Models, Information and Meaning

Marc Artiga*
Universitat de València

Abstract

There has recently been an explosion of formal models of signalling, which have been developed to learn about different aspects of meaning. This paper discusses whether that success can also be used to provide an original naturalistic theory of meaning in terms of information or some related notion. In particular, it argues that, although these models can teach us a lot about different aspects of content, at the moment they fail to support the idea that meaning just is some kind of information. As an alternative, I suggest a more modest approach to the relationship between informational notions used in models and semantic properties in the natural world.

1 Introduction

The idea of explaining the origin and development of meaning in terms of some kind of evolutionary process has been popular for some time. Nonetheless, it has been difficult to spell this idea out in detail. Partly for this reason, many have welcomed the use of game-theoretic models, which has already provided more precise definitions, new arguments and suggestive ideas that have greatly enriched the debate and contributed to a better understanding of this phenomenon. Some striking results concern the emergence of meaning (Skyrms, 2010a; Hutteger, 2007a, 2007b), the evolution of perceptual categories (O'Connor, 2014), concepts (Barrett, 2014) or moral norms (Harms and Skyrms, 2008), among others.

*Paper forthcoming in *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*

A clear virtue of game-theoretic models of signalling is that they provide a simplified and indirect representation of much more complex real-world dynamics, which makes it possible for the theorist to give precise definitions and analyse specific questions. In particular, current research usually quantifies the information (or some other correlational measure) between variables in the model in order to learn about meaningful relations in the natural world. This fact has encouraged some people to defend that a naturalistic analysis of meaning can be provided in terms of information or related concepts. This is the main claim I would like to assess in this paper. More precisely, I would like to discuss the relationship between semantic notions that are defined in models and meaning.

The paper has three main parts. Section 2 frames the discussion, puts forward a key assumption ('MODEL-INDEPENDENCE'), and briefly explains some notions of information that have been put forward in the literature. Section 3 discusses the main target of the paper, the 'immodest view', according to which the relative success of game-theoretic approaches can support an informational theory of meaning. Finally, in section 4 I present and defend a more plausible perspective on the relationship between information and meaning, which I label the 'modest view'.

2 Learning from Models

Meaning is primarily a property ascribed to entities in the real world. For instance, scientists describe one of the alarm calls of vervet monkey's as meaning something like *snake approaching* and the firefly's flashes as roughly signalling *I am a female willing to mate*. Crucially, meaning attributions play some important explanatory roles. For example, representational content is supposed to contribute to an explanation of behavior: why do vervet monkeys go into bushes when hearing a certain call? Because the call means *snake approaching*. Why do male fireflies approach females that are emitting light? Because the flashing means *I am a female willing to mate*.

One strategy for improving our understanding of semantic notions is to build models. The nature of scientific modelling is currently a hotly disputed issue (Frigg and Hartmann, 2012; Frigg and Nguyen, 2017), but for simplicity in this paper I will frame the problem using something like Giere's (1988) approach. On this picture (summarized in figure 1) the practice of modelling involves three elements. On the one hand, scientists describe a model system by means of words, mathematics or some other kind of representation. This description specifies a model system that bears some resemblance to the target system. The nature of models is an open question; it can be thought, for instance, as an abstract object or as a fictional entity. Similarly, what kind of resemblance relation holds between the model and the target system is unspecified. I will not assume any definite answer to these questions.

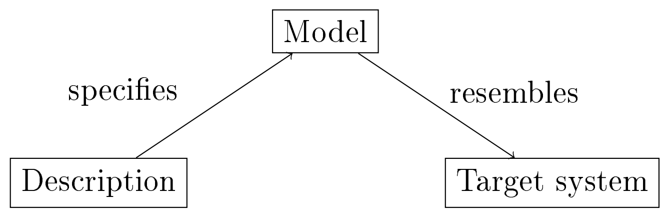

Figure 1: Modified from Giere (1988)

This picture is important because it helps clarify the main goal of this paper: to examine the relationship between a certain class of models and a particular property of target systems, to wit, meaningful signals. More precisely, I would like to discuss whether the undeniable success of formal models of signalling (which employ the notion of information in the model description) can reveal to us the nature of meaning, which is an aspect of the target system.

In a sense, the question we are interested in is whether current models can provide a 'theory of meaning', but note that this expression can be understood in at least two different ways. In a broad sense, a 'theory of meaning' is an account that seeks to answer questions such as the following: what is the most efficient way to communicate meaningful signals? How can vague concepts evolve? Can communication evolve between organisms lacking common interest? By now, it should be obvious that models of signalling can provide very interesting insights into these questions. There is, however, a narrow sense of 'theory of meaning' (which connects with the naturalist literature), whose main question is the following: what it is for a state S to mean p? Can we reduce meaning to information (such that, e.g. for S to mean p just it for S to carry information about p) or causation or some combination thereof? In this paper, I am interested in discussing this last set of questions, so in what follows I will use 'theory of meaning' in the narrow sense.

There is a further issue that needs to be clarified. The perspective on modelling illustrated on figure 1 suggests that properties of the target system can be defined and attributed independently of the models that we employ to understand them. When this idea is specifically applied to meaning, I will label this claim 'MODEL-INDEPENDENCE':

MODEL-INDEPENDENCE (MI) The fact that signals possess meaning does not depend on the theorist's interests or modelling decisions.

Various arguments can be provided in favour of MODEL-INDEPENDENCE (MI). I previously mentioned that attributions of representational content contribute to an explanation of behavior. Now, some might think that for meaning to play this explanatory role, the fact that a signal has certain content should not depend on our explanatory interests (this is true, for instance, if an explanation in terms of content is a specific kind of causal explanation - see Shea, 2018). Furthermore, naturalists who seek to explain content in terms of more fundamental properties should probably embrace MI: since the modeller's psychological states (interests, decisions, etc.) involve semantic properties, any theory that rejects MI fails to explain semantic properties in terms of non-semantic ones. Consequently, in what follows I will assume that the meaning of a message is observer independent, in the sense that a signal's content does not depend on how we as a theorists interpret or model it.

This brief discussion on models helps structure the discussion. In the next section I will briefly survey some models and notions of information that have been employed in game-theoretic analyses of signaling. I will focus on three approaches: Skyrms' very influential account, Birch's original variation and Godfrey-Smiths's alternative proposal. Once these elements are on the table, we will be in position to discuss whether there is anything to be learned on the nature of meaning from the success of these models.

2.1 Models and Information

Since the considerations I will put forward are pretty general, for our purposes a brief description of the simplest Lewis signaling game will do (Lewis, 1969). Suppose a game with two players, 'sender' and 'receiver', such that only the sender can observe how the world is an only the receiver can act on it. Suppose there are only two possible world states (S_1 and S_2), two messages (M_1 and M_2) and two acts (A_1 and A_2); nature randomly chooses a world state, the sender then emits a message and, conditional on the signal, the receiver chooses an act. Both sender and receiver get a payoff of 1 just in case the subscripts of state and act coincide. Senders can follow many strategies: $< S_1 \rightarrow M_1,\, S_2 \rightarrow M_2 >$; $< S_1 \rightarrow M_1,\, S_2 \rightarrow M_1 >$; and so forth. Receivers also have many strategies available: $< M_1 \rightarrow A_1, M_2 \rightarrow A_2 >$; $< M_1 \rightarrow A_1, M_2 \rightarrow A_1 >$, etc. The evolution of this simple game has been extensively studied with replicator dynamics. Suppose we have a population of individuals, which sometimes play the role of senders and sometimes the role of receivers, and in which all possible strategies are played. It has been shown that senders and receivers quickly learn to coordinate their actions and, if certain provisos hold, a signalling system in which both get a payoff of 1 is certain to evolve (Skyrms, 1996, 2010a; Hutteger, 2007b; Huttegger et al. 2010; Bruner et al. 2018).

Now, suppose there is a population in which senders always follow the strategy $< S_1 \rightarrow M_1,\, S_2 \rightarrow M_2 >$ and receivers always act according to the following rule $< M_1 \rightarrow A_1, M_2 \rightarrow A_2 >$. Intuitively, in this model M_1 tells the receiver that S_1 is the case and M_2 says that S_2 holds. Modellers have usually appealed to informational notions in order to capture this intuition. Skyrms' (2010a), for instance, defines the *informational content* of a signal as a measure of how the signal moves probabilities. More precisely, if $P(S_i)$ is the unconditional probability of state S_i and $P(S_i|M_i)$ is the probability of S_i conditional on the presence of M_i, the informational content of a message is a vector specifying how much the presence of a signal changes the prior probability of the state:[*]

$$< log_2 \frac{P(S_1|M_1)}{P(S_1)}, log_2 \frac{P(S_2|M_1)}{P(S_2)} >$$

For instance, in our previous populations signal 1 carries the following informational content: $< 1, -\infty >$, whereas message 2 has the informational content $< -\infty, 1 >$. Skyrms also suggests that when the informational content of a given state excludes some world states, then it carries *propositional content*. Propositional content just is a special case of a more general notion of informational content (Skyrms, 2010a: 42). Thus, in this model the propositional content of M_1 is $< S_1 >$ and the propositional content of M_2

[*] I will focus on the information that messages carry on states, but a similar analysis could be given on the information that signals carry about acts (see Skyrms, 2010a, 2010b, 160; Huttegger; 2007b).

is $< S_2 >$. This is supposed to capture the intuition that M_1 means that S_1 holds and M_2 that S_2 obtains. This claim will be examined with more detail below.

Skyrms also provides a way of measuring the *quantity of information*, which basically specifies how much a given message moves probabilities. More precisely, the quantity of information is a weighted average, with the weights being the probabilities of the states conditional on getting the signal. In general, the quantity of information in M_1 on a set of states S can be defined as follows:

$$I(S|M_1) = \sum_{n=i} P(S_i|M_1)\ log_2 \frac{P(S_i|M_1)}{P(S_i)}$$

Let us consider again the 2×2 game. Suppose we have a population in which individuals sometimes play the role of senders and sometimes play the role of receivers. Assuming replicator dynamics (which is supposed to mimic evolutionary processes), as individuals interact, signals carry more and more information about the world. The population asymptotically approximates a situation in which signals carry the maximum amount of information. The fact that signals carry more and more information runs parallel to the intuition that they are become more and more meaningful.

Birch (2014) has developed an alternative proposal, according to which a signal's propositional content is determined by the informational content that it *would* carry at the separating equilibrium (i.e. one-to-one mapping between states and messages) nearest to the population's current state (Birch, 2014: 505). In other words, if in a population there is a one-to-one mapping between signals and states, the content of a signal M_1 is determined by its information (as in Skyrms' view). However, if this separating equilibrium is not reached, then the content of signal depends on the information that they would carry in the closest scenario in which is a one-to-one mapping between signals and states. Birch argues that this proposal can accommodate the possibility of false propositional content, whereas Skyrms' approach cannot (see section 3.3).

Godfrey-Smith (2012) suggests an alternative analysis (developed in more detail in Shea et al. 2018) according to which the message's informational content is a vector that contains the post-signal probabilities of states. For instance, in the previous example, M_1's informational content is defined as follows:

$$< P(S_1|M_1), P(S_2|M_1) >$$

This approach differs from Skyrms' in that informational content is not a measure of how the message changes the probability of states, but of the posterior probability of the states given the message. Godfrey-Smith argues that this notion better captures the role of content in guiding action: it makes sense that the receiver's behavior is sensitive to the probability of the world state given the signal, but it is much harder to see why it should take into account how much the signal changes prior probabilities. † Thus, for instance, if someone already knows that tomorrow it will be sunny and receives a signal telling her that it will be sunny, on Skyrms' account the signal does not carry information (because the signal does not change probabilities), whereas in Godfrey-Smith's account it does (O'Connor, forthcoming, ch.5.2).

These and other informational concepts (e.g. mutual information) have been used to study various aspects of meaning. Just to mention an illustrative example, Godfrey-Smith and Martinez (2013) argued that communication can be maintained even if there is complete conflict of interest among participants. In their model, it is assumed that the existence of communication can be tracked by measuring the information that states carry about acts (Godfrey-Smith and Martinez, 2013: 3). This is a common practice, that has been shown to be very useful and productive. Even in cases in which information is not used or explicitly defined in the models, the information that a state carries about states (or acts) is often supposed to reveal it semantic content.

Now, assuming that this practice has been successful, can it be used to vindicate an informational theory of meaning?

3 The immodest view

According to what I will label the 'immodest view', game-theoretic models that employ the notion of information yield support to the claim that meaning just is (some kind of) information. Skyrms (2010a: 34), for example, seems to suggest this idea:

†Shea et al. (2018) also define what they call 'functional content'. Although the notion of functional content might share some of the difficulties I will outline below, the main arguments of the paper are not directed at it.

> A new definition of informational content will be introduced here. Informational content, so conceived, fits naturally into the mathematical theory of communication and is a generalization of standard philosophical notions of propositional content.

I take Skyrms to be proposing that his technical notion of informational content (and propositional content) corresponds to the notion that philosophers use when talking about propositional content, i.e. (some sort of) meaning (see also Skyrms and Barrett, forthcoming). Thus, Skyrms not only applies informational concepts in order to study the evolution of meaning; he also seems to embrace the more ambitious project of explaining meaningful relations between organisms in the natural world (vervet monkeys, fireflies, meerkats, bacteria, humans, and the like) in informational terms. As he points out (Skyrms, 2010a: 43-44), 'neither intentionality nor teleosemantics is required to give an adequate account of the informational content of signals. Here I stand with Dretske. The information is just *there*.' In any case, I want to discuss whether models can provide a theory of meaning in the narrow sense, and I think the immodest view is at least compatible with the view of some authors (see Shea, 2014; O'Connor, forthcoming, ch. 5).

Likewise, consider Birch, who not only follows Skyrms in employing informational concepts, but also seems to share the ambitious project of providing an informational theory of meaning:

> In this article, I want to suggest a new way out of what appears a bleak predicament for information-theoretic accounts of content. Though I will focus on content in simple signalling systems – of the sort we find among vervet monkeys, lemurs and meerkats – my hope is that the solution I develop will extend to more complex cases too. (Birch, 2014: 495)

Other authors seem to suggest an interpretation of their work that could be understood as endorsing the immodest approach (e.g. Isaac, 2019). Does the immodest view provide a plausible approach to meaning? Is the fact that semantic properties in models can be used to learn interesting facts about meaningful relations in the world best explained by assuming that meaning just is information? What are the consequences of adopting this perspective? These are the questions I would like to address in what follows.

3.1 Target-Directive Modeling and Modeling Without Targets

Before discussing the arguments for an against the immodest view, however, a last issue needs to be addressed. Models can be used in different ways, two of which are relevant for our purposes. Sometimes the modeller has a specific system in mind that she seeks to explain. The famous Lotka-Volterra model, for instance, was build in order to understand better the population density and species abundance in the Adriatic see at the beginning of the XXth century. Weisberg (2013) calls this practice 'Target-directed modelling'. In other occasions, however, models are used to study more general phenomena, such as sexual selection or parasitism. As an illustration, consider Schelling's (1978) model of segregation, which was not intended to track the segregation process of any particular community, but to study a general process that might lead, for example, to ghettos. Similarly, models about the evolution of altruism are not meant to describe any specific organism, but to draw conclusions about how altruistic behaviour could evolve. These cases involve modelling without any specific target.[‡]

I assume that whether a model is used in the first or the second way utterly depends on the phenomenon that the modeller intends to track, rather than on any intrinsic property of the model itself. For instance, despite the fact that the Lotka-Volterra model was firstly used to model a specific scenario, nowadays it is used as a general approximation to prey-predator dynamics. Thus, in principle game-theoretic models of signalling could be used in target-directed modelling or in modelling without a specific target. This is significant because I think that the set of considerations in favour or against the immodest view greatly depends on how they are understood. Consequently, in the following two subsections I will discuss models of signalling under these two interpretations and I will argue that at the moment neither of them can plausibly vindicate the immodest view.

3.2 Target Directed Modelling

Suppose that one embraces the immodest view and holds that models are interpreted as an instance of target directed modelling. Suppose, for instance, that the model is used to gain knowledge about the vervet monkey's alarm

[‡]In the context of how-possibly explanation, Forber (2010) draws a similar distinction between global and local explanations.

calls. In that case, the informational content that signals carry about states in the model could be used to ascertain the representational content that calls carry about certain events. The goal of this section is to present a difficulty for this approach. In a nutshell, I will argue that this view faces what Birch (2014) calls 'the partition problem' and that solutions to this worry fall prey to a dilemma: they either abandon MODEL-INDEPENDENCE or presuppose (rather than deliver) a theory of meaning.

3.2.1 The Partition Problem

After presenting and defending his own approach, Birch (2014) discusses a difficulty, which he calls the 'partition problem'. Again, suppose that we try to model the vervet monkeys' alarm calls. To define a model, we need to assume a certain number of states, signals, acts and players. Let us imagine that we have some way of distinguishing three different signals: How many world states should be included in the model? Should the state space just distinguish two states, predator and no predator? Or perhaps also different kinds of predators (snake, leopard, eagle)? And why not six different states (e.g. distinguishing snake being close by from snake being far away and so on)?

> In evolutionary game theory, it is usually taken for granted that we can help ourselves to such a specification-that specifying the set of relevant states is part of the modeller's tacit skill. But it is reasonable to ask for further justification if we are to base a theory of content on these foundations. We need a partition of states of the world that is not just a reasonable one for most purposes, but the right one for the specific purpose of individuating semantic contents. (Birch, 2014: 508-509)

We might think of the partition problem as involving two different difficulties. On the one hand, the question concerns the number of states. How many states should be included? This is a pressing question for anyone embracing the immodest view. Since changing the number of states in which a situation is partitioned might change the information that signals carry about states, a different partition of states could change the meaning of signals. For instance, according to Birch (2014) meaning is determined at the separating equilibrium (i.e. one-to-one mapping from signals to states) and this equilibrium only exists if the set of states and signals have the same cardinality, so (as he rightfully admits) we need some criterion for establishing the number of states. I will call this difficulty the 'quantity problem'.

On the other hand, a model can be used to learn about a target system only if there is a specification of how the parts of the model map onto parts of reality, what we might call an 'assignment' (Weisberg, 2013: 29). Thus, even if we know the number of states of our model, there is a second question that needs to be addressed: which states in the target system are represented by states in the model? For instance, supposing that only three states are included in the model, do they stand for eagles, leopards and snakes? They could equally stand for feathers, spotted fur and faked skin or they could represent eagle screech, leopard growl and snake hiss. This is what we might call the 'assignment problem'.[§]

Birch admits that he has yet no satisfactory answer to these worries. Skyrms, of course, is also well aware of this problem. After presenting and defending his own approach, he adds the following footnote:

> This is information content within a given signalling game. It is implicit that this vector applies to the states or acts of this game. For a different game, the content vector shows how the signal moves probabilities of different states, or different acts. Content depends on the context of the signalling interaction. It is a modelling decision as to which game is best used to analyse a real situation (Skyrms, 2010a: 40).

Skyrms suggests that which states or acts are included in the game depends on the modeller's decision. This is certainly true, but it can hardly be a resting point, since we need an answer to the following question: what determines the modeller's decision?

Suppose that there are multiple ways of partitioning states compatible with any scenario and we put no restriction on the modeller's choice. Then, it is not hard to see that there is an important problem: if one embraces the

[§]Note that the difficulty I am pointing out here is not the general worry of whether the conceptual distinctions we draw correspond to something real. I'm assuming that, in general, assignments correspond to real kinds. Rather, the worry is that there are too many kinds to latch onto and we should find some principle criterion for choosing one or another. I would like to thank a reviewer for helping me realize that this clarification was required.

immodest approach, employs a model to understand a specific target system and also holds that the solution to the partition problem depends on the modeller's decision, then the signal's content utterly depends on the theorist's preferences. This consequence is in tension with MODEL-INDEPENDENCE, i.e the idea that the meaning of signals does not depend on the choices or interests of the modeller. The fact that the firefly's flashes mean *female willing to mate* or the vervet monkey's call represents *snake approaching* does not seem to depend on how we as a theorists decide to model this situation. Probably, only if we grant this mind-independence can semantic properties fulfil their explanatory role and vindicate naturalism. Consequently, adopting this proposal would imply that even if models are useful, they cannot show what meaning really is. Hence, unless we can provide an criterion for partitioning states that does not depend on the our interests and decisions, the immodest view is in jeopardy.

Therefore, for the immodest view to be compatible with MODEL-INDEPENDENCE, there has to be some objective criteria that enable us to solve the partition problem. Fortunately, I think that these criteria can probably be found, so there is likely to be a principled way of solving the partition problem. I will argue, however, that plausible solutions to this difficulty cannot vindicate the immodest view.

3.2.2 How to solve the partition problem

A solution to the partition problem requires the following: given a real scenario and a set of signals we are interested in, it has to deliver a procedure for specifying the number of states in the model and their assignment. Intuitively, this recipe should provide the set of states that are good candidates for being the referent of signals.

One way of addressing this difficulty is by paying attention to actual scientific practice, since this is a problem that needs to be solved when assessing particular cases. For instance, in a landmark work that employs game-theoretic models to study animal signals, Searcy and Nowicki (2005: 3, emphasis added) claim:

> The difficulty with this formulation is in ascertaining what the signal is 'supposed to' convey. 'Supposed to' in this context must be interpreted from the viewpoint of the receiver rather than the signaller; what matters is whether the signal *conveys something that the receiver would benefit from knowing.*

According to Searcy and Nowicki, what a signal is supposed to express (what we might call its 'content') depends on the states that the receiver would 'benefit for knowing', in the sense of being states that causally effect the fitness of the organism (see also Birch, 2014: footnote 23). Indeed, that looks like a very plausible idea: the reason the state space in the model includes a representation of eagles and not feathers or screeches is that eagles make a difference concerning their survival and reproduction of monkeys, whereas feathers and screeches (as such) do not. There is a significant causal relationship between the presence of eagles and monkey's survival. Thus, in order to identify the state space, one needs to appeal to properties that are causally relevant for a change in fitness. Note that this proposal is not very far from the key intuition driving teleological theories of content, according to which (roughly) the content of a signal is the state that causally contributed to fitness when the signal was tokened in the recent evolutionary past (Millikan, 1984; Neander, 2012).[¶]

Now, suppose that one might need to assume some teleological theory in order to solve the partition and assignment problems. If that were true, it would have a striking consequence for our main question, since teleological theories are naturalistic theories of content; these theories are supposed to explain what meaning really is. Accordingly, if the use of models to learn about content requires solving the partition problem, and in order to address this difficulty we need to assume a theory of meaning, then models cannot be used to provide a theory of content.

I think the argument generalizes, since it is largely independent of the truth or falsity of teleological theories. Suppose that what is actually required for solving the partition and assignment problems is a theory T. T could appeal to fitness differences, functions, causal relations or whatever have you. My point is that T would probably count as a theory of meaning. Solving the quantity and assignment problems requires providing a procedure for specifying the states that are good candidates for content; in the vervet case, three states are assumed, which are supposed to correspond to eagles, leopards and snakes (rather than, say, feathers, spotted fur and faked skin, for instance). Only if three states are assumed, which are assigned to these three

[¶]One significant difference is that teleological theories usually appeal to past contributions to fitness, while this is not explicit in Searcy and Novicki's models. Nonetheless, this aspect might well be implicit; after all, animals are usually studied in their natural habitat, so what causally affects fitness in the present is probably closely related to what had a causal impact in the past.

kinds of predators can Birch's or Skyrms' models deliver the right content attributions. The question, however, is how could they offer a criterion for picking up these referents that falls short of qualifying as a theory of meaning.

Therefore, a plausible specification of the principles used by the modeller to solve the quantity and assignment problems probably requires providing a theory of meaning. If this reasoning is on the right track, game-theoretic analyses involved in target-directed modelling have to assume a naturalistic theory of content, so they presuppose (rather than deliver) a solution to the problem of the nature of meaning.

The argument developed in this section can be put in the form of a dilemma: if one embraces the immodest view and intends to use a game-theoretic model in target-directed modelling, one needs to solve the partition problem and the solution either depends on the modeller's decisions or it does not. If it does depend, then this proposal is incompatible with the mind-independence of content attributions (what I labelled 'MODEL-INDEPENDENCE'). If it does not, then this proposal probably presupposes a theory of meaning. Consequently, a model of signalling is probably unable provide a theory of meaning (in the narrow sense) that is compatible with MI. Thus, it is hard to see how the immodest view can be vindicated by means of target-selective modelling.

3.3 Modelling Without Targets

The previous section explored the interpretation of game-theoretic models as an instance of target-selected modelling. However, I argued that one can also understand them as a models without a specific target. It is not hard to find quotes suggesting that something along these lines is often intended. For instance, Huttegger (2007a: 1) argues that a goal of his models is to show 'how meaningful communication can emerge'. A similar idea is expressed by Zollman (2011: 160) when he claims that 'these evolutionary accounts offer a sort of proof of possibility that much human-like meaning can be had without appeal to intentions'. Other questions involve: 'Where do indicatives and imperatives part? Do some imperative have a grounding similar to that of indicatives?' (Huttegger, 2007b: 410); 'Why do vague predicates arise in the first place?' (O'Connor, 2014: 707); 'How much more efficient does salience make the evolution of communication? How might salience itself evolve — i.e., how might an evolving signal become salient?' (LaCroix, forthcoming: 9). In many cases the goal is to address some general questions about meaning, rather than modelling any particular system. Accordingly, let us suppose that models are used in that way. Can they lend support to the immodest view, i.e. the idea that meaning just is information?

As a preliminary remark, note that in this case we lack one kind of support that is usually provided for theories of content. In general, a common way of testing them (although, of course, not the only or most important one) is by considering particular cases (frog's mental states, warning calls of vervet monkeys, mating calls of fireflies, etc...) and analysing whether their predictions fit the content attributions of scientists and laymen. However, for the reasons outlined in the previous section, this is just not possible without getting trapped into the partition problem. An alternative suggestion is to look for analogies between the two notions: Concepts are supposed to play certain explanatory roles, so perhaps one could argue that information plays the same role in models as meaning in the outer world (e.g. Hutteger, 2007a, p. 413; Zollman, 2011: 162).

In section 2.1 I pointed out some broad similarities between the role information plays in models and representational content in target systems. For the sake of the argument, let us assume that there are some analogies between how information behaves in models and meaningful relations in the real world. Does that suffice for vindicating the immodest view? Although I doubt knock-down arguments can be provided a that point, in what follows I will present some considerations that suggest a negative answer. More precisely, I will provide two reasons for thinking that the success of current models probably fail to vindicate the immodest view.

First, note that there are various non-equivalent definitions of content in terms of information. Indeed, I think a close look at current models suggests that the multiplicity of concepts of information is not a sign of immaturity, but captures a central aspect of the modelling practice. The variety of concepts of information has some utility, since many of them have been used to learn about different aspects of meaning, as suggested in section 2.1. For instance, in some cases we might be interested in knowing how much knowledge we acquire from a signal (so we might use Skyrms' definition), in others we might be interested in capturing what a signal tells us about the world independently of our previous knowledge (so we might use Godfrey-Smith's notion) and still in others how current usage differs from certain pattern existing in 'ideal conditions' (so we might use Birch's or, alternatively, Skyrms and Barret's more recent proposal, depending on the 'normal conditions' we want to focus on). This list is of course not exhaustive. These notions seems to be suitable for a certain purposes, so it is not a case in which we have different, competing proposals, but one in which a variety of tools have been developed. This variety seems to be a feature, not a bug.

This plurality of useful informational notions, however, puts the immodest theorist in a dilemma, depending on whether she maintains that one of these concepts is the one that actually defines what meaning really is or whether she holds that all of them do. If one takes the first horn and seeks to identify meaning with one of these notions of information, one should be able to specify reasons for favouring one rather than another. It should say, for instance, why Skyrms' notion of informational content (which is a vector specifying how much the presence of a signal changes the prior probability of a signal) is more likely to capture what meaning really is than Godfrey-Smith's notion (which measures how the message changes the probability of states) or any other. I think that at the moment we lack convincing reasons for thinking one of them should be superior to others and, indeed, we should not expect any of them to prevail. Different notions have been successfully used to learn about various properties of representations, so why should we think only one of them captures what meaning really is?

This reasoning leads to the other horn of the dilemma: perhaps we do not need to choose between the different definitions; every useful analysis of information corresponds to a different important semantic relationship in the target system (for a version of this view, see Shea et al. 2018). Note, however, that this is a relatively extreme form of pluralism; the point is not just that we might have to accept more than one kind of content (something that some people are willing to accept anyway). Even if one admits that more than one concept of content is required, multiplying the semantic entities one is committed to just because one can build illuminating models seems to be too promiscuous. In any case, in the next section I will describe a different interpretation of what these models achieve that does not have this strong metaphysical commitment. Since the modest view provides a much more ontologically parsimonious explanation of this success, I think that, all things being equal, it should be preferred.

The problem suggested by the multiple notions of information is not the only consideration against the immodest view. Since the first informational theory of content was provided by Dretske (1981) there has been some discussion on the prospects of such a theory and some important difficulties have been identified (Neander, 2012; Adams and Aizawa, 2017; Artiga and Sebastián, forthcoming). Crucially, despite the undeniable success of recent models of signalling, the classical problems of informational theories of content have not been satisfactorily addressed yet. To justify this claim, let me briefly discuss two classical problems: liberality and misrepresentation.

A classical objection against informational theories of content relies on the fact that information is cheap: any state carries information about many other states of affairs. Do the informational notions developed in signalling models address this issue? Within the models, this problem is avoided by limiting the number of states, acts and so on. These idealizations might give the wrong impression that this objection has been satisfactorily addressed, but this is an artificial result deriving from model design. Only if one restricts the space of states, acts and others in certain ways, one might be in position to find some privileged informational relation between signals and referents. In this case, the theoretical assumptions are doing all the work in solving the worry; unfortunately, once we consider the idea that informational relations underpin meaningful relations in the target system, this difficulty reappears in all its strength.[‖] Although in the real world signals certainly carry information about their referents, they also carry information about many other states that do not fall under its extension, and it is not obvious there is any measure of information that can single out the message's content.

Secondly, whereas the possibility of error is a central property of meaning, a traditional problem of informational theories of content is that they fail to allow for misrepresentation. Similarly, informational notions provided in the context of models of signalling face serious difficulties in accommodating this aspect. Skyrms' and Godfrey-Smith's notion of information, for instance, cannot be false (Godfrey-Smith, 2012). Birch's analysis made some progress, since it leaves room for some cases of misrepresentation but, as he admits, error is still impossible when a system has reached a separating equilibrium[**] (Birch, 2014: 505; see also Godfrey-Smith, 2012; Ventura, 2017). On Skryms and Barrett's (forthcoming) approach, misrepresentation is impossible once

[‖] There seems to be a connection between the this worry and the partition problem. Perhaps one could envisage the partition problem as a version of the liberality objection.

[**] Birch's approach does a pretty good job in accommodating cases of misrepresentation thanks to the fact that he defines the content of a signal counterfactually (so, in this sense, his approach also falls into the category of modal theories, such as Fodor, 1990, Nanay, 2014 or Prinz, 2002). This theoretical choice, however, comes at a cost: for instance, in some cases the informational content that a state would carry in the nearest equilibrium fails to correspond to the content that the signal seems to be carry in the actual circumstances (see Skyrms and Barrett, forthcoming). AS a result, his theory might predict misrepresentation in cases in which it seems to represent correctly and vice versa.

the game has reached an equilibrium in a context of common interest.[††] But why should common interest preclude error?[‡‡] Thus, the possibility of misrepresentation remains one of the central features defining representations that informational approaches have difficulties in accounting for. Of course, I am not claiming that these problems cannot be solved; rather, my point is that the classical arguments against informational theories of meaning still need to be satisfactorily addressed before the immodest view can plausibly be vindicated.

In conclusion, I think the immodest view faces important difficulties. Even though we can learn many things by considering how information behaves in models, at the moment it is far from clear that this fact can support a naturalistic theory of content in terms of information, whether we are engaged in target-selective modelling or in modelling without target. Fortunately, there is an alternative perspective on the relationship between models and target systems that does not have these drawbacks. As I will argue in the next section, one can keep all the explanatory benefits of current models while avoiding these difficult problems. The price, however, is to give up the immodest view, and adopt a more modest approach.

4 The modest view

I argued that the immodest approach face serious difficulties, so I think it is time to consider an alternative. Modest accounts maintain that there are certain similarities between the different notions of information applied to models and semantic content in target systems that enable use the former in order to learn about the latter. Nonetheless, according to a modest perspective, this fact fails to vindicate informational theory of meaning: these analogies are robust enough for allowing us to gain interesting insights about content, but they are insufficient for telling us what meaning really is. In other words, informational relations within models are useful for tracking non-informational relations in the outer world. Let me elaborate.

[††] In various respects, Skyrms and Barrett's proposal seems to be analogous to Dretke's (1981) distinction between a learning and a post-learning period that he introduced to accommodate misrepresentation. This solution has been criticized by many people (McLaughlin, 2001).

[‡‡] It is slightly more plausible to think that common interest should preclude lying and deception, which is the main motivation for their proposal.

Recall that models are indirect representations of target systems. By studying models, scientists can obtain knowledge of worldly entities and processes, even though (or, perhaps, thanks to the fact that) models tend to involve abstractions and idealizations (Godfrey-Smith, 2006). One of the central properties of representations, however, is that they can work as such even if they differ in many respects from the entity they represent. Two analogies might illustrate one way of developing a modest account.

Consider, for instance, the famous Phillips Hydraulic Computer or MONIAC (Monetary National Income Analogue Computer) designed by Bill Phillips, which was an analogue computer that modelled the national economic process of the United Kingdom by means of water tanks and pipes. The flow of water from one tank to another (e.g. from the tank labelled 'income' to another one labelled 'savings') was supposed to correspond to the flow of money from one part of the economic system to another. The same equations that were thought to rule the national economy were supposed to describe as well water flow.[§§] Although manipulating and tracking the flow of water could be used to explain certain phenomena and make some predictions (Nguyen, 2016: 159), it would have been nonsensical to identify money with water flow in the real world just by pointing at this success. A model might be illuminating because some of its features can be used to track a very different kind of property in the target system.[¶¶]

If the MONIAC model provides a useful illustration of the idea that one can successfully model one aspect of the target system by using a very different kind of property, cartographic representations show that multiple perspective can be useful, even if all of them bear on a single entity. There is a wide diversity of maps of the Earth that employ different projections, and each of is useful for a different purpose. In Mercator maps, for instance, rhumb lines are straight, which is very useful for navigating, but it misrepre-

[§§]According to Bissell's (2007) description: "The water flow representing total national income enters at the top of the machine.(...) Similarly, flows in other pipes represent imports and exports, while tanks (acting as integrators, from a control engineering point of view) represent financial balances of various kinds. (...) In this way, various mechanisms in the economy are simulated, including feedback introduced by government fiscal control action."

[¶¶]Of course, in some sense, information theory seems to provide a more direct representation of the target phenomenon than MONIAC, but precisely for that reason I think the comparison with MONIAC is illuminating: it illustrates the fact that models can be successfully used to learn about a target system that is constitutively very different. I want to thank a reviewer for pressing me on that point.

sents the surface of continents and distances. The Gall-Peters projection does a much better job in representing relative sizes, but it distorts most shapes. Stereographic (azimuthal) maps are valuable for moving around pole, but it only provides a partial representation of the rest of the planet. There are dozens of different projects, many of them useful for different purposes and all of them represent a single object: the Earth.

I think informational content in models should be roughly treated in a similar way: as involving multiple perspective on a common phenomenon (like maps) using a different property (as in the MONIAC). We often use models and informational relations between variables in order to track and learn about meaningful relations in the real world, but the fact that this is very useful fails to show that meaning in the real world can be identified with some kind of information. It rather supports the idea that scientists sometimes employ informational measures within models to learn about another kind of relations in the natural world. Information within models might work as a proxy for meaningful relations, due to partial similarities between them (see 1). Some notions of information might capture some aspects of semantic relations, whereas other notions might be more suitable for understanding other dimensions. So far, however, we lack convincing reasons for thinking this success can provide a theory of meaning in the narrow sense.

Note that the modest view avoids all the previous difficulties we raised against the immodest view. First of all, it is compatible with a reasonable solution to the partition problem: if we are interested in modelling a set of signals, we can simply employ our preferred theory of meaning to identify the state space of our model. Furthermore, the theory of meaning we employ need not appeal to the modeller's interests or her theoretical decisions, so it is not in tension with MODEL-INDEPENDENCE.

Secondly, the fact that different notions of information are successfully employed fails to generate problem, given that the modest view does not assume that they need to correspond to different semantic relations in the target system. Models that employ different measures of information can be used to learn about different aspects of representational content even if all of them are about the same semantic property in the target system. I think this perspective is liberating: the modeller does not need to look for a definition of information within models that corresponds to an appropriate informational relation holding in the natural world between a sign and its referent. Thus, Skyrms', Birch's and Godfrey-Smiths' notions might all be useful for different purposes, and, again, this is compatible with MODEL-

Independence. Finally, we do not need to solve the difficult problems that informational theories of content face.

Certainly, these advantages of the modest view come at a price: we need to accept that game-theoretic models of signalling might be unable to provide a theory of meaning in the narrow sense. Despite the large amount of interesting insights provided by these models, they might just not be the right tools for providing a fully satisfactory theory of the nature of meaning.

5 Conclusion

Many of the recent game-theoretic analysis of signalling have been remarkably useful for studying different aspects of content. In this paper, I tried to analyse the consequences of this success for naturalistic theories of meaning. In particular I argued that the immodest view, which seeks to reduce meaning to some sort of information, faces important difficulties, whether it is interpreted as an instance target-selective modelling or as an example of modelling without target. In contrast, I think the modest view keeps the explanatory virtues of these models without inheriting any of their difficulties. At least, that is the modest claim I have been trying to defend.

References

[1] Adams, F. and K. Aizawa (2017) Causal Theories of Mental Content. *The Stanford Encyclopedia of Philosophy.* (Summer 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/content-causal/>.

[2] Artiga, M. and M. Á. Sebastián (forthcoming) Informational Theories of Content and Mental Representation. *Review of Philosophy and Psychology:.*

[3] Barrett, J. A. (2014) Rule-Following and the Evolution of Basic Concepts. *Philosophy of Science* 81(5): 829-39.

[4] Birch, J. (2014) Propositional Content in Signalling Systems. *Philosophical Studies* 171(3): 493-512.

[5] Bissell, Ch. (2007) Historical perspectives - The Moniac A Hydromechanical Analog Computer of the 1950s. *IEEE Control Systems Magazine* 27(1): 69-74.

[6] Brooks, D., R. T. O'Grady and E. O. Wiley (1986) A Measure of the Informartion Content of Phylogenetic Trees, and its Use as an Optimality Criterion *Syst. Tool.* 35(4): 571-581.

[7] Bruner, J. A. C. O'Connor, H. Rubin, S. M. Huttegger (2018) David Lewis in the Lab. *Synthese* 195(2): 603-621.

[8] Da Costa, French. (1990) *Knowledge and the Flow of Information.* MIT Press.

[9] Dretske, F. (1981) *Knowledge and the Flow of Information.* MIT Press.

[10] Fodor, J. (1990) *A Theory of Content and Other Essays.* MIT Press.

[11] Forber, J. (1990) Confirmation and explaining how possible. *Studies in History and Philosophy of Science Part C.* 41 (1):32-40.

[12] Frigg, R. and S. Hartmann (2012) Models in Science. *Stanford Encyclopedia of Philosophy*, Edward N. Zalta, (Ed.) URL: <https://plato.stanford.edu/archives/sum2018/entries/models-science/

[13] Frigg, R. and J. Nguyen (2017) Models and Representation. In Lorenzo Magnani and Tommaso Bertolotti (eds.)*Springer Handbook of Model-Based Science* Dordrecht, Heidelberg, London and New York: Springer, 49-102.

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

[15] Godfrey-Smith, P. (2006) Mental Representation, Naturalism, and Teleosemantics. In G. MacDonald and D. Papineau *Teleosemantics.* Oxford University Press, 42-68.

[16] Godfrey-Smith, P. (2012) Review of Brian Skyrms' Signals. *Mind* 120: 1288-1297

[17] Harms W. and B. Skyrms (2008) Evolution of Moral Norms. In M. Ruse (ed.) *The Oxford Handbook of Philosophy of Biology.*

[18] Huttegger, S. (2007a) Evolutionary Explanations of Indicatives and Imperatives. *Erkenntnis.* 66: 409-436.

[19] Huttegger, S. (2007b) Evolution and the Explanation of Meaning. *Philosophy of Science.* 74: 1-27 .

[20] Huttegger, S., B. Skyrms, R. Smead and K. Zollman (2010) Evolutionary dynamics of Lewis signalling games: signalling systems vs. partial pooling. *Synthese.* 172 (1):177-191.

[21] Isaac, A. (2019) The Semantics Latent in Shannon Information. *The British Journal for the Philosophy of Science.* 70 (1):103-125.

[22] LaCroix, B. (forthcoming) On salience and signaling in sender–receiver games: partial pooling, learning, and focal points.*Synthese.*

[23] Lewis, D. (1969) *Convention.* Harvard University Press.

[24] Mclaughlin, B. (1991) *Dretse and His Critics.* Blackwell

[25] Millikan, R. (1984) *Language, Thought and Other Biological Categories.* MIT Press.

[26] Millikan, R. (1984) *Varieties of Meaning.* MIT Press.

[27] Nanay, N. (2014) Teleosemantics without Etiology, *Philosophy of Science* 81(5): 798-810.

[28] Neander, K. (1995) Misrepresenting and Malfunctioning. *Philosophical Studies.* 79(2): 109-41.

[29] Neander, K. (2012) Teleological Theories of Mental Content. *The Stanford Encyclopedia of Philosophy.* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/content-teleological/>.

[30] Nguyen, J. (2016) *How Models Represent.* PhD Dissertation, London School of Economics.

[31] O'Connor, C. (2014) Evolving Perceptual Categories. *Philosophy of Science.* 79(4): 707-727.

[32] O'Connor, C. (forthcoming) *Games in the Philosophy of Biology.* Cambridge University Press.

[33] Papineau, D. (1993) *Philosophical Naturalism.* Blackwell.

[34] Prinz, J. (2002) *Furnishing the Mind: Concepts and Their Perceptual Basis.* MIT Press.

[35] Schelling, T. C. (1978) *Micromotives and Macrobehavior.* WW Norton.

[36] Searcy W. and S. Nowicky, B. (2005) *The Evolution of Animal Communication.* Princeton University Press.

[37] Skyrms, B. (1996) *Evolution of the Social Contract.* Cambridge University Press.

[38] Skyrms, B. (2010a) *Signals: Evolution, Learning and Information.* Oxford University Press.

[39] Skyrms, B. (2010b) The flow of information in signaling games *Philosophical Studies.* 147:155–165

[40] Shea, N. (2014) Neural Signalling of Probabilistic Vectors. *Philosophy of Science.* 81(5): 902-913

[41] Shea, N. (2018) *Representation in Cognitive Science.* Oxford University Press.

[42] Shea, N., P. Godfrey-Smith and R. Cao (2018) Content in Simple Signalling Systems. *British Journal for the Philosophy of Science.* 69(4): 1009-1035.

[43] Stegmann, B. (2009) A Consumer-Based Teleosemantics for Animal Signals. *Philosophy of Science.* 76(5): 864-875.

[44] Ventura, R. (forthcoming) Ambiguous signals, partial beliefs, and propositional content. *Synthese.*

[45] Weisberg, M. (2013) *Simulation and Simularity.* Oxford University Press.

[46] Zollman, K. (2011) Separating directives and assertions using simple signaling games. *The Journal of Philosophy.* 63(3): 158-169.