Networks in contemporary philosophy of science: tracking the history of a theme between metaphor and structure

Valter Alnis Bezerra

Departmento de Filosofia / Faculdade de Filosofia, Letras e Ciências Humanas, and
Research Project "Genesis and meaning of technoscience" (Fapesp) / Institute of Advanced Studies
Universidade de São Paulo — Brasil
e-mail: bezerra@usp.br

Abstract. Our purpose in the present work is to survey some of the formulations that the theme of networks has received in contemporary philosophy of science over a period spanning twelve decades, from the end of the 19th century up to the present time. The proposal advanced herein is to interpret the evolution of this theme in four stages: first, one that goes from a metaphor or expressive image to a notion aspiring at implementation, but still having a virtual character, in the orthodox view of theories; second, a shift towards networks viewed as explicit, with the notion still having an enormous metaphorical power, as in Neurath and Quine; third, from this towards a concept amenable to implementation, as in Thagard, but still in search of an epistemological rationale, supposedly provided by BonJour and Dancy; and finally, as a notion fully endowed with a precise structuralist implementation, in the scope of the structuralist metatheory of Balzer-Moulines-Sneed. The interpretive framework adopted in this exercise in the history of philosophy of science is a modified version of Gerald Holton's thematic model, adapted to the history of philosophy instead of its original domain, the history of scientific ideas. In our version of the thematic model, we posit two main dimensions for the appraisal of any given philosophical concept, besides the thematic dimension: the metaphorical one and the dimension of its systematic implementation.

Keywords: Coherence. Structuralist metatheory. Otto Neurath. Paul Thagard. Quine. Theory-nets.

Introduction. The notion of a *network* is not new in the philosophy of science. It was repeatedly revisited in previous twelve decades by numerous authors in quite different contexts. We could take as our starting point for this survey — a something conventional one, to be sure, but not entirely arbitrary — pioneer Pierre Duhem in 1892, and then pass successively through iconoclast Otto Neurath in the 30s of the 20th century, the logical empiricism of the 50s, and Quine between the 50s and the 70s, up to more recent work by authors such as Thagard and Moulines (associated respectively with coherence theory of justification and structuralist metatheory). Along this trajectory, several facets of the network idea were explored, at the service of distinct images of science (and knowledge generally).

The concept of network is seductive in philosophy of science basically because it allows one to capture the idea that *scientific knowledge forms a tightly interconnected whole*, one in which parts of the system are not mutually independent, but somehow "depend on each other", and what happens in one part of the network has repercussions elsewhere. The challenge, then, is to achieve a precise formulation of this idea. Moreover, it seems plausible to suppose that the notion of a network, properly conceived, can become a useful tool to capture the richness, complexity, the multifaceted and variegated character that characterize scientific knowledge and rationality.

Our purpose in the present work is to survey some of the formulations that the theme of networks has received in contemporary philosophy of science over the aforementioned period of time. The proposal advanced herein is to interpret the evolution of this theme in four stages: first, one that goes from a metaphor or expressive image to a notion aspiring at implementation, but still having a virtual character, in the orthodox view of theories; second, a shift towards networks viewed as explicit, with the notion still having an enormous metaphorical power, as in Neurath and Quine; third, from this towards a concept amenable to implementation, as in Thagard, but still in search of an epistemological rationale, supposedly provided by BonJour and Dancy; and finally, as a notion fully endowed with a precise structuralist implementation, in the scope of the structuralist metatheory of Balzer-Moulines-Sneed.

As to the interpretative framework adopted in this exercise in history of philosophy of science, we work here within a modified version of Gerald Holton's thematic model in the history of (scientific) ideas (Holton [1982, 1988]). The basic distinction from which we start here is that between *theme* and *concept*, its central hypothesis being that the same philosophical theme can spread over an extended period of time (decades or even centuries), while being *instantiated* in different concepts in diverse periods and intellectual contexts. We take here two dimensions as being central to characterize any philosophical concept: on the one hand, there is the *metaphorical* dimension (which also has a *heuristic* role, as we shall argue below) and, on the other hand, there is the dimension of its *systematic implementation* — its role, say, constructive or operational within a philosophical system. In addition to the metaphorical and systematic dimensions, philosophical concepts also have a third dimension, the thematic one, in which is inscribed the degree of connection of a concept with the theme that underlies it. It is the through the thematic dimension that we are able to trace a significant degree of continuity between different concepts — that may seem profoundly different if taken at their face value, but, on closer inspection, reveal themselves to be ultimately distinct instantiations of the same theme.[1]

The importance of the metaphorical dimension of philosophical thought has been already subject of reflection by authors like Paul Thagard and Craig Beam, e.g. in "Epistemological metaphors and the nature of philosophy" [2004]. Metaphors provide certainly a major impulse for philosophical thought, and they are a powerful engine in the history of ideas, both in philosophy *and* science. The underlying notion of the idea of a metaphor, of "something being in place of something else" (cf. Palma [2004], esp. Chapts. 1 and 3, Sects. 3 and 4). Thus, it is akin to another strategy of thinking and language that is *analogy,* which is based on the idea

[1] Originally, Holton speaks of *scientific* concepts, and these have, in addition to thematic dimension, also an empirical or synthetic component and a formal or analytical component. Thus, a common theme becomes manifest at different times and contexts in the form of different concepts, with different empirical and formal amplitudes. For example, many concepts of force previous to the Scientific Revolution — not much susceptible to formalization, as this wasn't taken as a priority in the images of science hitherto prevalent — might be interpreted as having a formal component with smaller amplitude than those formulated, say, from the 17th century on — such as the Newtonian one, inserted into a highly developed mathematical system, meeting the demands of the new images of science that arose concurrently. Kepler's concept of force provides an interesting case, where the formal component was still incipient (in spite of the highly mathematical nature of the rest of his astronomical system) but the empirical component was quite pronounced (solar force as the driving cause of planetary motion in elliptical orbits). The main difference between the present view and that of Holton is that one does not identify in philosophical knowledge, in any obvious way, an empirical component (except indirectly, as discussed e.g. by Granger [1989]). This is why I think that, in the case of philosophy, the two dimensions that would correspond to Holton's "empirical" and "formal" reside in the metaphorical (or heuristic) dimension and the dimension of implementation.

of translating problems and solutions across *different domains* (about this we've previously brought forth some reflections, in Bezerra [2011]). The semantic displacement made possible when using a metaphor allows a widening of the field of potential applications of the concept in question, and suggests paths to be explored (paths that might have gone unnoticed beforehand). Therefore, we can say that a concept with a markedly metaphorical potential usually has a *high expressive power* (in that it brings together many different meanings) and also a particularly high *heuristic* power.

Our intention in this text is to show how the theme of a network in contemporary philosophy of science shifts and mutates over a period of 120 years, and becomes variously instantiated in concepts spread over a domain parametrized by the dimensions of metaphor and implementation, both of them present to varying proportions in the different concepts. It is noteworthy that we do not take here a conceptual change that emphasizes, say, the metaphorical aspect at the expense of weakening the systematic aspect — or vice versa — as being intrinsically more "valuable" or necessarily "progressive" (whatever meaning "progress" could have in philosophy). We understand that, in philosophy, every conceptual elaboration is part of a story, subject to a variety of contingencies, and occurs in response to specific problems, with certain programmatic ideals and particular values. Thus, an emphasis on the metaphorical aspect of a concept may be adequate for a particular philosophical perspective, while assigning a greater weight to the systematic-implementational aspect might be desirable within another perspective. If, however, there are certain cognitive and intellectual functions that a metaphor is able to meet — perhaps being the best way for doing that, in some cases — metaphorical thinking also faces certain limits. In philosophical analysis there are surely roles that can only be fulfilled by a "crystallization" of metaphor in terms of a precise, constructive, systematic implementation.

A word might be said about the periodization adopted in this text, choosing Pierre Duhem as a starting point for our survey of the network concept. A justification for this lies in the fact that we wish to define a "window" which corresponds, in general terms, to the existence of an agenda of problems, terminology and categories of analysis with which we are able to communicate today, synchronously, and without major difficulty. On the other hand, one notes also that this corresponds to the existence of philosophy of science in the contemporary sense, as an institutionalized discipline.

Duhem, the orthodox view, Hempel and Feigl. Pierre Duhem, when advancing the classical doctrine of theoretical holism in "Some thoughts on experimental physics" (Duhem [1989a], Part 1, Section 3), presents a reflection on systems of hypotheses (although he does not use the term "network", but instead talks of a "system", the idea is clear enough). Hypotheses, he says, do not face the experience alone, but together as a whole. The manner in which Duhem expressed his reflections in this article published in 1894 remains, I think, unsurpassed today, so it's worth quoting the words of the author in full:

> «A physicist sets out to demonstrate the inaccuracy of a proposition. In order to deduce from this proposition the prediction of a phenomenon, in order to establish the experience that must show whether this phenomenon occurs or not, to interpret the results of this experience and to see that the predicted phenomenon is not produced, he is not limited to employ the proposition in dispute, he moreover makes use of a whole set of theories, admitted by him without objection. The prediction of the phenomenon whose production should settle the debate does not follow from the contested proposition taken alone, but together with this whole set of theories. If the predicted phenomenon is not produced, it is not that proposition taken alone that is considered imperfect, it is the whole theoretical framework that the physicist made

> use of. [...] The only thing that experience tells us is that, amongst all the propositions that were used in order to predict this phenomenon and then to see that it did not occur, there is at least one error. But what it does not tell us is where this error lies. Does the physicist states that this error is contained precisely in the proposition that he wants to refute and not elsewhere? Yes, because he implicitly admits the accuracy of all other propositions used by him, and that trust is as good as its conclusion. [...]
>
> In summary, the physicist can never submit to the control of experience an isolated hypothesis but only a whole set of assumptions. When experiment is in disagreement with his predictions, it informs him that at least one of the hypotheses constituting this group is wrong and must be changed, but it tells him nothing about which one should be changed.» (Duhem [1989a], p. 93, 95)

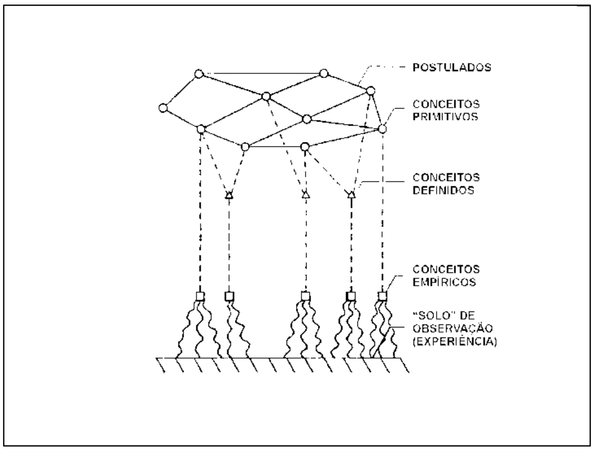
In this same text (Duhem [1989a], Part 2, section 1), and also in "Some reflections on physical theories", a 1892 article (Duhem [1989b], section 3), a similar concept appears, but Duhem refers there to hypotheses themselves as interrelations between concepts.

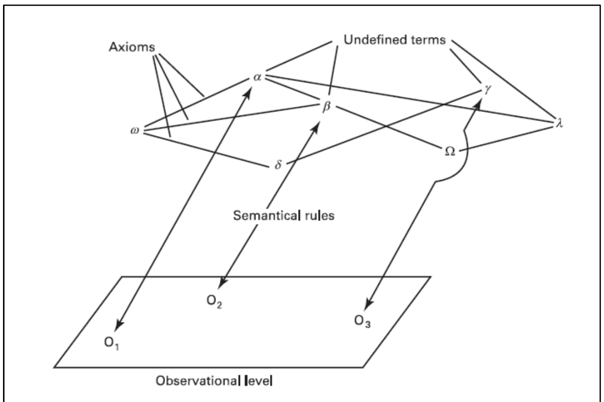
The ideal of deductive systematization has a long history, both in science and in philosophy, going back to Aristotle and the "axiomatic of the ancients" (to use Scholz's expression). This led De Jong and Betti [2010] to try conjoining together the most striking and enduring features of this image of science as a framework endowed with some degree of unity — the so-called "classical model of science" — in an *ideengeschichtlich* (history of ideas) register, as the authors say. Without delving here into the details of this model, we use this expression in an informal way (referring the reader to that work for details). It's worth recalling here, for example, the manner in which, in Euclid's *Elements,* each proposition refers to a cluster of previously established propositions: it's as we were seeing the constitution and unfolding of a network, step by step, before our eyes, as it were. Seminal works in the history of thought — be it in history of philosophy (such as Spinoza's *Ethics*), theology (such as Aquinas' *Summa Theologica*), natural philosophy (Archimedes' *On the equilibrium of planes*, Newton's *Principia Mathematica* and Descartes' *Principles of Philosophy*, to name but a few examples) — provide us with numerous examples of how that classical ideal was repeatedly pursued and updated throughout the history of thought.

In the 20th century, in the context of standard/received view of theories (also sometimes called the "statement-view" of theories), we can also find influential authors who employed or at least flirted with the notion of a network. Such authors were in many cases affiliated with logical empiricism — although the so-called "orthodox" conception is, in fact, not the work of a single individual, but rather the result of reconstruction and synthesis of piecemeal contributions from various authors. The central tenet of the orthodox view was the view that scientific knowledge is structured as a system of statements connected or linked by logical relations, particularly deductive relations, and this deductive system receives a partial interpretation. This idea allowed logical empiricists to benefit from the manifold achievements of mathematical logic in the early decades of the 20th century, as regards formal systems. The language to be used in theory reconstruction (at least according to the most widespread accounts) was usually first-order predicate calculus.

Carl Hempel, in one of his classic critical accounts of the orthodox view, the text published in 1950, "Problems and changes in the empiricist criterion of meaning" (Hempel [1959], section 5), also invokes the notion of network. Although Hempel uses the expression "all the logical relations [of a statement in an empiricist language] with the other statements", the network idea shows itself in a clear enough way. The kind of holism Hempel has in mind here is meaning holism, which, in principle, need not necessarily go along with testability holism (expressed in Duhem-Quine thesis). However, in this particular case, we must remember that, for logical empiricism, there was a close connection between meaning and testability, so what Hempel says about meaning has echoes with regard to testability also.

In this context one can also recall Herbert Feigl in his critical account of the received view, “The orthodox view of theories: Remarks in defense as well as critique” (Feigl [2004]). Much in the same way as John Losee would later do, in his *A Historical Introduction to the Philosophy of Science* [2001], Feigl alludes to the idea of a network, not refraining even to use diagrams, in order to illustrate theory structure according to the received view. The conceptual portion of the theory, represented by a network of relationships among primary concepts and defined ones, "hovers" above the "observational ground" that consists of empirical concepts, to which the conceptual part must refer, in order for contact with experience to be possible at all. Feigl’s and Losee’s iconic diagrams are reproduced below.





Hempel’s “Safety-Net” Image of Theories

As we all know, the list of pressing issues affecting the orthodox view was long. Among them was the practical impossibility of reconstructing any reasonably well developed theory in an axiomatic fashion close to that prescribed by the standard view. The few cases in which this was effectively done (such as Woodger’s reconstruction of classical genetics) boil down to a few, with limited and partial scope. Moreover, as most contemporary complex scientific theories usually presuppose extremely sophisticated mathematical apparatus, it would be necessary to precede every axiomatized theory with a ‘preface’ — orders of magnitude larger than the theory itself — which should contain all the relevant assumed mathematical theories, all of them axiomatized within the same language as the theory itself.

Neurath, *Ballungen*, protocols and coherence. Otto Neurath calls into question, in an incisive fashion, some well-entrenched theses of logical positivism about the structure and dynamics of science. Unlike his positivist colleagues, he advances the thesis that science is not structured in terms of “systems of clear statements” (i.e. clearly defined and limited as regards their principles, basis language, distinct subclasses of vocabulary, and deductive content). Rather, it would be organized in terms of much broader and more flexible structures, clusters which he calls "model-encyclopedias", and have the general outlook of "clusters" (in German *Ballungen).* We find a vivid description of this is in “Pseudo-rationalism of falsification" (Neurath [1935]), article published in the journal *Erkenntnis* that is a review of Popper’s *The logic of scientific discovery*, published in the previous year (Popper [2004]).

> «We, on the other hand, try to use models that give no scope at all to thoughts of an ideal of this kind [i.e. deductive systematization]. We start from masses of statements whose connection is only partly systematic, which we also discern [*überschauen*] only in part. Theories and single communications are placed side by

> side. While the scholar is working with the help of part of these masses of statements, supplementary additions [*Ergänzungen*] are made by others, which he is prepared to accept in principle without being quite certain what the logical consequences of this decision might be. The statements from the stock with which one really works use many vague terms, so that 'systems' can always be separated only as abstractions. The statements are linked to each other sometimes more closely, sometimes more loosely. The interlocked whole is not transparent, while systematic deductions are attempted at certain places. This situation is not open to the idea of an 'infinite regress', whereas Popper has to reject it especially in a certain connection (p. 90). If one wants to say that Popper starts from *model-systems*, one could say that we, on the other hand, start from *model-encyclopedias*; this would express from the outset that *systems of clean statements* are not put forward as the basis of our comments.» (Neurath [1935], p. 122, Section 2)

Given such networks as the clusters, one can say that the "field of force" that organizes them — that what gives cohesion to the cloud, we might say — is basically *coherence*, a view strongly defended by Neurath in "Sociology within a physicalist framework" [1931]. Neurath is a pioneer advocate of the coherence theory of justification. Although he occasionally refers to his conception as being about the *truth* of statements, it would be wrong to interpret it as a coherence theory of truth. The entire mechanism and dynamics outlined by Neurath refer undoubtedly to a conception of *epistemic justification* and methodological acceptability.

Neurath's coherentist view highlights the need to overcome a strict notion of coherence understood as mere logical consistency — and thus also points out a limitation of the axiomatic view — as shown in the following passage from his "Radical physicalism and the 'real world'":

> «Of course one tries to axiomatise science wherever possible.
>
> However, if [Thilo] Vogel [in "Zur Aussagentheorie Bemerkungen des radikalen Physikalismus" 1934] (p. 163) is of the opinion that one has to have the system axiomatised in order to have contradiction revealed, he overlooks the fact that in practice one proceeds much more clumsily and is mostly glad to have some contradiction pointed out *or* a greater number of conformities. It is precisely the history of physics that shows that our procedures are often quite consciously defective. It happens that occasionally two contradictory hypotheses about the same subject are used at two places with some degree of success. And still, one knows that in a more complete system only *one* hypothesis should be used throughout. We just resign ourselves to a moderate clarification in order to delete or accept statements later.» (Neurath [1934], p. 109, first italics mine)

This is a challenge that will be put to all coherentists from Neurath on: given that inconsistencies are bound to appear sooner or later in any sufficiently developed and complex theoretical system, how is it possible to define a richer and more flexible relation of coherence — in which context the presence of inconsistencies, albeit weakening the coherence, does not necessarily mean a descent into "epistemic hell" (to borrow the expression of Otavio Bueno [2008])?

Neurath's proposal is that the empirical basis of science should be constituted by the so-called *protocol statements*, whose format and properties he attempted to specify, e.g. in "Protocol sentences" (Neurath [1932]) and "The pseudo-rationalism of falsification" ([1935], section 7). While striving to articulate the notion of protocol statement, Neurath concedes that no part of the system of scientific knowledge is immune to revision — and this includes protocol statements themselves. Despite being an empiricist, Neurath admits, perhaps surprisingly, that even empirical statements (such as observational "recalcitrant" or "anomalous" results) can be discarded while retaining more well-entrenched theoretical assumptions ([1935], sec. 3) — although he admits that the rejection of a protocol is something that does not happen frequently ([1935], p 115). Thus, protocol statements are made fallible, and are subjected to a condition of coherence together with the remainder of the system. This introduction of fallibilism of the empirical basis is juxtaposed with the author's coherentism, and makes his system the target of various criticisms, e.g. by Popper and Schlick — who ask, in essence, what has been left of empiricism proper in this view.

Along with coherentism (a thesis about epistemic justification of beliefs), Neurath also advocates a form of *holism* (a thesis about the empirical testability of statements). For him, the whole system of science may be called into question — and he goes as far as stating that this includes not just all the *statements* (in particular protocol statements), but even the *methods* of science ([1935], sec. 3). The postulation of a sort of continuity, or lack of clear boundaries between, on the one hand, scientific clusters and, second, methodology and meta-science, can be viewed as an expression of the *naturalism* that characterizes Neurath's image of science.

I believe that, in post-Kuhn era, it is difficult not to interpret retrospectively the defense by Neurath of a form of flexible structuring of science in terms of "clusters", as a first step in the direction of *macrotheories* that characterize 20th-century philosophy of science. I think, in fact, that this does not amount to mere anachronism, and an answer in this direction can be sustained — although Kuhn does not explicitly recognize the inspiration of Neurath. More likely is that Neurath had envisaged the limitations of an excessively "theory-centric" philosophy of science, thus being an unrecognized pioneer of what we might call a "macrotheoretical turn" — a movement in which Kuhn himself participated, together with Lakatos, Laudan, Toulmin and others.

Quine. In Quine's "Two Dogmas of Empiricism" [1961], especially towards the final sections (5-6), we find the now familiar claims that "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (p. 41); that "the unit of empirical significance is the whole of science" (p. 42); and finally that "a 'recalcitrant' experience can... be accommodated by any of various alternative reevaluations in various alternative quarters of the total system" (p. 44). Most of all, we find the theses that "any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system" (p. 43) — echoing Neurath — and that "the considerations that guide [one] in warping [one's] scientific heritage to fit [one's] continuing sensory promptings are, where rational, pragmatic" (pp. 41-46). The network metaphor appears here in a clearer way than in any author so far: Quine [1961] employs terms like "network", "borders", "periphery", "field", "boundary conditions" and finally, the "web of science".

Quine's view of a system, network or web of beliefs is at one piece with his criticism of the analytic/synthetic dichotomy. For Quine what one has, rather than a dichotomy — as pressuposed by the traditional view — is a distinction of degree, instead of one of nature. The "more analytical" statements — metaphorically speaking, those more deeply imbricated within the system — would be less sensitive to the addition of new empirical information, while "more synthetic" statements — located near the "periphery" of the system — would be more directly sensitive to the impact of new information.

More than two decades later, in "On empirically equivalent systems of the world" [1975], Quine presents a new take on the same theme, writing:

> "Science is neither discontinuous or monolithic. It is variously jointed, and loose in the joints in various degrees. In the face of a recalcitrant observation we are free to choose what statements to revise and what ones to hold fast, and these alternatives will disrupt various stretches of scientific theory in various ways, varying in severity." [1975, p. 314-315]

In *The Web of Belief*, Quine and J. S. Ullian [1970] still defend the network metaphor. They reaffirm the thesis, already present in "Two dogmas", that belief assessment is done collectively, and not of each belief individually by itself. When a conflict arises among beliefs within the system, is not always easy to decide which way among the various alternative ways to reestablish consistency should be adopted; anyway, the fact is that our body of beliefs is, in general, perpetually in flux (Quine & Ullian [1970], pp. 6, 8).

However, beliefs derived from observation still constitute the *boundary conditions* of the system ([1970], Chapter 2) — a notion that has been employed in "Two dogmas". Put another way, they are the bottom of the system — although this cannot be the "rock bottom". (Recall that the notion of boundary condition has appeared in Quine [1961], and the metaphor of the blockhouse on a swamp, which never reaches the "rock", already shows up in Popper [2004], towards the end of Chapter 5.) The observational beliefs are peculiar, according to Quine and Ullian, in the sense that, while the other beliefs of the system face the "tribunal of observation" together, the observational face this "tribunal" individually ([1970], p 13). Although they are "almost foolproof" there still remains, however, a "a hint of fallibility" (pp. 16, 17). (The details concerning the status of observation sentences are worked out by Quine in "On empirically equivalent systems of the world" [1975].) In rare cases, the roles of "tug of observation" and "ship of theory" are exchanged, with the tug boat being pulled by the ship, so to speak, instead of pulling it (as it was supposed to usually happen).

Some such cases are solved summarily, assigning the conflicting observational beliefs to hallucinations or otherwise unknown interferences, thus discarding them. However, what about the cases in which the conflicting observational belief stubbornly persists? If the theory is well entrenched, and possess a certain "inertia", tension certainly arises. Quine and Ullian ([1970], p. 17) quote a long passage from Kuhn's *The Structure of Scientific Revolutions*, on the state of "crisis" that arises as a result of an accumulation of anomalies, without, however, the theory being abandoned before a replacing theory is available. One possibility, of course, is to *problematize* observation — what can be done *reinterpreting* it in terms of a modified theory. Observation still plays the role of boundary condition, albeit in a more sophisticated way. One must emphasize, however, that a theory that can be sustained *only* at the cost of repeated rejections of conflicting observations reveals itself to be an unreliable cognitive instrument ([1970], p. 18). The pressing problem that remains for Quine and Ullian is, then, to specify precisely the limits, mechanisms, and conditions for modifying the system.

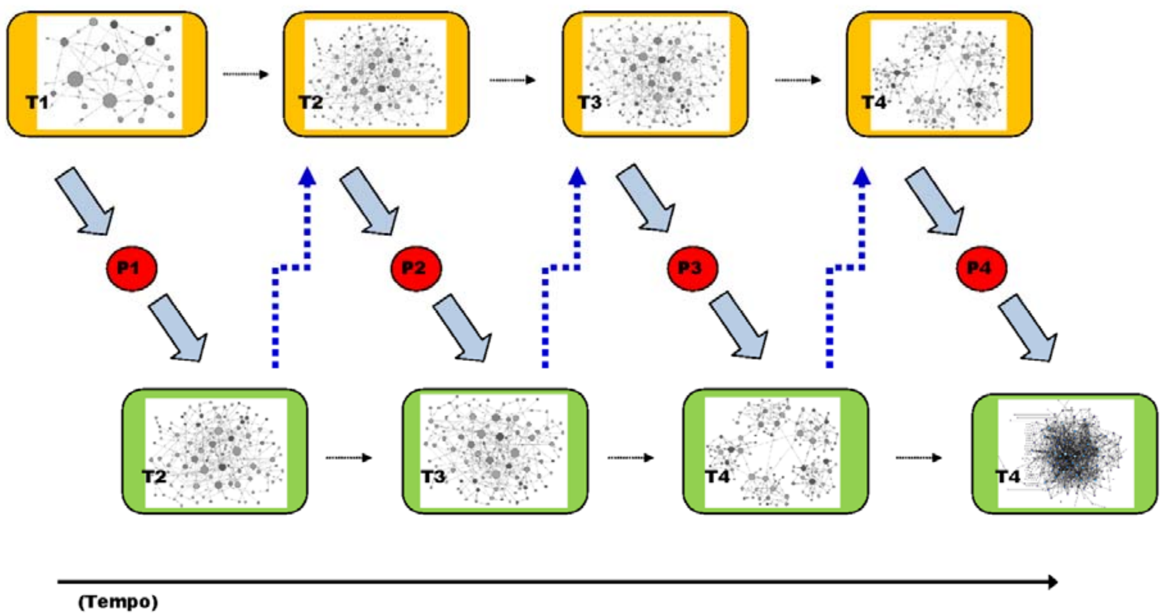
Networks, perturbations and the dynamics of problems. Images of science in terms of networks, such as Quine's and Neurath's, allow one to give, in a natural manner, a *rationale* for a diagnosis that had been expressed earlier by Hume and Kant (and more recently paraphrased in the 20th century by Nicholas Rescher [1999] in his "principle of propagation of questions"). Both classic authors from modern philosophy claim, in essence, that every solution to a problem creates new problems. First, consider Kant in *Prolegomena*:

> "Wherever we may begin, any answer given according to principles of experience always begets a new question which also requires an answer." (Kant [2004], § 57, p. 103)

This statement about knowledge of the world somehow resonates with Hume's dictum concerning philosophical analysis:

> "Each solution still gives rise to a new question as difficult as the foregoing, and leads us on to farther enquiries." (Hume [2000], IV.2, p. 28)

How could a contemporary approach in terms of networks lend support for the thesis anticipated by Hume and Kant? If problems do not exist in a vacuum, but are always identified as such against a certain previous theoretical background, and theories are viewed in general as attempted solutions to problems, then it is appropriate to imagine that, given a problem, its solution may require the redesign of structures or the introduction of new theoretical constructs. Such a (re-)formulation corresponds to a reconfiguration of our belief system for a given research domain. This reconfiguration can provide new opportunities but also creates new deviations from equilibrium and instabilities within the system — in other words, it creates new problems, which thereafter become part of the research agenda. This process is illustrated in the following diagram, where T_1, T_2, T_3, ... are successively reformed theories to account for problems, and P_1, P_2, P_3, ... are questions raised by new theories.



The coherence theory. The network metaphor has a natural affinity, one could say, with a coherentist view about justification. The network metaphor is important for coherentist epistemologies, first, because it places an emphasis on the internal mutual connections between beliefs that constitute the system, and, second, because it expresses the idea — very close to the coherentist mindset — that statements are not justified alone, but instead due to their pertaining to a system, which would be in turn justified due to connections distributed throughout the system. Coherentism already appears in an informal way in Neurath's and Quine's images of science; however, both of them still fail to provide a crucial element for giving this coherentist mechanism specificity and workability. In particular, they do not answer a question that remains open: *how is it possible to be a coherentist while still being an empiricist*? This challenge requires the development of a coherence theory of justification of a certain sort. This task was carried out, especially from the 80s of the 20th century onwards by authors such as Jonathan Dancy, Laurence Bonjour and Paul Thagard.

Insofar as coherence theory wants to talk about an *objective measure* of degrees of coherence, and about conditions that *increase* or *decrease* objectively this degree of coherence, it is essential that *networks be formulated in an explicit manner*. In the case of Paul Thagard's theory of coherence as constraint[2] satisfaction, networks are made fully explicit, and aren't virtual anymore (*Coherence in Thought and Action* [2000]). Interested as he is in connectionist models in order to implement the concept of coherence in scientific cognition and scientific change, by means of neural networks — going as far as running simulations of specific historical episodes — Thagard has it that the elements of the system in question (beliefs, propositions, etc.) are to be stated explicitly, and likewise with the various constraints (both positive and negative ones) between elements.

In the theory of coherence as constraint satisfaction, it is assumed that the network elements taken as the "empirical basis" are endowed with a certain initial plausibility, which causes them to be included in the system at the outset with a value greater than zero; the other network elements distribute themselves across a range of values of acceptability. The various connections distributed through the network can promote either reinforcement or weakening of the ties between elements, with different intensities. The system is then let to evolve based on these initial conditions, and it can be shown (cf. Thagard & Verbeugt [1998]) that in a significant class of cases, the system is able to converge towards an equilibrium state, which can be taken as representing the (provisional) state of knowledge at that time, under the prevailing conditions. It's worth noting that assigning positive initial values to certain "empirical" constraints, however, does not necessarily preclude that other constraints may eventually prevail over them later.

Thagard describes this version of the coherence theory as a "discriminating coherentism". However, even granting that he has, on the one hand, a framework for *describing* and *modeling* cognitive situations, on the other hand, however, he still does not provide a properly *epistemological* justification for this discriminating coherentism. An attempt to establish such an epistemological justification can be found in J. Dancy and the "first" L. BonJour (respectively in *Contemporary Epistemology*, Dancy [1985], and *The Structure of Empirical Knowledge*, BonJour [1985]), whose views I would like to describe as varieties of *asymmetric coherentism*. Such proposals are based primarily on the distinction between the *genetic* register and the *justificative* register as regards beliefs and belief systems, and also on the distinction between *antecedent warrant* and *subsequent warrant*. A belief could be *generated* by a process that is not completely inferential (i.e. it is somehow prompted by what is the case in the world), being thus endowed with a certain degree of antecedent warrant; in other words, it is "external" to the system in the genetic sense. And yet, this very same belief is to be *justified* in the same manner as all other beliefs are — namely, by virtue of its relations with the rest of the system — thus preserving monism (in a coherentist fashion) in the justificative register. As a result, then, coherence is endowed with an asymmetry that allows it to have an empirical character without falling back on foundationalism.

Interim balance. The metaphorical and the systematic. The need for a shift towards structures. In the hands of authors from Duhem up to Thagard, the theme of networks is variously worked, and instantiated in different philosophical concepts, with varying emphasis

[2] Incidentally, it should be noted that he does *not* use this term in the same way as structuralist metatheory does.

on both aspects — those of metaphor and implementation. We can distinguish three main stages in the process up to this point.

In the context of the standard view (and also, in part, already in Duhem), notwithstanding the urge on a programmatic level towards implementation (expressed in the demand to provide axiomatizations), networks are, nevertheless, still to be taken as inevitably virtual, in the sense that one can never completely exhaust the "logical content" of a set of axioms. (It is symptomatic of this that the standard view talks about *logical content*, *empirical content and confirmed confirmed* of theories — or, in Popper's case, *corroborated content* — though these "contents" are always left as open and indefinite classes, subject to expansion). On the standard view, a network is given *in an implicit way* when one gives *explicitly* a set of axioms. The network would consist, then, of all theorems that can be deduced from the axioms, i.e. all deductive consequences of the axioms, including all empirically testable consequences. Such a network would be potentially infinite, because, of course, one can never make explicit the *whole chart* of logical relations among axioms, theorems, lemmas, auxiliary hypotheses, etc. (After all, for example, one can always demonstrate a new, hitherto unknown theorem, say, within the system of Euclidean geometry.)

While networks are conceived, on the one hand, within the images of science associated with the classical model of science with a view to their *potential* implementation — not being able to “cross the border” due to them being given in an implicit, virtual manner only — in Neurath’s and Quine’s images of science, on the other hand, networks do not have a potential character anymore, and one could think about them in their entirety. The network metaphor is then pushed to its maximum expressive power. For Neurath, this allows setting up a critique of foundationalism and deductivism, and prepares the ground for an epistemological naturalism and methodological pluralism. For Quine, this becomes a vehicle for his critique of reductionism, and also a launching pad for his holism and naturalism. The gain achieved with these conceptions is extraordinary as regards expressive and heuristic power. This is important in its own right. However, since this still takes place mainly at the metaphorical level, a significant "degree of resolution" in the analysis is lacking. In these images of science, the notion of a network becomes somewhat amorphous, indistinct, without a "fine structure", as a result of the inability of these approaches to chart its internal structure. Neither there is a proposal for a definite mechanism governing the dynamics of its operation (in spite of Quine’s sparse comments on certain "pragmatic considerations" that would guide the process).

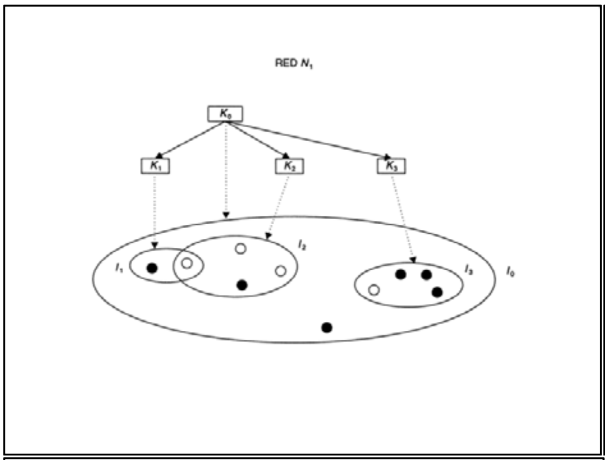
With Thagard, the metaphorical aspect ceases to be the dominant one (although it does not vanish entirely); there is, at the same time, a sustained concern with a well-defined *implementation* and a *mechanism* that gives operationality to the network proposal. However, along with an explicit project and an ability to describe and model cognitive situations — within a framework that, as the reader might suppose, has a flavor of epistemological naturalism to it — there still lacks a philosophical (or metaphilosophical) justification for his epistemology. In our view, such justification can be provided by other coherentist conceptions such as BonJour’s and Dancy’s. However, perhaps due to the fact both stem from more traditional epistemologies, a concern with formulating the networks in an explicit manner is lacking in them both.

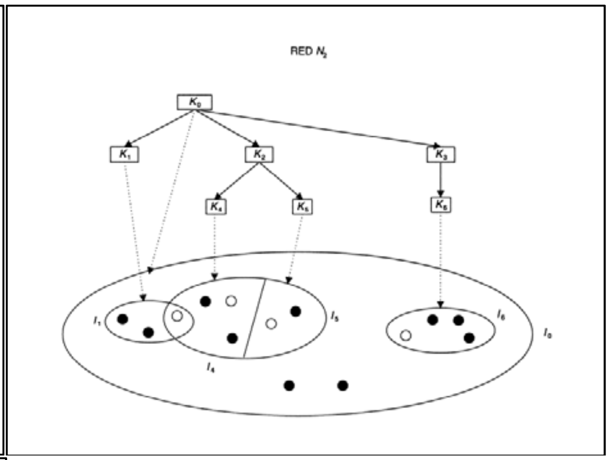
The challenge still stands, therefore, to explore within the conceptual space defined in the present study the possibility of formulating *an image of science based on the notion of network, that offers the prospect of explicit implementation, without losing the expressive and* *heuristic potential of the notion of a network taken as a metaphor, while, at the same time, providing a philosophical justification* (i.e. programmatic and critical). It is out contention that such alternative exists, and it is provided by structuralist metatheory, as a fourth stage in the development of the theme, to which we now turn.

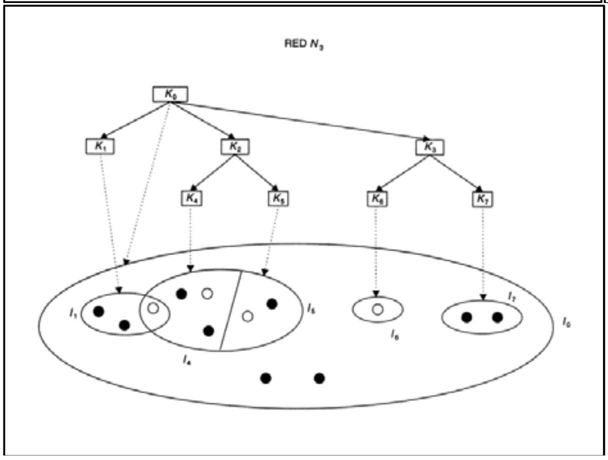
Structuralist metatheory. Structuralist metatheory (see e.g. Balzer, Moulines & Sneed [1987]; Balzer & Moulines (eds) [1996]; Díez & Moulines [1999]) was born out of: (a) a criticism of the standard view of theories (and, more generally, certain tenets of the classical model of science), as well as (b) from a new understanding of theoreticity of scientific concepts (and, related to it, a new take on the relationship between theory and experience), and, one could say, (c) it springs out of taking the Bourbakian notion of *structure* seriously. Scientific theories are no longer understood as classes of statements, but instead as classes of models. The relation of empirical adequacy is not analyzed in terms of testable deductive consequences ("empirical content"), but in terms of the embedding of certain structures (partial models) into other structures (full models). The interpretation of the formalism is not given by a system of correspondence rules anymore.

In structuralist metatheory, the main construct designed in order to capture the structure of scientific knowledge in their synchronic aspects, is the *theory-net*. Theory-nets are *explicit* networks: each component of the network (called a theory-element) is constructed with definite characteristics, has a precisely specified structure (in terms of classes of models related to each other by means of so-called constraints), and this structure is the result of well-defined choices. In other words, each theory-element in a theory-net is an *intentional object*, because it includes in its structure a *class of intended applications*. This is an indispensable component of the theory-element, which allows one to formulate precisely the *empirical claim* associated with the element (which can be formulated in a definite way). The connections within a theoretical network are introduced in an explicit and precise way, through a range of intertheory relations such as *specialization* and *theoretization*. A theory-net represents the theory in terms of a hierarchy of levels of specificity, by adding to the fundamental law (encoded in the basic theory-element) other special laws, present in the various specializations.

A theory-net is a synchronic formal construct, capturing the structure of the theories as “frozen” in time, so to speak; on the other hand, a *theory-evolution* corresponds to the unfolding of a theory-net in time, being thus a diachronic construct. The following diagrams taken from Diez & Moulines [1999] represent three stages in a generic case of theory-evolution. There is one theory-net (together with its attending classes of intended applications I_n) corresponding to each stage.







A key role is played in structuralist metatheory by the connections between elements of *different* theory-nets, called *intertheoretical links*, that supply a given theory T with terms from another theory T' — terms that are nontheoretical with respect to T (and join forces together with the "new", “autochthone”, properly T-theoretical terms).[3] The notion of T-theoricity and the presence of intertheoretical links allow one to clarify the sense in which the "empirical basis" of science can be both non-theoretical with respect to the theory whose empirical application or testing is in question, and at the same time theoretically impregnated by other theories. Intertheoretical links, by allowing the transfer of information between theoretical elements, contribute to the establishment of yet larger webs of connections between theory-nets, forming large-scale macro-theories called *theory-holons*.

Structuralism and coherence. How do structuralist theory-nets stand as regards the issue of epistemic coherence? It would seem that such connections as constraints (between models within an element) and intertheoretical links (between theoretical elements), together with the relations constitutive of theory-nets (i.e. specialization and theoretization) provide relationships of precisely the kind that would be typical of the coherentist view. It would seem that all these relationships contribute to the "cement" of coherence, that which makes the system of scientific knowledge "hold together." Finally, the more “macroscopic” structuralist view of the global structure of science in terms theory-holons, who get their cohesiveness from intertheoretical links, further enhances a vision of unified science as a large coherent

[3] The distinction between T-theoretical and T-nontheoretical terms departs significantly from the traditional views on the so-called "problem of theoretical terms." Besides not being considered as absolute, but always taken as relative to a given theoretical context, the distinction is not of an epistemological nature, but instead of a methodological character: in particular, it has nothing to do with the question of “observability”.

system, not unlike (in spirit, if not in the letter) the one advanced by Neurath in his proposal of a "Unified Science".

But at the same time, there would seem to exist a certain whiff of empiricist foundationalism, due to the privileged place of the "empirical claim" associated with each theory-element within a theory-net. Towards the end of *An architectonic for science* [1987, sect. VIII.6], Balzer, Moulines and Sneed undertake an analysis of how structuralist metatheory can be thought to stand with regard to the foundationalism / coherentism debate. They interpret the debate in terms of graphs of theory-elements connected by intertheoretical links, with the domain of intended applications of each theory-element being determined by concepts "imported" from other theoretical elements via interpretive links. The issue concerning foundationalism, then, is restated as the question of whether paths within the graph of theory-elements are all loop-free or if there are loops (more precisely, if there are no paths without loops). In other words, it becomes the question of the existence or nonexistence of certain theoretical elements that do not have interpretive links flowing from lower-level theories; the theoretical hierarchy "stops" with these, so to speak, and therefore they could play the role of "bed-rock elements". After a detailed analysis of two possibilities — either without loops (i.e. with terminal elements) or with loops (i.e. without terminal elements) — the authors finally come to the conclusion that, from a structuralist point of view, the coherentist position is more plausible than the foundationalist one.[4]

References

BALZER, Wolfgang & MOULINES, C. Ulises (eds) (1996). *Structuralist Theory of Science: Focal Issues, New Results*. Berlim: Walter de Gruyter, 1996.

BALZER, W., MOULINES, C. U., SNEED, J. D. [1987]. *An Architectonic for Science: The Structuralist Program*. Dordrecht: D. Reidel, 1987.

BONJOUR, L. [1985]. *The Structure of Empirical Knowledge*. Cambridge, Mass.: Harvard University Press, 1985.

BUENO, Otavio. [2008]. “Why inconsistency is not hell: Making room for inconsistency in science”. In: OLSSON, Erik (ed.) - *Knowledge and Inquiry: Essays on the Pragmatism of Isaac Levi* (Cambridge: Cambridge University Press, 2006), pp. 70-86.

BUENO, Otavio; FRENCH, Steven & LADYMAN, James (2002). “On representing the relationship between the mathematical and the empirical”. *Philosophy of Science* v. 69, pp. 497-518, 2002.

DANCY, J. [1985]. *Epistemologia contemporânea*. Lisbon: Edições 70, 1990. [Portuguese translation of *Contemporary Epistemology*.]

DE JONG, Willem & BETTI, Arianna [2010]. "The classical model of science: A millennia-old model of scientific rationality." *Synthese* **174** (2010), pp. 185-203.

DÍEZ, José A. & MOULINES, C. Ulises [1999]. *Fundamentos de filosofía de la ciencia*. 2nd. ed. Barcelona: Ariel, 1999. [In Spanish.]

---

[4] Nevertheless, it is worth noting that, in this analysis, one still stands in the realm of the constitution of knowledge that is under the aegis of structure, in other words, about what is already conceptualized in the form of theories. In the manner in which Balzer, Moulines and Sneed formulate the question, one does not take a position in epistemologically more traditional terms, say, about whether the empirical content of terminal elements has anything to do with a "direct access to the experience" — in other words, about what lies "below" the terminal elements. The issue about the fine structure of the theory-experience interface would probably better be discussed within the framework of an hierarchy of “models of data” a la Suppes-Woodward-Bueno (see Suppes [1962], Woodward [1989], Bueno et al [2002]).

DUHEM, P. [1989a]. “Algumas reflexões acerca da física experimental”. [Portuguese translation by N. de Carvalho of “Quelques réflexions au sujet de la physique expérimentale”, 1894.] *Ciência e Filosofia* no. 4 (1990), pp. 87-118.

DUHEM, Pierre [1989b]. “Algumas reflexões sobre as teorias físicas”. [Portuguese translation by M. Rocha e Silva and M. Fuchs of “Quelques réflexions au sujet des théories physiques”, 1892.] *Ciência e Filosofia* (São Paulo) no. 4 (1990), pp. 13-37.

FEIGL, Herbert [1970]. “The ‘orthodox’ view of theories: Remarks in defense as well as critique”. In: RADNER, M. & WINOKUR, S. (eds) - *Minnesota Studies in the Philosophy of Science, Vol. IV - Analyses of Theories and Methods of Physics and Psychology*, p. 3-16. Minneapolis: University of Minnesota Press, 1970.

GRANGER, Gilles-Gaston [1989]. *Por um conhecimento filosófico*. [Portuguese translation by C. M. Cesar and L. M. Cesar of *Pour la coinaissance philosophique* (Ed. Odile Jacob, 1988).] Campinas, SP: Papirus, 1989. [In Portuguese.]

HEMPEL, Carl G. [1959]. "The empiricist criterion of meaning." In: AYER, A. J. (ed) - *Logical Positivism*, p. 108-129. New York: Free Press, 1959.

HOLTON, Gerald [1988]. *Thematic origins of scientific thought*. Cambridge: Harvard University Press, 1988.

HUME, D. [2004]. *An enquiry concerning human understanding – A critical edition*. Ed. by Tom Beauchamp. Oxford: Clarendon Press, 2000**.**

KANT, I. [2004]. *Prolegomena to any Future Metaphysics That Will be Able to Come Forward as Science, with Selections from the Critique of Pure Reason*. Trans. Gary Hatfield. Cambridge: Cambridge University Press, 2004.

LOSEE, John [2001]. *A Historical Introduction to the Philosophy of Science*. 4th ed. Oxford: Oxford University Press, 2001.

NEURATH, O. [1931]. "Sociology in the framework of physicalism." In: Neurath, O. [1983], p. 58-90.

NEURATH, O. [1932]. "Protocol statements". In: Neurath, O. [1983], p. 91-99.

NEURATH, O. [1934]. "Radical physicalism and the 'real world'". In: Neurath, O. [1983], p. 100-114.

NEURATH, O. [1935]. "Pseudorationalism of falsification". In: Neurath, O. [1983], p. 121-131.

NEURATH, O. [1983]. *Philosophical Papers 1913-1946*. Trans. and ed. by R. S. Cohen and Marie Neurath. Dordrecht: D. Reidel, 1983.

PALMA, Héctor A. [2004]. *Metáforas en la evolución de las ciencias*. Buenos Aires: Jorge Baudino Ediciones, 2004. [In Spanish.]

POPPER, K. R. [2004]. *The logic of scientific research*. Portuguese trans. by L. Hegenberg and O. S. Motta. São Paulo: Cultrix, 2004.

QUINE, W. [1961]. "Two dogmas of empiricism." In: *From a logical point of view*, p. 20-46. 2nd ed. New York: Harper & Row, 1961.

QUINE, W. [1975]. "On empirically equivalent systems of the world." *Erkenntnis* **9** (1975), pp. 313-328.

QUINE, W. V. & ULLIAN, J. S. [1970]. *The Web of Belief*. New York: Random House, 1970.

RESCHER, Nicholas [1999]. *The Limits of Science*. [Revised Edition.] Pittsburgh: University of Pittsburgh Press, 1999.

SUPPES, Patrick (1962). “Models of data”. In: E. Nagel / P. Suppes / A. Tarski (eds) – *Logic, Methodology and Philosophy of Science – Proceedings of the 1960 International Congress*, pp. 252-261. Stanford, CA: Stanford University Press, 1962.

THAGARD, P. [2000]. *Coherence in Thought and Action*. Cambridge, Mass.: MIT Press, 2000.

THAGARD, Paul & BEAM, Craig [2004]. “Epistemological metaphors and the nature of philosophy”. *Metaphilosophy*, v. 35, n. 4, pp. 504-516, 2004.

THAGARD, Paul & VERBEURGT, Karsten [1998]. “Coherence as constraint satisfaction”. *Cognitive Science* **22**(1)(1998), pp. 1-24.

WOODWARD, James (1989). “Data and phenomena”. *Synthese* v. 79, n. 3, 1989.