or computational structure from scratch), and a domain D (with an intended interpretation in mind). This distinction gives the scientists a workable principled distinction as to which parts of the model are the ones to be revised or refined, as dictated by empirical data and which on the other hand are "not negotiable". *Only* some aspects of the model, D, are thus fitted to data, and the model makers are very concerned

[10] This could be e.g. Bayesian reasoning, but Bayesian models themselves are likely to turn out to be highly idealized, defining a limit of rationality that the actual mechanisms implemented in real human endeavors only approximate.

about getting *them* right. Others, T, are treated as *constraining idealizing assumptions* that enable the modeling process to proceed[11]. It seems to us that this kind of calibration and data fitting *of only some aspects of the model* – with accompanying idealization abstraction and approximation - are typical features of model building. When the models are built and reported, the scientists usually are signaling that they are making some very *specific* idealized assumptions, and that only some parts of the model commit them to empirical predictions.

Facing the three problems

The asymmetry problem. The information semantic view can deal this is easily. The representational relation is defined as a directional relation, and the information gathering means required connects the models to the world. To deal with asymmetry, it is important that observation is a *causal* and hence it is a directional process. However, informational connectedness is a *statistical*, not merely causal relation, and this is required to define the semantics of a model. Therefore this view should not be equated with a causal theory of reference (e.g. Kripke, 1980), where a proper name refers to whatever (token) occasioned the original use of the name. Scientific representations are not proper names, but universals describing the type structure of the world, and the statistical properties of the information gathering method that fixes the reference of models, *not* just the causal history of model making.

The problem of circularity. If truth as well as reference is *both* defined in terms of similarity, a model cannot possibly radically misrepresent or be completely false about its target. This is because target is fixed, by definition, to be whatever there is in the world that resembles the model, and what the model is hence also true of. However, one of the key theoretical developments in information semantics has been precisely to disentangle reference from truth. Reference X of model element D is defined information semantically as statistically the type of X for which mutual information between the referent and the model (or: domain element) is maximized. Now, factors such as observational noise or observer bias may well lead to situations where the actual target (from which information is being extracted) *does not* correspond to the referent, making the model *false* about the target. (See Usher, 2001).

The problem of relevance. Similarity based views face the problem that a model might resemble many things which, intuitively, we would not consider to be among the model's targets. Constraints on the arbitrary ways that a model and some system might resemble each other need to be put in place, since what is important for the modelers, and for assessing the semantics and/or truth of the model is that the model and the target should be similar in the *relevant* respects.

One solution would be to refer back to the modelers' intentions to use the model A (or the template T) as a model for B (rather than some X it just happens to resemble). However, we feel there are good reasons to question this. One motivation for the similarity view is that a model can be useful for an indirect analysis of the world only if it is structurally or dynamically similar to its target. However, the information theoretic perspective accommodates similarity quite naturally. The information semantic account requires that there be a reliable information processing mechanism M that supports the information connection between A and B, the model building

[11] In Kuhnian terms, these assumptions are part of the paradigm, and are not questioned in the process of normal-science problem solving which concentrates on the information gathering process of fitting D to the target and are only revised in a "scientific revolution".

process, which ensures *non-accidental convergence* between parts of the model (the domain) and parts of the world. That is, for the model to be useful as a stand-in for the world, the "similarity" must be empirically validated, and this is (along with template construction) is an essential part of the model building process. It is the only sound reason we have to trust the results.

On our view, the semantically relevant respects, on which similarity counts, are the parts of the model that are subjected to empirical confirmation. A consequence of this is that one cannot always identify the relevant (and on the present view semantics-constituting) parts of the model simply by probing the scientists intuition, i.e. asking the scientists to identify them - although often this might be the most reliable method.). This is because *semantic relevance* is used here in a special technical sense. An individual scientist might consider his or her pet template as "the most relevant" part of the model, or might consider the parts that are required to make the model cohere with his or her preconceived world view or general metaphysics as the parts that it is most important to get right for the model to come out true. But nature does not care for our preconceived views. On the information theoretic perspective She is the one, who has the final word.

Discussion

According to Frigg (2006), there are three desiderata that a theory of scientific representations should meet. First, the theory should tell us *what kinds of objects* models as scientific representations are. This is *the ontological problem.* Second it should tell us *in virtue of what* a model is a representation of something else (and what it is a representation of). This we will call *the semantic problem*. Finally, the theory should address the plurality of *different kinds of models* (mathematical models, scale models, photographs, analogies…), taxonomize them, and tell us the ways that they represent reality. This we call *the taxonomical problem*.

According to the information theoretical account the representational nature of models is based their nature as models are information carrying artifacts (the ontological problem). They are not purely abstract, mental representations or naturally evolved phenomena. They represent their targets in virtue of carrying information about them (the semantic problem). We do not believe that the taxonomical problem can be addressed only on the basis of semantic considerations. However, in general one might say that there are many kinds of models in science, and hence many philosophical uses for the term model as well. One usage defines model as ideal or abstract structures satisfying a set of axioms or mathematical assumptions. Having settled the ontological question by defining models as concrete artifacts, this is clearly precluded. There are highly abstract model structures, with no specific "intended targets" (e.g. cellular automata). These are what we call model *templates*, but they are not fully models, until they have an information gathering means and associated observational elements (a *domain*) associated with them. Also, it follows from our view that such "models" do not have fixed semantics, as semantics is on our account based on the *informational relations* between the model and the world. However, it must be added that a template certainly has information structure and thus can be said to encode or carry *structural information*. The information relation we have used to define information is the notion of *statistical information* (Shannon, 1948), which is a notion of information that is importantly different from the "structural information". Models thus have structural information (in e.g. their template) with statistical information embedded in them. (E.g. the template may have places for various variables whose values will depend on observation, these variables would be part of the domain).

We have presented an information theoretic view of models as scientific representations, where models are understood as information carrying artifacts. These are constructed in order to be able to represent and to indirectly study properties of real world phenomena. We have suggested that the semantics of models should be traced to the information coupling of the model to the world, rather than the intentions and interpretations of the models' users.

We have suggested that the view that parts of models carry statistical information about parts of the world can be used to counter the antinaturalistic critiques, and develop a detailed account of model building and representation with the added benefit of direct relations to parallel work in the philosophy of mind. From this perspective, a crucial aspect of models, or at least precisely definable parts of them, is that they *carry information* about the properties of their targets. When models are viewed as information carrying entities, this property of models can be used as a foundation for a representational theory analogous to information-theoretic naturalization of representation in the philosophy of mind.

Of course, there are many problems left open by an information theoretic account (based on 4b, above). For example, it is not trivial to work out the details about which aspects X of B a model making mechanism M makes the product, A, to represent and which not. However, these problems are not insurmountable, and less dramatic than the ones that, for example, the similarity view has to confront. What is more, many of the problems are strictly analogous to the problems that crop up in information semantics in the computational philosophy of mind, and have been extensively discussed there since the 80's (Dretske, 1981; Fodor, 1992; Millikan, 1998), with significant recent developments (Eliasmith, 2005; Usher 2001). Joining forces with philosophers of mind and cognitive scientists on these matters might offer ways for philosophers of science make advance, but also potentially to make useful contributions to the discussions in philosophy of mind and cognitive science.

Acknowledgements

We thank the participants of POS- meetings (especially Uskali Mäki, Tomi Kokkonen, Samuli Pöyhönen, Jani Raerinne, Matti Sintonen and Petri Ylikoski) and Tarja Knuuttila for many fruitful comments on an earlier draft of this paper.

References:

Arroyo-Santos, A. & De Donato-Rodríguez, X. (2008). "*Idealization and the Structure of Theories in Biology"*. In [2008] Philosophy of Science Assoc. 21st Biennial Mtg (Pittsburgh, PA): PSA 2008 Contributed Papers.

Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, MA: MIT Press.

Dretske, F. (1981). *Knowledge and the Flow of Information.* Cambridge, Massachusetts: MIT Press.

Eliasmith, C. (2005). "Neurosemantics and Categories". In H. Cohen & C. Lefebre (Eds.), *Handbook of Categorization in Cognitive Science*: Elsevier Science.

Fodor, J.A. (1992). *A Theory of Content and Other Essays*. MIT Press, Cambridge, MA.

French, S. (2003), "A model-theoretic account of representation (or, I don't know much about art…but I know it involves isomorphism)", *Philosophy of Science* 70: 1472–1483.

Gentner, D., & Stevens, A. N. (Eds.). (1983). *Mental Models*. HIllsdale, NJ: Erlbaum.

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

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

Goldman, A.I. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

Humphreys, P. (2002). “Computational Models”. *Phillosophy of Science*, 69, S1-S11.
Humphreys, P. (2004). *Extending Ourselves - Computational Science, Empiricism and the Scientific Method*. Oxford: Oxford University Press.
Jones, M.R. (2005). “Idealization and Abstraction: A Framework” in Jones, M. & Cartwright, N. (eds). *Idealization XII: Correcting the Model--Idealization and Abstraction in the Science*s (Poznañ Studies in the Philosophy of the Sciences and the Humanities 86) Amsterdam/New York, NY, 2005.
Jones, M.R. (Unpublished). “Models, the Semantic View, and Scientific Representation".
Knuuttila, T. (2005). *Models as Epistemic Artefacts: Towards a Non-Representationalist Account of Scientific Representation*. Philosophical Studies from the University of Helsinki. Department of Philosophy, University of Helsinki, Helsinki. Doctoral Dissertation.
Kripke, S. (1980). Naming and Necessity. Cambridge, Mass.: Harvard University Press.
Mäki, U. (2009a). “Models and the locus of their truth”. Forthcoming in *Synthese*.
Mäki, U. (2009b). “Models and truth. The functional decomposition approach”. Forthcoming in
European Philosophy of Science 2007. Springer 2009
Millikan, R. (1989). “Biosemantics”. *The Journal of Philosophy*, 86: 281–97.
Putnam, H. (1988). *Representation and Reality*. Cambridge, MA: MIT Press.
Ryder, D. (2004). “SINBAD Neurosemantics: A Theory of Mental Representation”*. Mind & Language*, 19, 211-241.
Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
Shannon, C. (1948). “A Mathematical Theory of Communication”. *Bell System Technical Journal*, 27, 379-423, 623-656.
Suárez, M. (2003). “Scientific Representation: Against Similarity and Isomorphism”, *International Studies in the Philosophy of Science*, 17, 3, October 2003, pp. 225-244.
Suárez, M. (2004). “An inferential account of Scientific Representation”. *Philosophy of Science*, 71:767-779.
Teller, P. (2001): “Twilight of a Perfect Model”. *Erkenntnits* 55:393-415.
Usher, M. (2001). “A Statistical Referential Theory of Content: Using Information Theory to Account for Misrepresentation”. *Mind & Language*, 16, 311-334.
Weisberg, M. (2007a). “Who is a Modeler?”. *British Journal for Philosophy of Science*, 58, 207–233.
Weisberg, M. (2007b). “Three Kinds of Idealization”. *The Journal of Philosophy*, 104 (12) 639–59.
Vosniadou, S., & Brewer, W. (1992). “Mental Models of the Earth: A Study of Conceptual Change in Childhood”. *Cognitive Psychology*, 24, 535-585.