Logique & Analyse **164** (1998), 343–363

INTERPRETING REALITY: MODELS AND REFERENCE

EVANDRO AGAZZI

Abstract. The notion of model has different meanings in common language and in the context of the sciences. In the formalized sciences this meaning has been established in the special sense of the *extensional* semantics of mathematical logic and is the basic notion of "model theory". It is characterized by the fact that the model must satisfy the conditions of the linguistic construction. When this construction, however, is meant to characterize its different models (or even a single "intended model"), several well-known limitations exist. In the case of the empirical sciences the situation is opposite: a linguistic structure (a theory) must be such as to satisfy its intended model (or domain of objects). This opposition can be overcome if one considers a model as an *intensional* reality (an intellectual representation) encoding a certain amount of properties, that are exemplified by concrete objects to a certain extent. The model (understood in this way) constitutes a realm of *sense* for which a *reference* is looked. From this point of view a scientific theory is the linguistic explicitation of its intellectual model, and the referents one intends to understand and explain by means of the model are captured by non linguistic "criteria of referentiality" of an *operational* nature. In this way it is possible to see that modifications of theories in the presence of unsatisfactory empirical evidence may happen without totally dismissing the theory, but simply by slightly modifying it, or its model. The ontological status of the entities to which a theory refers can be correctly established by considering the different criteria of referentiality.

Some distinctions

The concept of model has a variety of meanings, especially if one considers its use in common language rather than in the context of the sciences. In the most colloquial sense, a model means either an abstract idealization, that is meant to express the genuine nature of some "kind" of reality, or a concrete

entity, that is meant to realize in an outstanding way the characteristics of that "kind" of reality to which it belongs. In this case, the concrete entity is precisely thought to *exemplify* in a particularly significant measure (that we also often qualify as an "exemplar" manner) that which is *encoded* in the abstract idealization. It is in this sense that we say, for example, that a certain person is a model of honesty, that the recital given by a pianist was a model of musical performance, that a murder whose authors could not be detected was a model of perfect crime. But we also say (and this corresponds to the first and "abstract" sense of model) that the model of correct logical inference is that of a reasoning in which only what has been explicitly assumed in the premises is used for deriving the consequences, or that the model of an economically reasonable behaviour is that in which the amount of costs does not exceed the amount of benefits.

In both these meanings a certain notion of perfection is implicit, since an imitation of, or conformity with, the model is meant to be the appropriate way for attaining the "best" realization of the kind of reality (object, performance, human action) to which the model refers. This "axiological" aspect is present even in those cases in which other types of value judgments clearly imply a negative evaluation (think, for example, of the idea of a "perfect crime"). When we pass from common language to the terminology of the sciences, we can observe that the two fundamental meanings are essentially retained, and that the axiological flavour is strongly attenuated, but not totally dismissed either. In the sciences, however, takes a great importance a third sense of model, that is not unusual, but rather limited, in common language: the sense according to which a model plays the role of an *analogy* that is heuristically useful for the understanding or explanation of a certain domain of objects.

The concept of model in the sciences

The notion of model has received a specialized and technical sense in the domain of mathematical logic where it is the basic concept of what is denoted as the *semantics* of formal languages and theories, and is strictly related to the concept of *interpretation*. Given a formal language, a model of this is provided by an interpretation that relates its non logical symbols to certain entities or classes of entities belonging to a domain that is normally different from the language itself (in special cases, however, like in the case of Henkin's proofs of the semantic completeness of certain logical calculi, the language itself can provide the domain of reference for its metalinguistic interpretation). If the model provided in this way for the language happens to be such that an expression or sentence formulated in this language (according to its formation rules) becomes *true*, we say that this is also a model

of this expression or sentence (certain logicians used to say that in this case the interpretation is a model of the sentence, but this is a rather ambiguous way of speaking). The same discourse easily applies to a set of sentences and to a formal theory, if one considers a (formal) theory simply as a set of sentences: one can say that a model (provided by a certain interpretation of the language) is a model of the set of sentences, or of the theory, if all the sentences of the set or the theory are "true in this model".

All the said summarizes the most elementary and well-known ideas of *model theory*, as it is understood in mathematical logic. It is interesting, however, to consider this from the more general point of view outlined above. The formal theory (or the set of sentences) plays the role of the ideal or abstract structure that *encodes* the features its different models *exemplify*. This accounts in the simplest way for that multiplicity of models which is characteristic of formal calculi and theories. A deeper scrutiny shows us that even the "light" axiological aspect mentioned above is present here, and this is manifest when we *look for* a model of a set of sentences (or a formal theory). In this case, if we find that a certain model of the language is *not* a model of the set of sentences, we *discard* the model, and look for another one. If, after many efforts, we are unable to find such a model, we may remain unhappy, and even become a little perplex about the legitimacy of the theory (mainly because finding a model has become a standard way for securing a relative warranty of consistency for formal theories, but it is not said that this is the *only* warranty). Nevertheless we are not obliged to dismiss the theory. In a pictorial way we could say that, in the case of formal theories, the language "dominates" over the models, that they must conform to the linguistic structure, that the legitimacy of the model depends on its capability of "satisfying" the linguistic structure, and not the other way around.

Quite opposite is the situation in the case of the *empirical theories*. For them a "model" (in the sense described above) is *given*: it is the domain of reality they are *intended* to investigate, and they "must" speak with truth about this model. This means that the model "dominates" over the language. Even without taking the step leading to a "formalized" empirical theory, already at the level of an "intuitive" empirical theory we must admit that it is a *linguistic* construction that (in a rather general sense that we shall not analyze now) is "obliged" to be true of its intended model, to "conform" to it. If the theory does not "satisfy" its model, we must *discard* it, and look for a better theory. We never discard the model. Clearly, these remarks apply even more strongly when an empirical theory is formalized.[1]

[1] The ideas sketched here have been already presented in Agazzi [1976] with more details.

We want to stress that we have spoken of "formal theories", that must not be confused with "abstract theories": these are different from the empirical ones only because of the particular "kind of objects" they investigate, but they nevertheless investigate an *intended* domain of non-empirical objects (therefore, their difference lies in the *ontology* of their respective objects, but not in their *cognitive* attitude). The best examples of such abstract theories are provided by several branches of mathematics, such as arithmetic, geometry, analysis, probability theory, and many other. In such cases we are used to distinguish between an "intuitive" and a "formalized" arithmetic, geometry, analysis, set theory, and so on, and, what is more important, we speak of an *intended model* of them (that is, the model of natural numbers, of geometric entities, of analytic functions, of sets, and so on). The formal systems designed for formalizing those mathematical theories (usually by means of a suitable axiomatization) are expected to "capture" their intended model and describe it as faithfully as possible (this is why, on other occasions, I have spoken of "concrete" mathematical theories, for denoting this type of theories, and of "abstract" mathematical theories, for denoting those that do not aim at describing an "intended" model, but are open to a variety of interpretations, adopting a terminology that does not overlap with that I am using here).[2]

The difficulty of capturing the model

Mathematical logic contains several results showing how complicated and unsatisfactory is the task of linking a formal system with its models. The Löwenheim-Skolem theorem is only one among those proving the theoretical impossibility of bringing a formal theory to cope with its intended model. More elementary metatheorems already show this "limitations" of formal systems (as they are often called). Gödel's incompleteness theorem for elementary arithmetic, for example, already shows that there are infinite many propositions that are true in the intended (or standard) model of natural numbers, but not "captured" by any formal system invented for axiomatizing this arithmetic (the formal system system tells "less" than what is true in the standard model). However, since Gödel's theorem can be formulated and proved also in first order logic, and this logic is semantically complete, we must recognize that if a formal proposition is not formally derivable from a set of axioms, it is not their "logical consequence", that is, is not true in *all* models of these axioms. Therefore, those arithmetical propositions that are true in the standard model, but are not formally derivable from the axioms, must be

[2] For this different use of the notions of "concrete" and "abstract" mathematical theories cf. Agazzi [1977], reprinted in Italian in Agazzi [1978].

"false" in other models of these axioms, and this is already an indication of the existence of "non-standard" models of arithmetic (which were also actually constructed). But this amounts to saying that our axioms are unable to characterize their intended model in a deeper sense: they imply "too much", in the sense that allow for the justification of properties incompatible with those of our intended domain.[3]

This limitation is stronger than that already emerging from the "isomorphy theorem", that shows that, if a system of propositions admits a model, it also admits all models that are isomorphic with this. This fact means that a formal system can characterize its intended model only "up to isomorphism", that is, only as far as its "structure" is concerned (and this could be tolerable). A formal system enjoying this property is said to be "categoric". But the non-standard models are mutually non-isomorphic, and this is highly undesirable. The precondition for establishing an isomorphism between two models is that they are of the same cardinality or power, but the Löwenheim-Skolem theorem shows that, if a set of propositions of first order logic admits a model, it also admits a denumerable model, and this result, in a way, "conflates" the cardinality of all models and renders their isomorphism impossible. This explains why the research of categoricity often restricts itself to a "categoricity in power" (that is, within a certain cardinality). A remedy to this situation can be found in passing to higher-order logics. For example, Peano's arithmetic is proved to be categoric in the second order, but then the semantic incompleteness of the second order logic follows from this result, combined with the semantic incompleteness of arithmetic proven in Gödel's theorem, that does not depend on the adoption of a specific logical calculus. In brief, and without going into technicalities that are not needed for the development of our reflections, semantic completeness and categoricity of formal system are hardly compatible, and this indicates that the formal capturing of an intended model is a desperate enterprise. The reasons for this situation are deep and complex, and we estimate that they chiefly reside in the "extensionalist" approach adopted in model theory. We cannot dwell on this question, and some indirect light will come from considerations we will present in the sequel.[4]

[3] For additional considerations regarding the philosophical implications of Gödel's theorems, cf. also Agazzi [1992b].

[4] Some more details regarding this issue can be found in Agazzi [1979] and Agazzi [1994].

A broadening of the concept of model

The concept of model we have considered until now is essentially derived from the meaning it receives in the "model theory" of mathematical logic. It is a rather specialized meaning that, in particular, significantly differs from the meaning attached to this concept in the empirical sciences, such as physics. The best way for appreciating this difference is the consideration of the language-world relation. In the formal disciplines, the priority is given to language, and something is looked for in the world that is capable of satisfying the linguistic construction, of "corresponding", in a certain way, to it. This something is the *model*. In the case of the empirical sciences the dynamics is inverted: priority is given to the world, and "something" is looked for, in the effort of understanding and explaining the portion of the world we want to investigate. This something is expected to offer us some kind of intelligible picture of this portion of the world, and we call it again a *model*. This model, like in the first case, is expected to "correspond", to a certain extent, to the reality to which the priority is assigned (in this case, to the world), but is normally quite different from a linguistic structure, though, as we will see, also certain linguistic structures may be considered as models of the world, according to this new meaning.

Confirmations of the said comes from the consideration of some meanings attributed to the notion of model in the empirical sciences. The lowest-level of such meanings is probably used when a scientist, when he provides an account of a given field of investigation of whose inadequacies and partiality is fully aware, says: "this is *just* a model". In this sense the model is understood as a rough first approximation, with no greater pretention than that of offering an initial schematization and a few general ideas, that can serve as guidelines for the further deepening of the investigation, and would probably be dismissed or, at least, essentially refined. Therefore, the model (understood in this sense) has a pragmatic role and, from a cognitive point of view, some heuristic value. More engaging, and adopted, is the meaning of model we find when in certain fields of research, "models" are constructed using elements derived from *other* fields, in order to understand and explain the facts of the investigated field (for instance, when hydrodynamic models were employed for the study of electric currents, or several mechanical models of the ether were proposed for the understanding of the electromagnetic field). In this case, models are used not simply because of their heuristic force, but for their *analogical* nature: if the properties already sufficiently explored in a given field appear to be "similar" to those characterizing a different and better known field, it is very natural to try to "broaden" the similarity, that is, to extend other properties of the better known field to the field under investigation, hoping that this similarity continues to hold. This

intuitive idea can be easily translated into more rigorous terms: the appearing similarity between the two fields, since their ontologic nature is different, cannot be but a similarity in *structure* and, since it has been ascertained only on a part of the field under investigation, it cannot be taken to be an identity of structure (isomorphy) but, precisely, only as a similitude in structure, that is, as an *omeomorphism*. If the force of the omeomorphism turns out to be such as to convince the scientists that it actually is an isomorphism, the result may even be that of a unification of the ontology of the two domains, or the reduction of the one of them to the other (as it happened, for example, with the electromagnetic theory of light, that actually meant a substantial reduction of optics to electrodynamics). These reflections indicate how strictly models are related to theories: what we have said suggests that, at least in certain cases, the adoption of a model may evolve in the construction of a theory, but we cannot insist on this point now, since we need to rely upon a notion of theory that we have not elaborated yet (the rudimentary concept of a theory as a set of propositions provisionally adopted above is totally insufficient for a rigorous analysis). Once this meaning of model (as something admitted on the ground of a structural similarity or identity with the investigated field) is accepted, the way is open for enlarging the notion of model up to include also *abstract* models, whose ontological constituents are, for example, mathematical entities. In this case we can speak of the mathematical model of a certain domain of inquiry, and can even say that the system of equations of a given physical theory is a model of the domain this theory investigates. Since these equations are often reduced to be part of the *language* of the theory, it often happens that the theory itself be qualified as a "linguistic model" of its domain. Several misunderstandings are implicit in these passages, but we need not spell them out here. The interesting fact is that, in such a way, we have reached the end of the inversion of perspective mentioned above: the language becomes the model of the world, in the sense that it is meant to provide structures that are (and must be) able to conform to the world.[5]

Two remarks may be appropriate. The canonical presentation of the empirical theories as uninterpreted formal systems to which a physical interpretation is then provided by means of correspondence rules or something of this kind (a view advocated by Carnap and adopted by the great majority of analytical philosophers of science) is basically a *misrepresentation*, since it gives the primacy to the language, and considers the empirical "content" as a model of this language while, as we have seen, the dynamics of the empirical theories is precisely the opposite: the primacy goes to the empirical content, while those linguistic structures that constitute theories are

[5] A detailed discussion of these meanings of model in the empirical sciences is contained in chap. IX of Agazzi [1969].

accepted only to the extent that they are "models" (in the *second* sense clarified above) of the empirical content. As an excuse for this misrepresentation logical empiricists often said that their theory of science aimed simply at a "logical reconstruction" of the structure of scientific theories. Nevertheless we can say that this was indeed an inadequate logical reconstruction since it disregarded the basic difference we have discussed, and which is a logical difference.

Of this fact there are, moreover, certain not negligible symptoms: while it is characteristic of the formal theories that they admit a plurality of different concrete models (and they are even incapable of strictly characterizing one single model), in the domain of the empirical theories we find the *opposite* situation: a plurality of theories is always possible for describing and understanding a given empirical content (this is the well known fact that any theory is "underdetermined" by the empirical evidence). To put this difference in a sharp form (and using the concept of model in its *first* or model-theoretic sense): in the formal disciplines it is impossible to determine one single model for a theory; in the empirical disciplines it is impossible to determine one single theory for its intended model.

The expression "intended model" reminds us that also several mathematical theories have their intended model, and we had already the opportunity of considering this fact when we saw that formal systems fall short of characterizing such models. But another fact is also well-known: one and the same intuitive mathematical theory (such as arithmetic, geometry, set theory, and so on) can be formalized through several different axiomatizations, and this actually means that these different "formal theories" have the same intended model, and are, on the one hand, "legitimated" by the fact of admitting this model, but, on the other hand, are "underdetermined" by the model itself. This situation is indeed identical with that which we find in the case of the empirical sciences, and reflects the fact that the "working mathematician" investigates his domain of objects with an intellectual attitude very similar to that of the physicist, that is, he is able to "see" inside this domain, he tries to "discover" its features, he finds "counterexamples" to his conjectures that play the same role as the finding of empirical facts contradicting a physical hypothesis, and so on. All this shows the inadequacy of a purely "formalistic" conception of mathematics and, by the way, has encouraged the rise of that "empiricist" conception of mathematics that has known significant developments in the last decades.

Meaning, sense and reference

Almost all the differences (and difficulties) we have taken into consideration are rooted in a certain distinction (and in an insufficient attention paid

to it). This is the well-known distinction between *sense* and *reference*, with its allied distinction between *intension* and *extension*, whose contemporary clarification is chiefly due to Frege, but which was very familiar to ancient classical logic under the terminologies of "intentio" and "suppositio", and "comprehension" and "extension", respectively. Though initially applied to concepts, these distinctions were applied to linguistic *signs* already by medieval logicians, and therefore are in keeping also with the "linguistic turn" of our analytic philosophy. Roughly speaking, the sense and the intension constitute the "content of thought", the "idea", the "intellectual representation" attached to a linguistic expression, while the reference and the extension constitute the individual "object", or the class of objects, or the fact, that are attached to different kinds of linguistic expressions. As a consequence, we can say that the *meaning* of a linguistic expression contains two distinct, but not separated, aspects: sense and reference, or intension and extension. Analytic philosophy was affected by a deep empiricist mentality, which in particular implied an equally deep diffidence against "mental" entities. Therefore the obvious tendency was that of avoiding sense and intension (that were equated with "mental" subjective and "inscrutable" entities, in spite of Frege's efforts for avoiding any form of "psychologism" and distinguishing, for example, the objective content of thought and the subjective "representation" or *Vorstellung* of this thought). This tendency or preoccupation is explicitly present also in Tarski's "semantic" theory of truth, and accounts for its extensional approach to semantics, that has remained the cornerstone of model theory in mathematical logic.[6] The basic move of this approach consists in the pretension of "attaching" directly to a sign a "concrete" object (or class of objects), without the intermediation of the "mental" sense or intension. But precisely this pretension is highly problematic, as we will see in the sequel.

The unawareness of this difficulty, and of the ambiguities it contains, can already be found in the fact that, as we have seen, the fundamental operation for assigning a "model" to a language is called *interpretation*. Now, interpreting a linguistic expression is usually meant as "understanding" what it "means", that is, as grasping a certain intellectual content it expresses. Only as a consequence of this understanding one can try to see whether there are some objects that "exemplify" this intellectual content, and are possible "referents" of the expression. On the contrary, as we have seen, in the semantics of formal languages it is usually said that we interpret a sign on an individual object, or *in* a given domain of objects, a way of speaking that is clearly at variance with the usual meaning of "interpreting".

[6] Cf. Tarski [1944].

We must recognize, however, that such a confusion was already implicit in the way of understanding the commonplace statement that, in the modern formalistic view of mathematics, the terms occurring in the axioms, and hence the axioms themselves, had to be taken as linguistic expressions "devoid of any meaning". This way of speaking is used in the literature sometimes as equivalent to "devoid of any sense", and sometimes as "devoid of any reference". Therefore, it is no wonder that (as Tarski did), if we understand semantics as the enterprise of "giving a meaning" to linguistic expressions, this was understood as the effort of providing these expressions with referents. The open question, however, remains that of knowing whether it is possible to single out references without the intermediation of the sense. We know that Frege maintained that the sense is the "guide" for finding the referents, but we cannot rely on the simple authority of Frege, and will investigate this issue in some detail.

Models and language

When we try to understand a linguistic expression we estimate we succeeded in our effort if we can form us an intellectual picture of "what is said" in this expression or, equivalently, "what it intends to say". This is obviously very close to the Fregean notion of *sense*, but we can say that this is the most pregnant meaning of *model* as well. Therefore the model is primarily an intellectual construction, and has the characteristics of *intensionality*, first of all because it is intensional in the sense already explained above, and secondly because it is the result of that human capability that has been called *intentionality* from the time of the medieval philosophers down to Brentano, Husserl and the representatives of the phenomenological school. The meaning relations between intention and intension have been sufficiently discussed in the pertinent literature, and we need not detain us on them now. Still it is worth noting that these are also transparent in the expression "what it *intends* to say" by which we wanted to clarify what amounts to the understanding of a linguistic expression. But now an easy conclusion is at hand: if interpreting a linguistic expression consists in grasping a model, this cannot be whatever model, but precisely the *intended* model of the expression concerned. An evidence that this is really the aim of any interpretation is given by the fact that, if several different models can be attributed to an expression (that is, if it is open to different "interpretations"), we say that it is *ambiguous*. By the way, Frege's and other logicians' justification for advocating the use of artificial languages was precisely that common language is ambiguous. But what happens with the allegedly non-ambiguous formal systems? As we have seen, they are unable to admit just one single model, or of characterizing their intended model, but this clearly depends on having conceived the

model *extensionally* (with an additional incoherence: the "intended model" of which it is spoken in the context of mathematical logic is an unconsciously admitted intensional model since it corresponds to what is thought of in an "intuitive" theory).

Our considerations may suggest the wrong impression that the language comes first, and that a model is *then* looked for. The real situation is almost opposite. A language never is a pure collection of signs, but a system of signs that come to light already equipped with some sense: a sign is always created to communicate a sense. Therefore, the debated question of the primacy of thought over language, or of language over thought, is fully ill-posed. It is true that our thinking is continuously moulded by the language we use, but this happens precisely because the language is the expression of thoughts, and it is obvious that these thoughts may interact with other thoughts, and in such a way also determine the "formation" or modification or our thoughts.

But what about reference? Reference is, in general terms, the *exemplification* of thought, since it is constituted by "concrete" objects that actually realize the properties *encoded* by different thoughts. Two remarks must be immediately made: by this we do not suggest that thoughts come first, and referents are somehow produced by thoughts: we are considering here a pure structural relation, and will consider later its genetic aspect. Secondly, when we say that the referents are "concrete" objects, we do not suggest that they are "material" objects: their proper ontological status will again be investigated later. Our structural analysis amounts to a distinction that does not mean a separation, but that a distinction must be made is apparent from some easy examples. An expression such as "the golden mountain" is perfectly understandable without ambiguities (once its linguistic context is sufficiently determined); this means that we can form a "model" of it, in which its properties are encoded (the property of being a certain kind of geographical configuration of the earth surface, and of being totally made of gold). Yet no referent in our experience exemplifies this model. The same can be repeated of a lot of similar expressions, such as "the fountain of youth", "the square circle", "the perpetuum mobile", of which no exemplifications can be found in the domains where they should respectively be found. This way of presenting the situation is a rather good clarification of the (a little too vague) assertion of Frege, that the sense "guides" us in the search for the referent, and in particular clarifies that *looking* for the reference "depends" on sense, but *finding* it does not depend on the sense at all (we will see later what is needed in addition).

We are now in the position of understanding what actually are the extensional "models" considered in the semantics of formal languages and model theory: they properly are *exemplifications* not so much of linguistic expressions, but of the intensional models of such expressions. Calling them models is not a true abuse of language, but at least a too hasty convention. If one

understands that they cannot do but exemplifying the properties encoded by an intellectual model, one can better understand certain results of model theory. The plurality of the extensional models of a formal systems is simply the consequence of the very common fact that several objects can exemplify one single thought. Moreover, these examples exemplify their "encoding" thought only to a certain extent since this encoding necessarily includes only a finite number of properties, and the exemplifying objects usually contain a lot of other properties: this is why an explicit characterization inevitably says "too little" and "too much" with respect to any individual "concrete" model (or, better, exemplification). It is fully reasonable, and indeed inevitable, that the concrete objet contains several properties that are not captured by the intensional encoding that, nevertheless, it really exemplifies. This happens in all situations of our life, but is at the same time an easy explanation of the reasons for the semantic incompleteness of formal systems. That also other objects, with very different and even "undesired" properties, may exemplify a certain encoding that is exemplified by some very "acceptable" objects is again something we often come across, and this, again, is an easy explanation of another "limitation" of formal systems, that is, of their lack of categoricity.

Idealization, encoding and exemplifying

We shall now eliminate a wrong impression, possibly suggested by the foregoing considerations, that is similar to the one already overcome above, when we have seen that it is not correct to think that we "first" have a linguistic expression, and "then" interpret it by giving it a sense. A similar illusion would consist in thinking that we "first" have an intensional representation, an encoding thought, and "then" try to find some exemplification of it. This may occur occasionally (exactly as in the case of the interpretation of certain artificially constructed linguistic expressions), but is not what characterizes the "natural" situation. Indeed, thinking that we first have concepts, ideas or intellectual pictures, and then find concrete examples of them, is tantamount to adhering (consciously or unconsciously) to *innatism*, and this is notoriously a very weak philosophical position (it can be defended with some success only when it is understood in the sense that man has "innate capabilities", or even certain "innate structures" characterizing his mind, but not in the sense that man comes to life with "innate contents" of knowledge). Exactly like a linguistic sign comes to life as a means for communicating a certain thought with which it is originally linked, a thought (concept, idea, intellectual construction) comes to light as an encoding of certain salient properties found in a concrete object. This assertion, however, must not be understood in a chronological sense, that is, as if "first" comes the world,

then come the ideas, and then comes the language. This was indeed the way of thinking that characterized "modern" philosophy from Descartes to Kant, that was dominated by the epistemological tacit presupposition that we "know our ideas", and must therefore determine *from where* they come to our intellect (this is the famous problem of the "origin of the ideas", that concerned the philosophers of the said modern thinking, and produced several unsatisfactory solutions simply because it was an ill-posed question). Indeed there is no evidence nor argument for maintaining that ideas are "produced" (be it by God or the "external" world). Ideas are *present* in us, and they simply are the particular *way of being present* in us of the different objects. These are not "physically present" but, as ancient epistemology had already stressed, they are *intentionally* present, and this is why the world can by "physically external" to our body, but is at the same time within our mind in a *cognitive* sense.

In this presentation intentionality is recognized as the distinguishing capability of the mind in a general sense, a capability that is typical of certain *kinds* of existing beings. Sub-human animals are already equipped with this capability since they are endowed with perceptions that enable them to have the world "present" in them without being physically, but precisely "intentionally", present (in this sense we can speak of an "animal mind"). Perceptions are the intentional presence of single items of the world, but those living beings (typically human beings) that are endowed with a higher-level intentionality are also able to to have *general* representations of the world. It is customary to call *intellect* such a higher-level form of intentionality, and to call *thoughts* the way of being present to the intellect of the different features of the world. If perceiving the objects is the activity of the mind that corresponds to the first level of intentionality, *understanding* the objects corresponds to the second level of intentionality, and it provides those *intentional* representations that *encode* those general features of the objects of which the single objects are *exemplifications*.

Many things can become clear: if our purpose is that of *interpreting* a linguistic expression, we *factually* start from language, proceed to find intentional intellectual representations (thoughts) that are or can be attached to the linguistic expression, and may then look for objects that concretely exemplify the properties encoded by the thoughts concerned. But *genetically* (not chronologically) things are the other way around: the starting point are the objects, which are intentionally present in the mind under the form of thoughts, and these are expressed by means of language. The world of the *intentional* entities, however, has a kind of existence of its own: perceptions and thoughts can be associated, retained in memory, combined, such as to give rise to *intellectual constructions* (especially when a third capability of higher-level intentionality is available, that is, the capability of *reflecting* on

the abstract intentional entities, and submitting them to additional manipulations, of which *logical reasoning* is the most significant example). At this juncture, however, the genetic aspect mentioned above shows its importance: when such additional constructions are *intended* to have a *cognitive* function, we do not remain content with the fact that they are perfectly understandable or "intelligible", but are interested in finding *whether* they are actually *exemplified* in the world, or, even more often, in the domain of objects we *intended* to know.

The genetic process we have outlined may be, and has been, denoted by means of different denominations, depending on the intensional or abstract entities on which attention was focussed: abstraction, abduction and idealization are among the most usual terms adopted. The term idealization may be the most suitable since, on the one hand, it contains a reference to the "reality" that has been idealized but, on the other hand, points out that this reality was not simply "mirrored", but "understood" according to certain "ideal" patterns globally organized in a certain *model* of it. According to what has been said, this model encodes a display of properties that are meant to be actually *exemplified* by the reality that has been idealized, though only in a partial way, owing to the reasons already explained. It is important to stress that, since this model is the result of an intellectual construction, it is different from a simple organization of our perceptions: precisely because it aims at permitting us to *understand* and *explain* the perceived objects, it contains elements that are of a genuine intellectual nature and, nevertheless, are meant to be exemplified in the referent. Therefore the crucial problem appears to be that of clarifying the way of securing such a reference to the intended domain of objects.

Models and theories

Many ways have been proposed for characterizing a scientific theory, and we will certainly not spend time in overviewing them here. The most influential conception in contemporary philosophy of science has been proposed by logical empiricism, and chiefly consisted in considering a scientific *empirical* theory as a *linguistic* construction in which *sentences* are interrelated by means of formal *logical* dependences. A certain set of such sentences must have a "model" (in the model-theoretic sense explained at the beginning) granted by means of certain correspondence rules (that can be called rules of reference) while the other sentences are "justified" by the fact of permitting the logical deduction of such referential sentences. This kind of justification is notoriously not strong enough to permit the absolute solidity of the theory, and many developments of this awareness have produced several conceptions regarding the validity of scientific theories, their capability

of providing a reliable knowledge of reality, and so on. Again, we are not interested in overviewing such developments, and shall simply note that, as a consequence of several difficulties emerging from that general conception, this has been increasingly criticized and, under the label of "received view", also more or less explicitly abandoned and replaced.

Our position regarding this issue can be summarized as follows: (a) we can accept that theories are *linguistic* constructions (therefore, we do not believe that the linguistic point of view implies a misrepresentation of science, though being in need of several complementary considerations); (b) we reject the strict "sentential view" of theories to the extent that it maintains that theories are sets of sentences whose only connecting links are those of *logical deduction;* (c) we accept that theories try to describe linguistically an *intended model*, but consider the model according to its *intensional* characterization, and not in the extensional model-theoretic characterization of the current semantics; (d) we advocate the non-linguistic nature of *reference*, and try to relate models and reference through operational criteria of referentiality. Let us now develop a little these different points.

As I have explained at length in several papers and books,[7] any scientific discipline derives from the intention of investigating reality under a certain specific point of view, and in such a way determines its proper *domain of objects*. Therefore this domain, being obtained through a certain *conceptual cut* in which only a restricted amount of *attributes* of reality are considered, already consists in an *abstraction* or *idealization* of reality to which, however, it intentionally *refers* itself. As far as this reference remains a private affair (as it is the case, for example, for individual observations), the level of *intersubjectivity*, that is a distinguishing mark of scientific knowledge, is not attained. This is why every scientific discipline (be it in the field of the natural sciences, of mathematics, of the "human" sciences) is characterized by the fact of admitting a set of *standardized operational tools* that "link" certain fundamental *predicates* with the domain of reality under investigation, and in such a way allow for the determination of *concrete objects* that *exemplify* such predicates (or concepts), as well as the determination of concrete *facts* that *exemplify* certain *propositions* (propositions that, for this reasons, can be said to be "immediately true" of those referents). For these reasons it is appropriate to call "criteria of referentiality", as well as "criteria of immediate truth", such operational criteria. This is, however, only the starting point (from a conceptual, not from a chronological, point of view) of any scientific investigation, for it intrinsically contains a proposal of *understanding and explaining* the field of reality so circumscribed, and this is done through

[7] For the purposes of the present paper it is enough to consider the works of Agazzi already referred to here.

the construction of a *model* in the *intensional* sense described above. The model is, indeed, an *idealization* of the said field of reality, that offers a certain *interpretation* of it within the specific point of view adopted (therefore, the model has a basic *hermeneutic* nature). It does not simply contain the basic predicates that are directly linked with the "criteria of referentiality", but also contains other *theoretical* elements that are believed to *causally account* for the observed (better, for the *referred*) facts. In this sense, these models are the hermeneutic framework within which the *logical explanation* must take place.

Here is precisely the point where the *theory* appears: logical explanations must be explicitly and carefully stated, and this needs a *linguistic formulation:* this is the theory which, in the first place, has to be a linguistic elaboration of the intellectual picture contained in the model and, in the second place, must show explicitly how the model explains the domain of facts of which it is an idealization. Therefore, the theory is directly *the* theory *of its model*, and only indirectly a (particular) theory of its referents. This is not only in keeping with our previous considerations, in which we noted that a linguistic expression is directly attached to the conceptual content it "expresses", but justifies in a double sense the common notion of "intended model": the model is intended, first, because it is the content of an "intentional" act of the human capability of "intentionality" (that is, it is an *intensional* construction); secondly, it is intended also because it is that which the producer of a theory "wanted" or "intended" to express linguistically.[8]

Models and reference

The analytically inspired philosophy of science of past century has been dominated, as we have seen, by two hardly compatible influential views: radical empiricism and the "linguistic turn". The first has produced the extensional semantics we have already examined, and which essentially consisted in the effort of directly relating language and referents without resorting the the "mental" entities consisting in thoughts or representations. One could expect that, from the shortcomings of this one-sided stress laid upon the Fregean "reference", the other way should be seriously considered, that is, a stress laid upon the Fregean "sense". But this was not the case: the other way was inspired by the same mistrust against the "mental" realities, and the *meaning* of a linguistic expression was considered to result from the *linguistic context* alone, within an *holistic* perspective that was totally intralinguistic. As is clear, this semantics cannot be called "intensional" in any

[8] A detailed presentation of this view can be found in Agazzi [1992].

proper sense, for the intensionality of *sense*, relying upon its being the product of the human capability of intentionality, remains completely alien to this different approach. Some of the best known consequences of this strategy are the impossibility of attaining the "inscrutable" referents, the incommensurability of scientific theories, a fundamentally anti-realistic conception of science, and so on. We shall not discuss these consequences, that must honestly be considered as disappointing by whoever considers science as a form of knowledge (indeed the *only* genuine form of knowledge, according to the tenants of analytic philosophy). We will now try to show how, within the approach developed in the present paper, a much more satisfactory account can be given of the nature of science, and how several difficulties affecting the two different approaches of analytic philosophy of science simply disappear.

The starting point of our considerations is the absolutely obvious remark that whatever empirical science (and, more generally, any science intending to study a certain domain of objects of any particular ontological nature) has a fundamental genetical *referential root*, that is represented by the immediate reference to certain kinds of reality that are meant to exist within a certain admitted cultural background (or *background knowledge*). This referential link must not be lost already in that initial idealization that consists in considering reality from that *specific point of view* which (as we have seen) characterizes every scientific discipline. This requirement is satisfied by linking certain *basic predicates* of the said discipline with standardized criteria of referentiality that are of an *operational nature*. This means that they are different from the mere *perceptions* that were meant to be the basis for distinguishing observational and theoretical terms in the logico-empiricist philosophy of science. Nor do they consist in mere *ostensions* which are certainly too rudimentary (as Quine has stressed in his celebrated, but also simple-minded, examples related to his thesis of the radical impossibility of translation). The fact is that standardized operations are suitable for an *intersubjective* agreement in the *use* of certain notions (overcoming in such a way the privateness of observations and perceptions), and at the same time are really linked with thoughts (concepts and propositions) because it is included in the *intensional* sense of certain concepts, for example, that, in order to see whether they are exemplified, we must perform certain precise operations that are concrete, and *different from purely linguistic operations*.

When we propose an *intensional model* of our intended domain of objects, we admit in this model (as already stressed) several entities that are not accessible by means of the original criteria of referentiality, but are admitted because we think that they help us in understanding and explaining that domain of objects. How can we be confident that such entities exist (in the sense of existence we attribute to the original referents)? The answer to this question is rather simple: the confidence in their existence is proportional to our conviction that the proposed model *really* captures the deepest structure

of our domain, also in its not immediately apparent features, and *because of that* offers the most satisfactory understanding and explanation of it. In order this confidence to be more than a subjective faith, *rational arguments* must be provided, and this is possible only through the explicitation of the features of the model in the *linguistic formulation of a theory*. In such a way the non-immediately accessible entities and features of the model are expressed through special *terms* of the theory, that can now be appropriately called *theoretical* not because they are non-observational, but because they are characteristic of the theory, and formulated with the purpose of "theorizing" (that is, of understanding and explaining). Therefore, an empirical theory must contain *operational predicates*, as well as *theoretical predicates*, that are the constituents of those theoretical sentences that are the linguistic expressions of the fundamental features of the intensional model or idealization. What follows can then be easily accommodated in the standard view of empirical theories: from the theoretical sentences we try to *logically deduce* "empirical" sentences (that is, sentences that are "immediately referential") and, in particular, those immediately referential sentences that are *immediately true* of their referents (already known empirical true sentences, or predicted empirical true sentences). At this point a plausible consideration emerges: if we can be confident that a sentence from which several *true* consequences of different kinds have been *deduced* is itself *true*, it cannot but be true of its *referents*, and in such a way the existence of these referents is also granted.

A well-known objection to this "plausible consideration" is based on elementary logic: the fact of having correctly deduced even a great deal of true consequences from a premiss does not offer an absolute warranty of truth for this premiss. What is really implicated in this fact, however, does not concern truth properly, but *certainty:* we cannot be absolutely certain of the truth of the premisses, if we have deduced from them even a great deal of true consequences. But is this a serious objection? Do we have at all an *absolute* certainty about the truth of any sentence (even of the empirical ones)? In all the circumstances of our life we rely on certainties that simply mean the acceptance of those sentences whose truth appears to have been established *beyond any reasonable doubt*. Why should we be more extremistic with science, and pretend from it an *absolute* certainty, rather than a *reasonable* certainty (perhaps because a mythical capability of attaining an absolute certainty has been attributed to science with no real ground)? Therefore, if we can reasonably believe that a scientific theory is true, we are also reasonably entitled to believe that its *referents* exist, also regarding its theoretical concepts and features: this means that the *model* of which the theory is the linguistic expression is a good representation of the domain of objects investigated.

It is not accidental that we have never spoken of the "truth of a theory". This way of speaking is customary in the literature of the analytic philosophy of science, where theories are considered as "sets" of sentences uniquely related by formal logical links. In particular, it is very frequent that a theory be presented as the *logical conjunction* of all its sentences. According to this view, it is sufficient (at least in principle) to find one empirical sentence contradicting the conclusions logically derived within a theory for dismissing as *false* the whole theory. This never occurs in the history of science, and this fact has been often interpreted as the expression of a certain methodological laxism, or as a pragmatic attitude that prevents us from dismissing a theory if no better theory is available at a certain moment. The real reason is different. A theory is not just a set of sentences connected by formal logical links. Its unity and coherence directly depend on its being the linguistic expression of a *global* intensional model, on its being the explicitation of a given *Gestalt*. In this sense it is not correct to attribute it truth or falsity. When we find a *logical* incompatibility of the theory with an empirical evidence, it is not said that we loose our confidence in the model as such. In certain cases we may discover that our theory was an inaccurate linguistic elaboration of the model (and we slightly modify the theory in order to make it more adequate to its "intended model"), in other cases we may be led to certain refinements or little modifications of the model itself (in this case, the *referential* import of the model imposes those modifications); only in extreme cases we are led to radically change the model and, as a consequence, also the theory. The *holistic* crisis of a theory is only an exceptional event, and in any case is not produced by simply *logical* reasons: it is a *gestaltic* intensional appreciation of the model that which guides us in evaluating whether the discovered "anomalies" are so "strategic" as to imply the abandonment of the model, or whether they are only "marginal" and rather suggest some kind of accommodation of the model. We shall not insist on this point, that we have examined on other occasions.[9]

Reference and ontology

We shall devote a couple of observations to an issue we have treated at length elsewhere.[10] We have said the the referents are *concrete* entities that exemplify the intensional features encoded in a thought or system of thoughts. We have already said that such a concreteness must not be understood in a "materialistic" sense, and we have also provided some hints for the correct

[9] Cf. Agazzi [1992a].

[10] Cf. Agazzi [1997].

understanding of this issue, when we have said that the referents are found by means of *specific* operational criteria of referentiality, that are at the same time criteria of immediate truth, and that are specific, precisely, for every discipline. In the case of physics or other natural sciences, for example, these criteria consist in the performance of material observations, measurements, preparations, by means of concrete material instruments. But in the case of history, for example, these criteria (that enable the scientific community to admit something as a "historical *fact*") consist in the reading and evaluating of documents, inscriptions, coins, archaeological findings, whose referents may be physical persons, institutions, customs, ideas, events, processes, and so on. Similarly, when we say that it is *true* that Hector is a Trojan warrior in the *Iliades* we certainly refer to Hector, but not as a physically existing person, nor as an historically documented personage, but simply as someone having a "literary existence", of which we become acquainted by appropriate literary operations (reading the Homeric poem, studying the Greek language, etc.). An analogous discourse can be repeated as to the psychic referents, the mathematical objects, and so on. In a few words: the *reality* of the referents is something that depends on their being neither purely linguistic, nor purely intellectual constructions (they must be encountered by means of extralinguistic and, at least partially, extra-mental operations), but the *kind of reality* to which they belong (that is, their *ontological status*) depends on the operational criteria of referentiality by means of which we may come into contact with them. Of course, there is also a realm of purely intentional objects, but these are met precisely by pure intentional operations, such as introspections, conceptual analysis, memory, and so on, and there are also linguistic objects, that can be detected by the operations of linguistic analysis and all its related techniques. But even in these cases it is normal that we formulate certain "models" for interpreting and explaining such domains of objects, and then try to check these models be means of referential criteria that must show if these models are a good idealization of what they intended to understand and explain.

Department of Philosophy
University of Genoa

REFERENCES

[1969] Agazzi, E., *Temi e problemi di filosofia della fisica*, Milano, Manfredi, 1969, Roma, Abete, 1974 2nd ed.

[1976] Agazzi, E., The concept of empirical data. Proposal for an intensional semantics of empirical theories, in M. Przelecki, K. Szaniawski, R. Wójcicki (eds), *Formal Methods in the Methodology of Empirical Sciences*, Dordrecht, Reidel, 1976, pp. 143–157.

[1977] Agazzi, E., Les mathématiques comme théorie et comme langage, in *Langage et pensée mathématiques*, Luxembourg, 1977, pp. 15–32.

[1978] Agazzi, E., Le matematiche come teorie e come linguaggio, in *Epistemologia* 1(1978), pp. 165–182.

[1992a] Agazzi, E., Intelligibility, understanding and explanation in science, in C. Dilworth (ed.),*Idealization IV: Intelligibility in Science*, Amsterdam-Atlanta, Rodopi, 1992, pp. 25–46.

[1992b] Agazzi, E., Some philosophical implications of Gödel's theorem, in D. Miéville (ed.), *Kurt Gödel. Actes du Colloque, Neuchâtel 13-14-6-1991*, Neuchâtel, Centre de Recherches Sémiologiques, 1992, pp. 129–159.

[1994] Agazzi, E., On formalism, in G. Floistad (ed.), *Philosophical Problems Today*, Kluwer Acad. Publish., Dordrecht, 1994, vol. 1, pp. 74–137.

[1997] Agazzi, E., On the criteria for establishing the ontological status of different entities, in E. Agazzi (ed.), *Realism and Quantum Physics*, Amsterdam-Atlanta, Rodopi, 1997, pp. 40–73.

[1944] Tarski, A., The semantic conception of truth and the foundations of semantics, in *Philosophy and Phenomenological Research*, 1944, pp. 341–375.