From *Annali del Dipartimento di Filosofia*, IX [1993],
Università di Firenze, Olschki, Firenze 1994, pp. 225-334.

Alberto Peruzzi

From Kant to entwined Naturalism

This work received the financial support of a ‘40%’ research fund from MURST (Italy).

a mia madre

CONTENTS

PREFACE

Sag an, du kleiner gro□er Mann,
Der Turm, von dem dein Blick so vornehm niederschauet,
Wovon ist er, worauf ist er erbauet?
Wie kannst du selbst hinauf - und seine kahlen Höhn,
Wozu sind sie dir nütz, als in das Tal zu sehn?

The meaning of these verses by Schiller is only apparently univocal. Progressively, I felt they expressed a tension long hidden through philosophical subtelties. Could contemporary evolutionary cosmology be reconciled with Kant's Copernican Revolution? I assign a primary role to epistemology in framing a satisfactory theory of logic and language. But the primacy of that role is to be grounded in a naturalistic setting. Such conviction has numerous and deep-seated feedback—effects concerning semantics and epistemology. The need to make explicit at least a few of them gave rise to this essay.

Its main task is to provide a philosophical framework for theoretical research in the semantics of natural language and for the development of mathematical models of the mind. To this end an appeal is made, at a small number of decisive points, to the resources provided by the approach to logic and language based on category theory. Through the discussion of various alternative views of the relationship between mind, language and world, what is eventually offered is a naturalistic perspective on epistemological and semantical problems: one that sums up a long path in my intellectual life, as is mirrored in the title.

I do not conceal from myself that many readers will probably consider the architectural form of the essay as either misleading (for example, by claiming that the *proper* sense of 'transcendental' is distorted), or not pertinent to the body of problems with which epistemology (as formulated in the linguistic setting propagated by analytic philosophy) *properly* deals; conversely, others will find it insufficiently resolute in severing the links with traditional standpoints confining philosophy to a meta-discourse about the zoo of isms. Since it is beside the point to try in the Preface to defend the overall aim and the main theses of this essay, entering into specific aspects of the arguments involved, here I wish only to say something about the general sense of the journey alluded to in the title.

As regards the starting point, Kant is chosen as the paradigmatic figure whose work sums up and coherently systematizes the epistemological direction typical of modern

philosophy. Those aspects of his position considered here are intentionally focused on in isolation from their context. This expedient may be regarded as somewhat unfair, but it is intended to highlight and confirm the persistence of some aspects of the Kantian background behind many styles of reasoning in contemporary philosophy.

As for the terminus of the journey, there are many kinds of naturalistic views in epistemology; they address different problems and use different theoretical frameworks to solve them. Although common points among such diverse positions can be identified *per viam negationis*, I am not concerned here with listing them, comparing their force or finding some greatest common denominator, since any such list would tend to obscure the fact that naturalism is not intended here as just another ism within philosophy as traditionally conceived. Rather, my strategy will consist in making my conception of naturalism progressively emerge through a process of ideographic emulsion, in the style that was Giulio Preti's[1].

The positive components characterizing the approach developed in the following are: i) the adoption of phenomenological strategies, mostly related to the interface of language and perception, ii) the appeal to a realistic conception of truth and meaning, iii) the use of category theoretic notions at points where the usual formalization (within set theory or artificial intelligence), intended to tie the analysis of language to theories of mind, turn out to be inadequate, iv) the proposal of a many-layered parallelism with regard to the modular architecture of complex cognitive systems, together with the abandonment of the idea that thought is a sort of language. None of these ingredients is really independent of the others; they all enter the scene as axes of one and the same frame, centered on the conception that philosophy forms a 'modal braid' between natural and artificial intelligence.

If modernism corresponds to the diffusion of a foundational project, autonomous from any genetic consideration, the perspective suggested by the Presocratic embedding of epistemological investigation within philosophy of nature is pre-modern, rather than post-modern. In terms of the dichotomy between lightness and heaviness, made famous by Milan Kundera, the core of this return to the Presocratics is not a nihilistic sense of axiological heaviness present in our culture, one caught up in an unending game of self-exegesis, but the sense of lightness of an opening to natural roots, through the transparency of mathematical form.

My task here is that of a ferryman: to bring the reader out of the problem grid of modern epistemology (and contemporary philosophy of language too, inasmuch it is simply reformulating classical epistemological questions in linguistic form), showing that both the relativist and the pragmatist options are vitiated by the same idealistic imprint, and

---

[1] Unfortunately, the contribution made by Preti to twentieth century philosophy is still not available in English, although his research was deeply concerned with fundamental topics of analytic philosophy and brought him to original solutions. See the references.

that irrationalist and metaphysical exits can be avoided. When this task is completed, a new domain of investigation is opened to the eyes. Here the ferry stops, and the essay ends. Perhaps it is just for this reason the following pages belong to philosophy, *i.e.*, for having displayed a path: the stoics said that the philosopher is *hodopoietikòs,* «he who is able to open paths».

The first draft of the essay was made in 1989, but it was only some months ago I had the chance to think again about its completion. I considered the opportunity of bridging all the gaps and making all the needed steps explicit, but I soon realized that this would increase the length of the essay excessively, and could also have the effect of losing sight of the unity underlying the three chapters; what is important from the point of view of academic presentation may not contribute to the philosophical substance. Consequently, I decided simply to revise as least as possible of the original project, even though I am perfectly conscious that this choice will leave some arguments cursory. The overall sense of the essay is easy to grasp, the real difficulties being rather within the new area of investigation whose borders the essay is intended to sketch. The actual presentation is, in effect, the result of its very genesis: it started as a parallel series of 'frame' remarks pointing to a naturalistic conception of mind and language, while working at the border between category theory and logic, until I realised that there was a common standpoint underlying all those remarks: which I name *entwined naturalism* (the reasons of the qualification will appear in the course of the argument, see in particular p.XX).

I am indebted to the referees for their valuable criticisms, and to all others who have read the typescript, making me more and more conscious of the wide range of philosophical perspectives that are at odds with the view suggested here. Finally, I wish to thank Mike Wright for his suggestions in improving the English form and for his stimulating comments on the content of this essay.

Chapter 1. Five Kantian Contra-themes

For a long time the philosophy of language and the philosophy of mind have remained relatively autonomous subjects. Their autonomy has been reinforced by the separation between logic and psychology, which provided the respective tools for theories of language and theories of mind. The separation of a normative science such as logic from a descriptive one such as psychology was the result of a tortuous history, and one through which, while some philosophical problems were rendered susceptible to fruitful treatment thanks to that separation, others were at the same time re-situated in such a way as to make them effectively unsolvable. In the light of recent developments in psycholinguistics and artificial intelligence, the appeasing dualism of the cognitive and logical aspects is going to be shattered, and this will induce a change of perspective on many traditional topics in philosophy of language and philosophy of mind. In the present mobile landscape, a new appraisal of the contribution provided by the phenomenological tradition turns out to lend support to a non-reductive view of the logic/mind link, and to a rethinking in depth of the epistemological bases of semantics. A persisting obstacle to such an appraisal is the fact that, with rare exceptions, the phenomenological aspects of logic and language have been studied within the framework of either a Platonist conception (some kind of formal *intuition* grasps ideas and principles) or an idealistic one (the constructivity of thought moulds a reality internal to it). In both cases a dualism of form and content is seen as basic to the mind's architecture. The difficulties embodied in the former conception are easily recognized. Those of the latter are not. This contrast furnishes the subject matter of the first two chapters. Their character is critical as well as exploratory. The object of the criticism is those aspects of the Kantian legacy which I regard as supplying the basic underlying tenets of both functionalism and holism in the philosophy of mind and the theory of meaning. The recognition of phenomenological elements is bound up with the appeal to the mathematical tools provided by category theory. It fits within a naturalistic view that looks on philosophy as an activity internal to the body of scientific research.

§ 1.1. Reason as separated

Contemporary thought presents many different uses, or revisions of Kantian ideas and theses relevant to the language/mind connection. They may be seen as balancing the distance from the Kantian *humus* taken by logical empiricism, structuralism, and the so-called «new philosophy of science». A second-wave *Zurück zu Kant* has stemmed from linguistics, cognitive science and philosophy of mind, compensating for the lack, on Kant's part, of a 'critical' analysis of universals of language, while a parallel lack, no less serious, namely that of a 'critique' of logic, has been compensated for by intuitionism.

These logico-linguistic developments might suggest that, on the 'formal' side, the main gaps in Kantian epistemology have been filled. However, even while leaving aside problems concerning the architecture of scientific knowledge (and in particular the philosophy of geometry and physics), there remained harsh criticisms that attacked Kant's commitments to anthropological biases and more specifically to psychologism: as is well known, analytic philosophy and logical positivism purported to avoid this public danger by re-defining epistemic issues in terms of their linguistic counterparts, and by stressing the role of conventions. Yet this was not the only way out: Konrad Lorenz's interpretation of comparative ethology, and Giulio Preti's interpretation of the transcendental subject as a dynamic network of (not only linguistic) historically rooted categories exemplify different approaches. Such approaches set out to maintain some basic tenets of the Kantian heritage, without necessarily following the course taken by those who executed the linguistic turn, conceived of not only as a methodological strategy, but also as reflecting a dualistic ontology (logical objects *vs* mental acts).

The Husserl Renaissance that has taken place since the middle of Nineteen Seventies has contributed to emphasizing the link between the phenomenological view of language and logic, intuitionistic constructivism in proof theory, and various aspects of transcendental idealism. Yet there are many points where the distance of Husserl from Kant is worth noticing: for instance, Kant did not admit any intellectual intuition. While this may have had the advantage of eliminating the ticklish question of how we can grasp a content of thought whose objectivity is external to the mind, nevertheless it blocked investigation of the vast phenomenology of insight, concerning non-linguistic phenomena. For Kant, reason has no content, only forms.

In addition, Kant admits only an active synthesis in the self-organization of perceptual materials, whereas Husserl insists on the role of passive synthesis too. This can lead, by extension, to the recognition that certain forms are neither intrinsically mental nor intrinsically physical but emerge by synergetic effects in suitable 'environments'. In effect, in *Noema* I traced back the self-stabilization of basic cognitive experiences to the existence of constraints acting both in mental and in physical epigenetic landscapes. Parallel

constraints of this character had been already conjectured by D'Arcy Thompson, Gestalt psychologists and more recently by René Thom. This is a line of thought more deeply separated from that of Kant than is the position of Karl Popper, who still concedes that the human intellect imposes its laws on data, even though he criticizes Kant for having held that the conditions for the success of this imposition lie within mind — hence, science would not only be made possible but would even follow (necessarily) from our demiurgic cognitive architecture. The reason why the previous claim concerning parallel constraints deviates from Kant far more than Popper does is that it implies that empirical data are not *amorphous* at all, and that the constraints on their possible forms are not a gift of the mind. This position does not necessarily lend itself either to a Feyerabendian conclusion regarding the variable theory-ladeness of experience or to some version of holism: it points rather to a search for the 'primitive' structures underlying semantic experience and to the natural conditions (of cognitive import) which underpin the ability to trace the connective tissue of global features to which holists assign such a key role.

It is true that, by opposing the empiricists' associationism, Kant focuses attention on *schemes*, whose function consists just in mediating intuition and thought: through schemes, not only can the synthetic unity of the sensitive manifold conform to mathematical principles, but imagination becomes indispensable to ordinary perception: the actual is pervaded by the possible. Nevertheless, sensibility and intellect remain sharply distinct and for this very reason the Kantian criticism of metaphysics is possible. On this point logical empiricists follow Kant very closely: the improper use of reason (leading to antinomies) is simply traced back, through language, to the absence of epistemic (empirical) meaning. Both strategies *exploit* the purely formal character of reason. Separation between sensibility and intellect is, on the contrary, prevented by the local view of phenomenological *epoché* advanced in Peruzzi (1988): even the freest use of categories is never totally devoid of empirical content if our understanding of any judgment is to be possible — ch. 3 below makes explicit some of the reasons for this thesis. It follows that metaphysics cannot pretend to transcend experience, being subject to the usual conditions for any presumed knowledge. (This has direct consequences for philosophy of science, but they will not be analyzed here.)

(Neo-)Kantians and logical empiricists share the assumption that structure and existence, or form and content, can be and have to be globally separated in our overall picture of the world. This is just the conclusion which I contest, holding it to be phenomenologically (and cosmologically) flawed by inconsistency, separation being possible only locally. For (neo-)Kantians and logical empiricists claim that conditions of judgment transcend the domain of (mental and physical) facts: what matters is the *function*, and not the nature, of such conditions within the system of knowledge. Here lies the link with both functionalism and holism, hence John Searle's attack on the syntacticism of

artificial intelligence (AI), and Donald Davidson's charge against Quine of preserving the third dogma of empiricism: «the very idea of a conceptual scheme». The roots of this entire debate can be traced back to the ancestral dichotomy between form and content.

In fact, Kant presents his critical method as circumventing the tendency to substantialize the pure Ego as a Soul, but his remaining functional Ego is a rose without a stem: it could have nothing to do with humans, with their language, or with their Earthly environment. One is simply faced with the tasks of a generic intelligence, without being able to specify which deictic parameters are really negligible, and how other, architectural, parameters can vary under suitable conditions. In claiming this, it is not my intention to suggest that the dimension of the possible is destined to collapse; what is needed is rather a finer description of how essential such a dimension is for understanding the way the world is, *and vice versa*. Thus, let us try to establish some qualifying points at which my deviation from Kant is more emphatic.

Kant tells us that reason is not a passive pupil of experience, but a judge who obliges witnesses to answer his questions. Fortunately, we are not in the detached and aseptic position suggested by this legal metaphor. Reason is also experimental, and this implies interaction: the judge can find himself in the dock. But how can reason be accused, or defended, unless by itself? The point is that reason is neither (being purely formal) a cosmic judge, nor a Humean slave of the passions. For thought does not occupy a uniformly higher sphere than perception and emotion and there is a flow of contents within reason, a flow which is not unidirectional; and more than that, the direction and the content of this flow are not at all transparent, as Freud realized. [2]

With the institution of the Kantian tribunal, a radical divorce of philosophy from science occurs. Philosophy is conceived as an autonomous analysis of the conditions on which the possibility of science is grounded, and becomes a separate field of research —apparently specialized. The linguistic version of the *Erkenntnisproblem* preserved this autonomy for philosophy (as did so called «continental philosophy») with the result that both the conception of philosophy as a metatheory of science and its conception as providing the ultimate clarification of the preconditions of inquiry in general alike led to the enstrangement of the philosopher from any empirical research, as well as to standardized images of science. Whereas, not only are the search for truth and our analysis of the notion of truth on the same footing (Quine *docet*) but they are also on the same

---

[2] One of the strongest blows to classical rationalism did not come from the crisis of its bulwarks at the turn of the century (applied non-Euclidean geometry, logical conventionalism, etc.), but rather from Freud's attack on the crystalline transparency of self-consciousness. In effect, Chomsky's neo-rationalism in linguistics adopted from the start the ‘opaque’ character of language universals.

theoretical level (as Quine *non docet*). Once reference is investigated as a natural phenomenon, truth can be also.

On the other hand, such Kantian claims as i) that the determination of our own existence in time is possible only through the existence of real external objects, and ii) that the consciousness of our own existence is at the same time consciousness of the existence of *other* things external to us, can perfectly well be inserted into a naturalistic epistemology: they act as moderating elements with respect to the promise held out by Husserl's global *epoché*, and follow more faithfully than Kant the spirit of Rousseau, allowing us to conceive of philosophy not as a way to place man over nature but to bring him back to his natural status.

One could object that Kant is perfectly conscious of a realistic attitude *within* ordinary experience. Kant speaks of his «empirical realism» as ultimately consisting of the claim that something real in space corresponds to our intuition; yet space is an aspect of the Subject, thus it is the objectivity of knowledge, not the autonomous existence of an extended world, that can be recognized through his Copernican Revolution. The world is thought of as an onion: peel away any descriptive/conceptual leaf and nothing remains. «How could it be otherwise?» — so the idealist story goes.

Finally, the place Kant assigns to transcendental investigation is misleading: one can avoid skeptic consequences of the naturalistic fallacy, by imposing constraints on the form/content link. When Kant speaks of the two marvels, the starry sky above us and the moral law within us, he is speaking, in effect, of the two sides of one and the same marvel. For we *are* stardust: our existence depends on iron atoms in our body, which come from stars that, billions of years ago, expelled their hulls in a Supernova phase.

Curiously, for contemporary idealists, the failure of the Kantian program has meant that, apart from the revision of details, only its sense of foundational optimism is affected. The essentials of Kantian arguments supporting the central place of the Subject remain untouched. Ultimately, the price paid both by idealists and many anti-subjectivist philosophers (particularly in France) has been the loss of rationality: naturalism is the attempt to overcome this polarization and its convergent outcome. So it has to address the challenge raised by the usual objection: namely that any reduction of rational necessity, values and rules, to mere facts (the contingency of the actual) prevents not only science and ethics, but even judgment.

§ 1.2. Reason as not mathematizable

For classical rationalism, even if nature were chaotic, reason would remain the realm of order, which can be described mathematically. Kant denies the pretence of isolating such

a formal science from knowledge of the world. Yet a science of the Subject is impossible. There are many arguments to support the view that reason cannot be completely described in mathematical terms. Among them, particularly interesting are the five that follow:

i) such a formal science would have to make use of the same capabilities it sets out to describe — thus setting up a non-explanatory circle;

ii) even if this objectification of rational 'functions' were possible, it would not be faithful, for the thinking self would be lost as *thinking* ;

iii) to be adequate, such a theory should be complete, while such a completeness claim conflicts with the transcendent use of *ideas* (Kant), with the intrinsic openness of *Erlebnisse* (Husserl), and lastly with the incompleteness results stemming from Gödel's work in logic;

iv) no theory of mind could be geometrical, since space would require it to deal with an *extensive* manifold and the 'objects' to be *localized*, whereas both aspects are lacking, for mind and concepts respectively;

v) a mathematical theory of non-geometrical character would not capture the *continuous* inner flow of mental life, but at most its discrete aspects, which do not exhaust it at all.

Thus, even aside from v), it would be out of point to resort to the functionalistic program, in order to show that there is already a version of the Kantian way, that accepts a particular mathematization of reason, as modelled on recursion theory. Since I have presented in *Noema* a rejoinder to i)-v), here I shall just recall some ingredients of that rejoinder. First of all, geometry is no longer confined to the description of physical space, having become a more and more general theory of abstract spaces — able to be applied in modelling purely qualitative phenomena. Moreover, nowadays *logic is subject to geometrization, through categorical sheaf theory*. Yet, this symbiosis of geometry and logic could be considered at most the basis for an as yet inadequate, because static, representation of mind, since it would not privide room for the emergence of selected discontinuities in the physical world as much as in categorization; thus we can make use of additional resources from differential topology, as illustrated by Thom (1980), in dealing with morphogenesis: they allow the application of a mathematical theory to the dynamics of object-forms and concept-forms (*logoi*) involving epistemic and linguistic structures — this, however, does not commit me to an acceptance of Thom's overall picture. Apart from the unsatisfactory character of such a juxtaposition of theories covering different topics (in order to deal with other aspects of the mind design than the recursive organization of discrete symbolic units), these ingredients still leave the door open to an overall Kantian framework, as is exemplified in Jackendoff (1983).

Since the publication of *Noema,* however, I realised that the ingredients taken from differential topology in order to deal with the stability of *logoi* can be reformulated in categorial terms, by weakening the requirements on the category of smooth spaces, cf. McLarty (1992) ch.23, which is the proper universe of discourse for differential geometry.

Thus if.a unified conception of the architecture of mind is to embody the interplay of the continuous and the discrete already exhibited by algebraic topology, and to capture the functional relationships between generalized spaces and generalized algebraic structures, then it is category theory which becomes the natural organon of such a unification. But the objective dialectical character of category-theoretic concepts is at odds with any idealistic epistemology.

§ 1.3. Mind as a system of rules

Kant's transcendental method inherits the classical form of functionalism as it was introduced by Aristotle and reasserted by Descartes: mind is characterized by a system of formal rules. They run autonomously in deduction and through them any information provided by the senses is organized, processed and stocked in memory, affording the basis for induction. If this is a point on which rationalists and empiricists agree, the ontological peculiarity of such rules is clearly stated by Kant: they correspond neither to physiological structures nor to some platonic realm. Conformity-to-rules becomes a many-purpose tool: e.g., in distinguishing waking from dream states, and justifying the assignment of a key role to the category of 'cause' — otherwise perceptions would constitute a chaotic sequence. In order to accomplish these tasks, such rules must have a source independent from the contents of experience and even from their actual instantiation in the brain. A similar view in AI has been vividly stressed by saying that «Intelligence is mind implemented by any patternable kind of matter», in the words of H. Simon and H. Newell. This standpoint purports to identify thought with the manipulation of formal tokens, viz. 'symbols' in some programming language.

According to modern functionalism, mind is to body as the software of a computer is to its hardware; what matters are *programs*. In Kant the role of the computational paradigm is played mainly by geometry and arithmetic, syllogistics being just an all-purpose inferential motor, independent from intuition, whereas the building blocks of geometry and arithmetic are provided, respectively, by the «pure» forms of intuition, i.e., space and time. Significantly, only the latter enters into the central notion of scheme, which orders the application of forms and categories to data. The necessary truths found within formal sciences such as geometry and arithmetic tell us that there exist properties of the mind that are definable only in non-physical (non-chemical, non-biological, ...) terms. Hence, physics (chemistry, etc.) does not exhaust reality. This thesis is not to be confused with the weaker admission that nature is *stratified* : there are properties of objects of level $j > i$ that are not definable in terms of properties of objects of level i. The latter amounts to a form of non-

extensionality (some entities are not determined by their elements), that in itself does not imply the dualism of mind and matter. The stratification is also weaker than the form/content dualism, because emergent properties at *j* share a common background with *i*-properties, whereas this is not the case with functionalism — if not trivially: functionalists can conceive anything as if it were a suitable machine, thus the same form can cover every content (think of a universal Turing machine).

In contrast to the position outlined above, any real understanding of mind requires us to isolate a set of (natural) constraints on possible software/hardware pairings, in such a way as to exclude the notion that at least certain programs are operative on any support and in any environment — e.g. one endowed with physical features radically different from those of the brain.

As for the regulative side of the mind, Wittgenstein's and Kripke's analysis of *following a rule* can be regarded as a tool to deal with the generalizion of skepticism from rules supposed to guide linguistic behavior to rules that would guide human thinking in general, in as much as it is supposed to be independent from social use (and identification) of rules: which conception should apply in Kant's case too. Yet, if we cannot identify a rule, then any game that contains it is even less identifiable, be it solitary or collective, linguistic or olympic, unless one invokes the magic of structuralism. But if it is only our phylogenetically inherited common hardware that blocks skeptical regress, isn't this a brute fact? Then, is it for no *reason* that we follow, or believe that we follow, the rules we happen to follow in reasoning etc.? Or could the notion of 'brute fact' be just a result of ignorance?

Since the birth of evolutionary biology many facts concerning life on Earth are not regarded as brute. Yet one can trace such a quality back to cosmology — according to the idea that physics concerns the *laws* of nature, while cosmology simply describes how these laws are contingently implemented. One can see here a parallel with the dualism of logic and psychology in relation to mind. Still, wouldn't the very existence of the Universe remain a brute fact? Actually, investigations on the initial conditions of the Universe have provided new understanding, and are now tending to alter the status of the dichotomy between (contingent) initial conditions and (necessary) laws of nature. Therefore the onus that lies on the functionalist is far more arduous than was formerly supposed, once mind is admitted as a part of nature.

Moreover, as Wittgenstein realized, skepticism is only meaningful locally, *i.e.,* by contrast: there are to be *specific* reasons in order to be skeptical about anything, whereas if everything is quaking under the skeptic's feet, there is no piece of ground *in particular* that shakes, thus it makes no sense to question the grounds of belief, or to ask why something is quaking in human knowledge. The claim that socially shared evidence is necessary and sufficient for ordinary purposes in identifying rules is hardly consistent with such localization, once it is applied to meaning itself. The anxiety of an absolute (un-)certainty

that pinches skeptics is the same that befalls metaphysicists: a side-effect of arm-chair fantasies *as much as* of the pretenses of «pure» reason. So, why should skepticism be fought by strategies of total war (*i.e.* transcendental arguments) instead of local tactical expedients? The problem in invoking transparency of use as an anti-skeptical argument lies in that it doesn't provide any theory of linguistic competence.

Children do *not* need any instruction to learn a language, while computers have to be provided with a whole battery of explicit rules. Is a network of computers in better position as for the rationality of rules? It is not simply a matter of built-in rules, nor does the appeal to plasticity solve the problem, since plasticity may destroy a system, while, were it perfect, we would be gods. Thus we have to take account of various and variously linked architectural levels at which the rules are given. For example, as emphasized by Chomsky, the universal grammar UG, biologically inscribed into the mind, might consist of meta-principles for constraining the range of possible syntactic rules. These principles, however, do not suffice for the mastery of language: the semantic component involves the contribution of other mental abilities. The trouble is not simply that these do not form a cumulative hierarchy (in fact our *ability* to escape rules rests on such pre-established disharmony); the trouble is also that it is not at all clear how something physical can be (or provide for the emergence of) a rule, and moreover, a flexible rule. His biological stance notwithstanding, Chomsky's view, if taken at face value, remains dualistic. Thus, meaning and the nature of rules are still to be explained.

§ 1.4. One and many I's

Kant argues that the category of substance cannot be applied to the Cogito, as it can to empirical entities. The resulting Subject turns out to be a pure system of representational and referential functions. In the *Critique of Pure Reason* (A 123-124 and B 17) this Subject is respectively described as a «stable and permanent I» (having psychological reality) and as a center of synthetic activities (having formal reality): so it becomes the «Choreographer» of ideas. This double character of the I persists in Husserl (cf. his *Ideen* II §§ 24 and 27, where the I is a center of functions and the concrete I is part of the pure I's environment). By constrast, from orthodox Neo-Kantians like Bruno Bauch to liberal ones as Giulio Preti, the structural aspects of mind have been desubjectivized.[3]

Here, the Humean criticism of the notion of self, whose unity is simply postulated by Kant, becomes particularly in point. This criticism remains pertinent even if the self is

[3] In this regard, Preti anticipated, in a conversation with Michel Foucault, that *phenomenography* advocated by Hao Wang to overcome the difficulties of analytic philosophy, in dealing with mathematical language, cf. Preti (1973).

desubstantialized. For Hume attacks the view of the self not only as a substance, but also as a necessarily coherent unity. What is thus called in question is both Descartes' *and* Kant's Ego: notwithstanding that Kant does not apply noumenal properties to the I (*i.e.*, it is no longer a *res cogitans* but the precondition through which phenomena acquire unity), the transcendental subject is and has to be the unitary source, the cristalline warrant, superintendent on all cognitive activities, and detached from the passions (hence, lit. impassible).

Yet, the proposal of a new *zurück zu Hume* would mean forgetting the aporias of empiricism diagnosed by Kant and, moreover, missing the idealistic failures common to Hume and Kant. The point is that, whereas for Hume the I is a bundle of different perceptions, a motley of contextual responses to an inscrutable environment, I take the term 'bundle' with its present-day mathematical meaning derived from the theory of fiber spaces. In the same vein a bundle can be generalized into a *sheaf* — for the definition of this crucial notion, see Peruzzi (1991a). Still, there is no *a priori* necessity in this passage from bundle to sheaf. As opposed to Kant (and Husserl), the unity of noetic functions is a *problem*, as is the unity of phenomenal apparencies, and not a prerequisite for organizing the latter into regularities. So, my answer to Pascal's question «Where is this I, if it is neither in the body nor in the Soul?» (*Pensées,* # 323) consists in giving up the I as a *separated locus*, controlling every specific subsystem of cognition. The I is rather (the result of) a certain coherence of the totality of these subsystems. Then, it becomes in its turn a possible object of investigation through the tools provided by the subsystems. In this sense I adopt Terry Winograd's term of 'eterarchy', exporting it from the domain of language understanding, and giving it specific mathematical structure through the sheaf concept.

The shift from hierarchy to eterarchy is already demanded by the analysis of very simple cases in which two mental agents need each other's outputs. A nice example is given by Minsky (1985) § 3.4. Suppose we want to decide whether the scene on the left represents two or three objects:

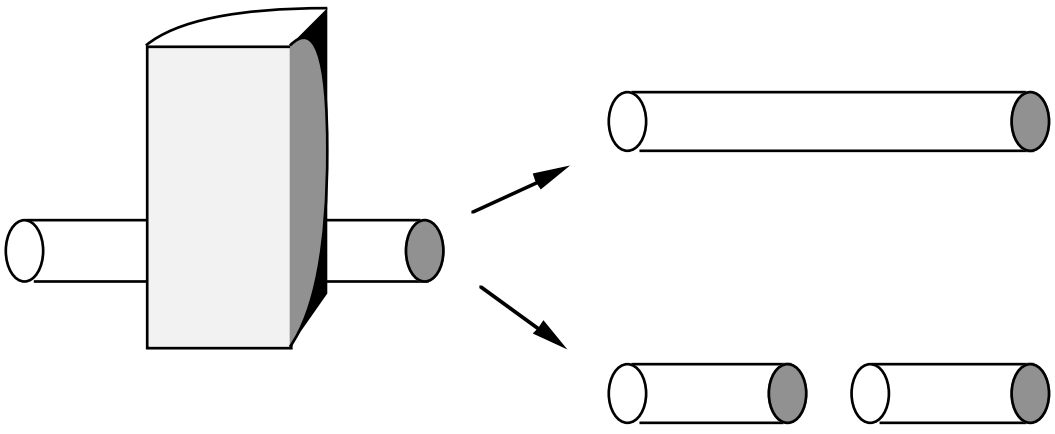


Figure 1

The agent `Seeing` could answer by subordinating the agent `Moving`, but in order to be applied to the front block, `Moving` needs `Seeing` to establish whether there is any obstacle in between; hence, instead of a rigid unidirectional hierarchy, we have a sequence of loops that give rise not to a vicious circle but rather to a feedback coordinated action. Still, doesn't this coordination impose a central eye? In effect, especially if the consideration of `Language Understanding` is added, one is led to admit the existence of an inner structure connecting the different rules of different subsystems of the mind. So, we have to specify how §§ 1.3 and 1.4 are related, in order to avoid inconsistency.

According to the computational model of mind, to possess mental states, as expressed by propositional attitudes (*viz*. to believe, to see, to expect that, etc.), means to have internal representations and certain computational relations to them, governed by recursive principles. Now, if the mechanisms underlying cognitive processes are computational, then the access to knowledge is obtained by the *form* of representations, since computational systems are *syntactic* . Isn't this at least a partial vindication of Kant?

Fodor (1983) grafts on this functionalistic background the thesis of the modularity of mind: mind consists of a network of modules, each one having specific features, independent from the others, still there is place for a central language of thought, formatting the cognitive resources traditionally covered by the term 'reason'. So, there are two kinds of systems: i) central systems, inhabited by consciousness, belief, logic and induction, and ii) input systems, that are «informationally encapsulated», peripheral, mutually autonomous and independent from central systems. Therefore, the analysis of any given output, as provided by systems of type ii) *is not concept-driven*..[4] For this reason the structure of small input subsystems and large modules is impermeable to belief, and thus co-reference is possible: the forms of intuition are pre-categorial.

Hence emerges the metaphor of the mind as an orange with modules as its slices.Yet what a strange orange, since it has a kernel, and this kernel is the softest part!

Modular input systems have specific domains, innately determined, hardwired, and not assembled by experience; their mechanisms are computational and possess *mandatory* character: one cannot see the source of a visual stimulus except as what one sees, one cannot hear a phrase but as a phrase, instead of as a simple noise, and the passage of fingers on a sheet automatically causes a tactile sensation. In normal situations, these systems are

---

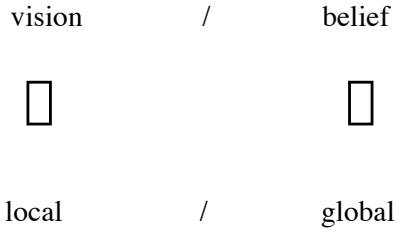
[4] This is not to say that it is bottom-up. A gestalt-psychologist would accept that perception is not concept-driven, but still claim that it is top-down — for the 'top' here is always local.

activated whenever it is (possible — apart from physiological damage, inhibition caused by drugs, or some sort of control acquired through forms of meditation.

Central systems have access to stimulus generated information only after this has passed through input analyzers; intermediate levels in perceptual processing remain impermeable to central systems. In contrast to the slowness of these (for example in problem solving), input analyzers are fast, and this advantage is obtained through the very shallowness of mandatory processes. The informational encapsulation of input systems implies that the so called «New Look», according to which beliefs and expectations largely affect perception (by general top-down, concept-driven processes) has to be revised: we discover the way the world is, even when it is in a way we do *not expect*. In particular, vision cannot be globally holistic. And this extends from vision to any other input system — for instance, to syntactic analysis.

Usually, this sort of argument is taken as a support of the view that syntax is formal, whereas a generalization of case grammar such as that proposed in Peruzzi (1992) shows that the involved ‘forms' depend on schemes of semantic interactions, although they yet remain modular. A similar point has been emphasized by Petitot (1983) in relation to the import of topological singularities in linguistics.

Now, since the property of being encapsulated means for a subsystem S that information external to S is excluded from S-performances, then S is encapsulated not only relatively to the ‘center' but also to any other S': a hypothesis on which evidence of intermodular coordination in the newborn sheds some doubts — as we shall see in the next section, these doubts are symmetrically inverse to those on holism with respect to central processes. Moreover, while Fodor is inclined to deny the existence of modular systems that are not input systems, I suggest *local* modularity also for central systems: so, even if there were a central processing unit (CPU), it would have a more reduced role than is usually supposed. The difficulty lies in governing the interference among modules but the key to the problem lies in the existence of *patterns* of such interference. One might reply, if there were just modular input systems, how could they interface? Interface patterns would have to access to outputs of all modular systems. Thus, it seems that the representational ingredients of each module have to be integrated, and corrected, in the light of general knowledge. Yet, there is no general knowledge which is not the memorized result of interface procedures among modules; that is, there is no strict correspondence such as



for there are local/global aspects at the various layers of the system's architecture, and they pertain to both vision and belief. Even for central systems, one could say that an intelligent system S results from the *connections* of less intelligent units composing S, each of them being in its own turn a structured net of agents, still less intelligent. As before, this decrease in intelligence is necessary to avoid circularity or regression in explanation — but it is not reductionist in spirit, as it seems to be for Minsky. Furthermore, these 'agencies' are relatively autonomous: each one does not know what another is doing at the same time, for such communication requires messages, which in most cases are lacking, and when they exist, the problem emerges of how they are mutually intelligible; anyway, if pre-established windows are built-in to the patterns of intermodular connections, then, in view of their specificity, they are not holistic. Thus, mind does not have a unique inner time either. Here the twofold internal/external perspective in conceiving sufficiently autonomous units assumes an essential role. E.g., describing the child's behavior in constructing a tower of blocks, we can say that from outside (looking at it as an agency) `Constructor-of-towers` knows how to act, whereas from within (looking at it as an agent) it does not. On these grounds, meaning is a question of interfacing relatively autonomous devices of reference and action, but since the external could always be internalized, the dichotomy between internal and external is insufficient for explanation as long as it is not correlated to local/global factors. This leads straight to questions concerning the foundations of semantics: in fact, such correlation is required in order to give account of *basic facts* concerning propositional attitudes, as shown in Peruzzi (1987).

On the other hand, the interface mechanisms among modules, agencies, and similar subsystems cannot be *completely* domain-specific; nonetheless, we haven't to represent any belief as if it involved the activation of every component, e.g., the belief that tomorrow it will rain cannot mean that the claim «Tomorrow it will rain» takes account of *all* information of any sort; it is simply sufficient that one source of information, competing with others to drive the responses of the system (each source being taken with its associated sort of *ceteris paribus* clauses), succeeds in overwhelming others: there is no need of a centrally guided weighting. *Tertium datur* between shallowness and pervasiveness, since matching patterns are an example of morphisms (in the language of category theory) respecting conditions of structural stability. These morphisms embody constraints on the objects they link and contribute to determining a functional environment, susceptible of precise mathematical description.

Thus, instead of Fodor's sharp separation of input systems from central ones, I argue for both i) the domain-specificity of features of logic itself, and ii) the integration of modules as not less direct than perceptive output, instead of being meta-theoretical, see Peruzzi (1989), since it is generated by intrinsic structures of parallel processes. The

conjunction of i) and ii) conflicts with the existence of a fixed central system, the source of order over any component, and supports rather a *sheaf* model of mind, of the kind already indicated. Of course, this does not obliterate relevant differences between lower and higher levels in modules' architecture. That systems corresponding to logic, problem solving, belief, etc., are not encapsulated *does not imply* that they have to be viewed as absolutely non-modular — contrary to the conjecture that domain-specific, but not encapsulated, systems do not exist. At the same time, by talking of 'sheaf', I intend to avoid from the start locating thought as a particular module among others: thought emerges from the harmonious running of modules, as well as from the varying suspension of *each* of them (cf. below, the notion of local *epoché*), but once emerged it acts as another module-like system, with its specific structures. It inherits meaning from the underlying modular activity and gains formal character through the lifting from basic perceptual features to their transposition in 'abstract' categories.

Of course this view of the mind is not immune to difficulties, but its advantage will prove as particularly relevant to various topics, e.g. to the significance of Gödel's Theorem. Moreover, it is consistent with the opposition to locating philosophy as a meta-T, for every theory T, and it constrasts with two sources of such a voyeristic attitude: semantic ascent and the dichotomy between a foundational activity towards the products of thought and descriptive activity towards thinking itself. Such ascent and such a dichotomy turn out to be neither necessary nor sufficient. Philosophical activity lives in the most specialized not less than in the most general regions of knowledge, even though its usual manifestation, in *present* Western culture, assigns a privileged place to the latter. It inhabits our sensomotorial regime no less than our theory of language. Furthermore, the very flow of meaning that manifests in language is first interrupted and then distorted by its persistent objectification *in vacuo*. Often the attempt to reflect on what you are thinking (or on your thinking) modifies your thought. This interference is sometimes pertinent but not always. If you are playing a tennis match, reflection on whether your way of hitting the ball is technically correct is a prelude to losing; likewise, if you are exercising geometric intuition in order to solve a problem, reflection on the logical principles you are using may be a prelude to losing sight of the solution to the problem.

Finally, there is another, not less intrinsic, manifestation of intersubjectivity, which undercuts the notion of an isolated subject: the self is structured *through* and *for* relations with other selves. Indeed, in order to establish a common reference with adults for his/her words, the child learns to orient (or rather matures in orienting) his/her attention to the target of others' attention, by following the glance of the others — or through different perceptual abilities in the case of the blind. As Bruner (1986) observes, the child has no difficulty in commuting deictics, whereas he/she *should* if the starting point of linguistic-cognitive development were egocentric; by the same token, we should expect to meet with

similar difficulties in explaining common reference if our methodology were as solipsistic as Fodor's. The ability to switch, in any given situation, from our to someone else's perspective is an intrinsic potentiality which is confirmed already in two year-old children by their capacity to *adapt* their own sentences to be more and more adherent to the shared situation and intelligible to others.

§ 1.5. Logical Opacity

Hobbes already made explicit what we now consider one of the basic tenets of functionalists, *viz.* that thought equals manipulation of inner symbols: ideas and propositions are *words* of some mental language. Thought has nothing to do with things, but rather with representations, and these are of a formal character. According to this view, reasoning is a matter of the rule-governed combinatory manipulation of symbols following computational rules. There are two cases: if mental symbols are manipulated as already meaningful entities, the processing of thought cannot be entirely mechanical, for meanings are not subject to physical forces, neither are they mastered according to a pre-established set of recursive rules; otherwise, manipulation is not, by itself, reasonable, for reason depends on what symbols mean, while manipulation would occur as a brute fact. This is a reformulation of the «paradox of mechanical reason», as Haugeland (1985) terms it.

The argument can be linked to what has been said in the previous section: if symbols are intrinsically meaningless, understanding seems to be miracolously produced by something shallow, while if they are meaningful, another difficulties arises: what is the status of the intelligent manipulator, the so-called *homunculus* ? The problem, with such a *homunculus*, is threefold. First, it must not only control the manipulation of symbols, but also inject meaning into symbols. Second, in order to perform such manipulation, it) has to have access to and be conscious of them. Such activity is in its turn of a representational nature, hence making it necessary to appeal to another, still more internal, source of such representations, and so on. Thus we risk an infinite regress of manipulators. Third, suppose there is a finite team of *homunculi* whose joint operation gives us understanding and meaning; going 'downward' to the simplest units, they are progressively more shallow, and finally a maximum of shallowness is reached: this could be identified outright with the physico-chemical level, the firing of neurons (or, possibly, other units). But how is it possible to pass, going 'upward', from shallowness at degree d to shallowness at degree $d' < d$? Where does meaning enter? For intermediate *homunculi* everything is simple: semantics is structuralistic and internal, so it reduces to syntax, while at the top level (of the

Self) it is no longer syntax, and at the bottom level (of 'brute' hardware) it is biochemistry, ultimately nothing else than physics. Thus, something in this picture is wrong.

Actually, one can be prompted to say that the human mind is endowed with a 'perspective' device for imprinting meaning, computers do not. Is this a gap 'in principle' or just a contingent matter of fact? We shall consider in the subsequent sections some anwers. Previous remarks related to intermodular constraints and the sheaf model suggest that meaning emerges out of multi-module amalgamated activity, exercised within the natural *ecos*. Still, the top level of this amalgamation does not support a globally separated entity. Nor does the bottom level, whatever it turns out to be, correspond to absolutely amorphous data, alien to law and structure.

There is no explanatory power in the above hierarchy of *homunculi* within the mind, each of them calling for a new screen on which a new smaller spectator can see (observe, represent, grasp, check...) how the preceding *homunculi* are acting, unless we take into account the genesis of their mutual interaction through their interaction with an environment extremely rich in structure.

So we should realize that such double interaction does not need any first *homunculus*, and this affects the Kantian view too. We have already hinted at some aspects of such a view that can no longer be maintained. Now we are facing what is one of the most influential steps in the direction of idealism. In the *assertion* «I know myself», it seems that the I-subject cannot coincide with the me-subject. The different levels of the Self appear in very concrete experiences: in front of a Frazer's spiral, I see it as a spiral, but I know it is a set of concentric circles. Though this same consciousness reveals some inner opacity, the basic tenet is confirmed. The knower, as such, transcends the known. The simplest way out of the resulting infinite regress consisted in positing a unique, central, indivisible, self.

This way of reasoning is not specifically Kantian. Kant merely provided a systematic framework for its use, transforming it into a basic principle of the mind's design. If philosophy has sometimes taken routes in sharp contrast to common sense, this is certainly not the case here. The idea of a unifying self is in fact the usual, comfortable view of common sense, which was converted into one of the bulwarks of Idealism. From Kant's synthetic unity of apperception to Gentile's «pure act», this core of selfhood has been conceived as an intrinsically active *source* of thought; and as such it cannot be completely represented. Thus Gödel's Theorem can be seen as a precise statement of such as Gentile's. Given the essential gap between theoremhood and truth which Gödel established, the former can at best represent an approximation to the latter; an approximation that cannot be improved on by rule-governed procedures, since rules are supposed to be recursive, whilst truth transcends the range of recursion. This suggests that truth is linked to meanings that are diffused and not localizable within the bounds of any isolated formal system. Hence the

attempt to fix meanings for separating them from the external, leads to non-terminating programs or to paradoxical loops.[5]

However, once we try to reduce incompleteness by appeal to the external (formal or informal) environment, we face a final difficulty: does not cosmology itself turn out to be incomplete as a formalized theory? Or, are there unforeseen constraints that intrinsically confine incompleteness to the very separation of the purely formal from content and existence? One could suggest that completeness is only a relative phenomenon, not merely in the sense that the classical completeness of $Th(\boldsymbol{M})$, for any structure $\boldsymbol{M}$, is just an instrumental hypothesis for discussing axiomatic theories with respect to features of their models, usually taken as richer than the given expressive devices, but also, and primarily, in the sense that categorial semantics has made clear: namely that completeness of any theory is parametric on the categories and the functors considered. The situation is critical in the case of a cosmological theory T, if the completeness considered is intrinsically related to the meaning of T-sentences: here we have a sort of transcendental loop of form and content. It is revealing that exactly this loop is active in the emergence of the self.

Now, the formation of the self-symbol is grounded on synergetical interactions of parallel subsystems, producing as a result what is *interpreted* as a coherent behavior of neuronal networks. Although each of them is of a mechanical type, their collective outputs are sensitive to threshold effects, whose number is such to make the exact global result unpredictable and yet such that we can know that if the complex system is of the right kind, phenomena of coherence on a global scale will be manifested: such phenomena affect and then dominate the contribution of single elements composing the system — as is already shown by natural systems of non-cognitive sort, such as laser beams, Benard's instability in fluids, etc.

In a similar sense, when each of us thinks he/she is an I, this 'I' is simply a comfortable simplification of a multiplicity of structures emerging from the cohesive activity of stratified layers of units. The hypothesis that this multiplicity can be enclosed in a system of recursive instructions is, at best, lacking evidence, even if it were possible to describe well-defined classical programs for each lower 'slice' of the mind, and we can consider incompleteness results as a sign that such confinment is not exhaustive of mind's architecture.

---

[5] In this connection, it is useful to recall Hao Wang's proposal of conceiving of truth as provability in an informal collection of formal systems. One would concede that such a collection can be specified for *each* given sentence of any theory under consideration, but cannot be so specified for all sentences and for the *whole* of our theories, once this 'whole' has achieved self-reference. Here the above appeal to the notion of sheaf can be made more precise: the 'sheaf' pasting of mental modules cannot, by symmetry with respect to local *epoché* (see § 3.1), but be local. For example, at any given time, attention can be directed on only some features of experience, just as imagination can de-activate only some, not every, cognitive component.

Although it differs from a methodological point of view, this synergetic approach shares some aspects with recent tendencies in AI and cognitive science, under the heading of 'connectionism'. Attention to cognitive processes that *look* rule-governed, while they emerge out of a system lacking codified rules, has been given by connectionists such as J.L. McClelland, D.E. Rumelhart, *et al*. However, even though mental representations are the effect of self-stabilizing processes (instead of the output of *following rules*), what we have before us is not necessarily a new sort of associationism, since, through the understanding of links relating each local structure to its environment, we are led to recognize that the associations are not freely varying, being *topogenic* (as in Waddington's epigenetic landscapes), and they are consistent with modular organization of higher cognitive structures.

The threefold renunciation of a meta-empirical Subject as an inner source of meaning, of a *homuncularis* hierarchy, and of a globally distributed (non-modular) architecture, still leave us the task of answering the question, in what then does our capacity for self-reference consist?. Hofstadter (1985) supplies an apparently circular answer: mind is a configuration perceived by a mind. That is, once the self-symbol has been codified by self-reference, a resonance occurs (among different levels of the system) that becomes progressively stronger, leading to the internalization of the auto-intentional stance. So the Tortoise concludes (ib., p.626) the dialogue with Achilles:

> The more that illusion of unity is cycled through the system, the more established and hardened and locked-in the whole illusion becomes. It's like a crystal whose crystallization, once started, somehow has a catalyzing effect on its further crystallization. Some sort of vicious closed loop that self-reinforces, so that even if it starts out as a delusion, by the time it has locked in, it has so deeply permeated the system's structure that no one could possibly explain how or why the system works as it does without referring to its 'silly, self-deluding' belief in itself *as a self* .

Apart from my disagreement as to the illusive character of such an attractor-oriented process, the presence of this «level crossing feedback loop», responsible for the reification of the self, requires that the sheaf model of the mind is suitably enriched, as seen as resulting not only from locally coherent integration of cognitive modules, but also from a fixpoint quality of the intermodular connections. On the mathematical side, this suggests the investigation of sheaves of theories and their meta-theories (simultaneously considered) — so we find in this problem a new impulse to deeper theoretical work. On the empirical side, we are led to clarify the relations among modules as components of a sheaf-type architecture and as components of a hologram-type architecture. Indeed, if modules are not completely encapsulated, a holographic model might appear more adequate, and in fact K. Pribram has proposed a holographic model of memory (though memory is not a module

in the sense described above). Albeit controversial, this model throws into sharp relief the empirical import of the choice we are facing. It highlights the problem of identifying the precise details of the relation between and the holographic trace of built-in local and global linkages and the sheaf model. This too requires both theoretical deepenings and experimental tests. In the meantime it is possible to recognize some important advantages of the hypothesis of I-as-a-sheaf. The leading idea is that the sheaf induced by the stabilization of the symbol for the self may undergo transformations allowing the plasticity of self-modification. The following picture can illustrate the situation:

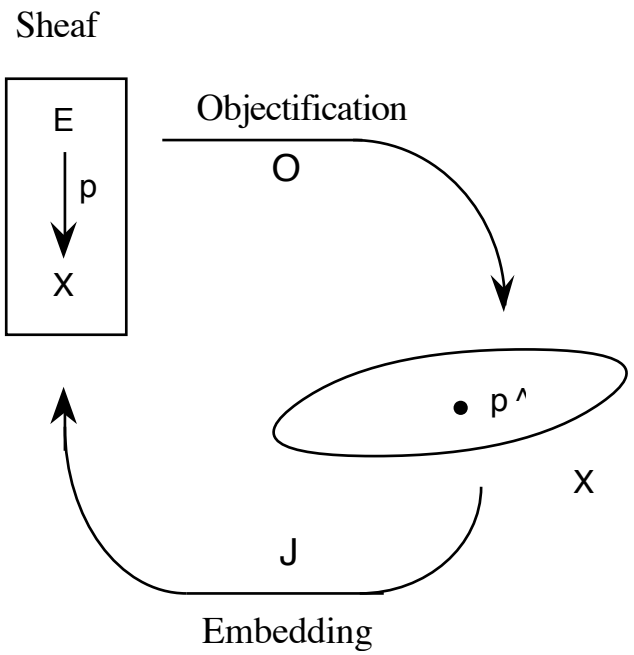


Figure 2

where, for simplicity's sake, I appeal to a spatial sheaf, as given by a local homeomorphism p on a space E to the base space X. The objectifying map corresponds to self-access and the loop $J(O(p))$ corresponds to the rise of the notion of 'I'. Since p can vary, keeping X fixed, possibly $E = X$. The structure of this model is such that it rules out the possibility that the classical universe of sets can provide room for such a dialectical loop: the only set X, such that $X^X \cong X$, is a singleton. Whereas we already have categorial models for particular cases of similar 'closure', related to (un-typed) lambda-calculus, X being an extremely rich space.

Moreover, once embodied into the system, the notion of the self as a self-stabilizing fixpoint of eterarchical processes may be conceived as an *attractor* in the dynamical

evolution of any sufficiently complex system. This gives rise to new phenomenological perspectives as well as new logical investigations relating sheaves and attractors.

Thus the symbol for the self is no longer confused with the inner eye of some *homunculus*. There is no *someone* who has perception of the functioning of the system: there is only the (locally determined) cohesive activity of the system. Coming back to the analogy with the foundations of mathematics, it is clear that this conclusion may be applied also to the case of sufficiently powerful mathematical theories, making us drop the idea of a *homunculare fundamentum* : theory, meta-theory, meta-meta-theory and so on. Consequently, if foundations of mathematics have to be tied in some way to mind's architecture (and this does not imply any sort of subjectivism!), it seems more feasible to develop a network model of mathematics as a self-stabilizing system of concepts and theories, reinforcing the phenomenological position from which MacLane (1986) deals with the evolving architecture of mathematics. Phenomenological considerations are essential to mathematical thought, but they are not, in most cases, meta-mathematical. It is rather the intertwining of cognitive activities that provides the proper material whereby mathematics can develop and be nourished by new insights into problems raised by interactions with nature. When the nature of mind is in question, the above loop becomes a crucial topic for mathematical investigation.

Chapter 2. Neither Functionalism nor Holism

Once thought processes are conceived as formal, according to the information processing model (IPM) of the mind, mainstream epistemology and psychology both support the idea that thought has access only to formal properties of what is provided as the output of the senses. (This idea lies the main *non sequitur* of which I ask the reader to keep track, but we shall come back to it only after a long detour.) Now, if understanding of meaning is located at the top level of mind architecture, a dilemma enters the discussion: either meaning is purely formal (be it recursively determined or not), or reference too is form-laden.

In the former case, thought processes do not have access to genuinely semantic properties, so, since mental activity can only confront representations with other representations, and not with what they represent, questions of truth-value or denotation have no place in the study of mind. (This excludes a priori the very project of phenomenology, as aimed at explaining 'aboutness', the basic notion of semantics, in terms of intentional acts *and* their formal structure.)

In the latter case, we are left with just another reformulation of the Kantian stance, that is ready to be either to be updated in the way suggested by Ray Jackendoff, or diluted in the variety of ways that semioticists, cultural relativists and post-logical-empiricist philosophers of science have propagated.[6]

The trouble of the former case is that the resulting exclusion of semantics is already called in question in the case of artificial systems operating on 'microworlds' — as specified by computer programs. These systems cannot even be said to have either true or false beliefs about the microworld, as far as truth and falsity differ from purely formal properties. Then, how are semantic properties of mental representations established? The second alternative answers by enlarging the range of syntax and making semantics a sub-region of the representational/computational world. The reasons why both solutions are

---

[6] These reformulations might appear to be in agreement with what has been said in § 1.1, pointing to a non-standard way to trace the demarcation line between syntax and semantics. Such appearence turns out to be misleading, as ch.3 will show.

inadequate will emerge through the following discussion of some aspects of functionalism and holism, focusing on opposite theses advanced by Jerry Fodor and Hilary Putnam.

§ 2.1. Psychosemantics

Recent studies in semantics have made it clear that interpretation depends on a structured *net* of facts, not simply on a causal chain or a tree of monotonic epistemic stages; such a net concerns the transactions of both the interpreting and the interpreted system with the physical environment. Taking this insight as firmly established, we shall look at whether and how it is to be explained, making a sort of theoretical 'sandwich' of opposite viewpoints.

Now, Fodor admits that there is no incompatibility between the 'rational psychology' he proposes and naturalistic psychology, but holds that, from the methodological point of view, the former is the only practically feasible option, the latter leading to meaning holism, and thus to skeptical consequences concerning the very possibility of semantic theory. He takes from Putnam the idea of such «methodological solipsism» (a term that, in fact, can be traced back to Carnap) according to which the only legitimate constructs in psychological theory are *states* in a strict sense, meaning 'non-transparent' states, *i.e.*, not involving any differentiation of the effects that truth or falsity could have on mental structures and operations. Thus psychology needs only 'opaque' attribution of mental states, and this is also sufficient to take account of what the subject 'has in mind' and the motivations causing his/her behavior. Notice that this methodological option favours the condition of formality above, by preventing the attribution of (transparent) semantic properties to representations. In this sense, were semantic properties necessary to describe mental states, functionalism should be rejected. (Still, the option does not exclude non-computational formal properties of mind. Vice versa, many semantic properties may well be recursive.)

Unlike Putnam, whose theory of reference admits causal elements external to mental contents as well as social distribution of semantic competence (while finally reinterpreting all talk on reference and truth internally), Fodor simply identifies mentalist semantics with functionalist syntax: holism would mean renouncing semantics, for the same reason that Leonard Bloomfield pointed out: if to grasp the meaning of 'salt' it is necessary to know what 'salt' refers to, then we must appeal to chemistry, and since semantics would involve chemical and physical and ... (cosmological?) terminology, it should consider any object and event in the Universe, requiring us to know what salt and anything else really is. Thus, semantics is impossible. The moral of this argument is assigned a wider significance, as indicating the sterility of any naturalistic psychology. Hence the necessity of the restriction to an 'opaque' description of mental states. The remaining question would be whether it is

also sufficient. However, I consider this entire argument as inconclusive, because it presupposes that holism is the correct general view about knowledge and science, while any sort of solipsism ends in blocking any explanation of how our understanding of everyday use of language in concrete contexts is possible.[7] And since functionalism is attacked by Putnam on the same holistic grounds, I will propose a view of meaning, different from both Fodor and Putnam. In order to show that the contraposition of holism and functionalism is not exhaustive, some aspects of these views have to be examined.

Fodor (1983) claims that, besides confirmation, (non-deductive) belief fixation also is «isotropic» and «Quinean»: this claim implies the sharpest negation of the thesis of cognitive encapsulation (typical of input systems), for isotropy means that any information can be relevant to accepting and checking a given hypothesis *h* . Unfortunately, the resulting view is far from adequate to reasoning in either science or common sense.

If our entire knowledge were involved in the matter of confirming *h*, then the attempt to separate the kind of phenomena to which *h* is supposed to refer from other kinds of phenomena, would be a priori doomed to failure. (So, at least another a priori truth would exist apart from that proposed by Putnam — *i.e.,* that not every sentence is inconsistent.) But if this holds for all phenomena, it holds also for the supposed laws describing them: hence isotropy (and anisotropy, for that matter), cannot be regarded as either true or false, since it is not even intelligible. Yet, suppose it is. Isotropy would reveal itself in the conviction that the world is a maximally connected causal system, whose connections we ignore forever, and thus we must be ready to modify our relevance estimations again and again. Were this true, how to explain that scientific practice does not start each day from scratch? One might retort in the spirit of Hume: because of the viscosity of our beliefs and the practical need for them. Once more we would live in a seemingly extraneous world. But how did we come to be there? Kant tried to offer an alternative to extraneousness by his transcendental method. We have already seen some of the flows in that response, and in chapter 3 still others will be met, dealing with the heritage of Kantian anti-naturalism. If, on the contrary, we stick only with beliefs and needs, from an empiricist standpoint, isotropy is even more untenable.

In order to say that we know something, we have to know that certain ‘connections' are not relevant, where this knowledge can be suitably tested by experimental tools. The application of isotropy to any case of problem solving involves confronting this point. To mention the irrelevance of telescopic observation claimed by some scholastics against Galileo is simply misleading: it does injustice to Galileo himself, who rightly would not admit Aristotelian psychology as pertinent to astronomy!

---

[7] In fact, Fodor seems to have changed his mind on the suitability of such option, even when conceived as purely methodological. But I will avoid here the discussion of his revised view, which presents other difficulties.

In discovery, a central role is played by analogical procedures (at higher levels through simile and metaphor, at lower ones via geometric intuition as realized in vision and sensorimotor ability), thus allowing the transfer of information across distinct domains. This also tells against isotropy. For tracing analogies is possible only thanks to the capacity of omitting as not pertinent several properties of the domains in question.

Fodor draws two different conclusions: i) the more intelligent a system is, the more isotropy it manifests; ii) the more global (isotropic) a cognitive process is, the less comprehensible it is. *Ad* i), I object that the valid side of holism is not isotropy, but its local relational dimension, anchored in properties of «regional ontologies» (in Husserlian terms). An intelligent system is capable of passing from isolating local qualities to a weighted connection of these qualities with other (non-arbitrary) qualities. *Ad* ii), I object that what matters are the *particular* patterns of notions of global character involved — first of all the notion of object and the relation figure/background that are involved in different situations, even if in a disguised form with respect to the prime spatial-visual context. Analysis and algebraic topology have investigated successfully for more than a century the links between local and global properties of spatial structures; the results are perfectly comprehensible and their range (cf. § 1.2) covers both physical and computational 'spaces'. Much less comprehensible is complete cognitive isotropy, and vast parts of contemporary mathematics have something to say in dissolving its appeal.

Now, besides isotropy of belief fixation, confirmation would be Quinean, in the sense of being non-monotonic: if evidence *e* confirms hypothesis *h*, it does not follow that the conjunction of *e* with *e'* still confirms *h* ; hence the degree of confirmation of *h* is sensitive to the whole mobile system of beliefs. Holism of belief fixation (isotropy) and such non-monotonicity are easily related: metatheoretical estimations, that assign different degrees of truth-likeness to *h* and *h*', are holistically indistinguishable from theoretical considerations.

Nelson Goodman's projectability could be cited as a property sensitive to the whole system of beliefs; yet the well known green/grue puzzle could also be interpreted as the end result of an unsatisfactory analysis of induction, rather than as a positive argument for admitting all-pervading 'Quinean' properties. I am not advocating backing up to an extreme compartmentalization, as if confirming or falsifying *h* were activities with uniformly fixed boundaries, predetermined independently from any other belief. Of course, central processes are not encapsulated. The point is simply that there are *covariance* cognitive relations, which are neither completely free nor of unlimited range.

For example, the relation existing between notions, such as 'dog' and 'bark', is stable (sufficiently insensitive to global properties of the system of belief), even if it does not necessarily hold. There is room for a *locally constrained* isotropy — if barking dogs turned out to be alien automata, barking would remain a phenomenologically prominent property in the formation of the 'dog' notion and its first lexicalization as 'bowwow'.

At this point it should be clear that holism is to semantic cognition as the inverted image of what Humean empiricism is to causal knowledge: the latter implicitly assumed complete separability of one phenomenon's properties from those of another (in conformity with the Aristotelian subject/predicate pattern, but having lost any link with substances), whereas the former conjectures complete inseparability of any belief from any other (in conformity with the logical homogeneity of all relations, but having lost any link with the actual structures of cognition). Both are overarching hypotheses about nature, both collide with our experience in everyday life and in scientific research.

Until now, I have listed a few considerations telling against the hypothesis that rational processes are holistic. An assumption has not yet been discussed, implicit in this hypothesis: *i.e.,* that belief fixation processes are inferential. In this regard, Gestalt theory, and more recently the theory of «mental models», see Johnson-Laird (1983), as well as connectionism provide arguments to the contrary, and even if these theories were wrong on other points, the evidence (concerning different areas of cognition) they have by now assembled against a purely logico-linguistic manipulation of representations as central to belief fixation cannot be underestimated. Furthermore, once a plurality of logics anchored to regional ontologies is admitted, 'formal' logic would no longer be such a privileged method for linking two arbitrary beliefs. Not only could the link turn out to be established according to a logic extrinsic to the given kind of objects, but also the way beliefs are activated, organized and manipulated could be governed by principles of a non-logical nature. (One may revise the range of what is commonly taken as logic, but this is another matter.) In fact the constraints characterizing my local holism are primarily not of logical, but rather of geometrico-topological nature; and since theories too are affected by localization, the notion of theory has to incorporate such geometrico-topological characteristics, in order to be more adequately understood than it has been till now through a purely logico-linguistic analysis, the output of which is extrinsically mapped to set-theoretic models. The conception of theories as categories and the categorial interpretation of deduction, cf. Peruzzi (1991a), are helpful in understanding how modularity enters central processes.

Both extreme compartmentalization (as exemplified in formalized versions of logical-empiricist *vulgata*) and global holism pass over the network of tuning procedures that interface the cohesiveness of the so called central processes and the local roots of any meaning. What secures stability for a dynamic system such as the human mind is not a monarchy of the Central Nous over slave peripherical modules, legitimated by the rigid behavior of input transduction systems, in contrast to the indefinite openness of reason: it is rather a question of constrained plasticity and interactivity. It is easy to adduce the 'frame problem' as an almost insuperable difficulty for those who intend to describe belief fixation as a local process; yet, as *Noema* made it clear, I am not saying there are unrevisable beliefs,

nor that, given any empirical domain *U*, there exists a *unique* set of beliefs $\Phi(U)$ canonically associated to *U* : what are canonical are rather the constructive operations underlying reference to *U*.

On the other hand, the holist cannot delight in contemplating the recursive unsolvability of the 'frame problem', for he is simply substituting for it the unsolvability of the 'expedition problem': wherever the place we wish to explore, we won't be able to start, since, in order to start, we must have the equipment suitable to features relevant of the place, but before getting there we ignore what these features are, and if we knew all of them, there would be no need to explore it. (The classical statement of this position, of course, is Plato's *Meno*.) 'Global' notions are important, but *in the appropriate ways*. Moreover, global features are sensitive to the internal/external dichotomy, relative to any given referential system. In this connection, the following diagram may be helpful in fixing ideas.

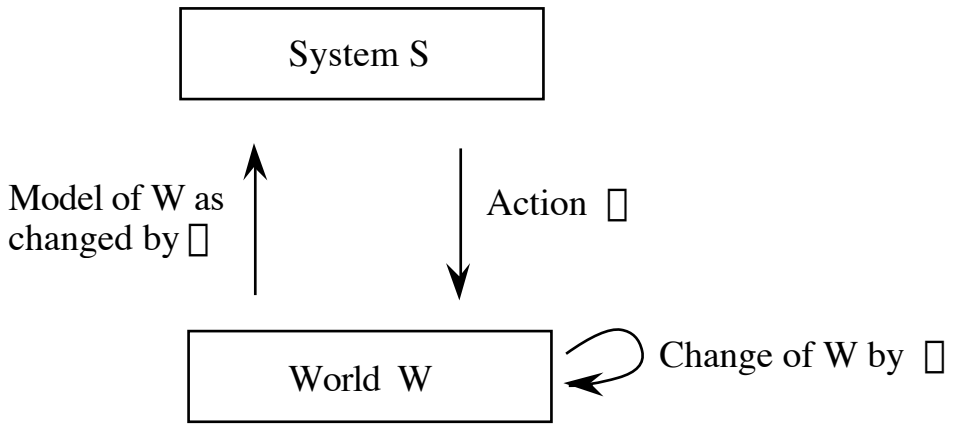


Figure 3

(Obviously, *S* is in *W*, but it is pictured as distinct from *W* for easiness of image.) The question consists of endowing *S* with two capacities:

1) (re)planning its action on the basis of how *W* changes;
2) taking account of possible alternative model/action pairings and their results.

Heuristics involved in 1) and 2) furnish the way to limit the combinatory explosion that holism, rather than being unable to avoid, directly theorizes. Heuristic procedures presuppose a *space* of the problem and a *space* of descriptions, linked to localization of . But by which criteria should they be governed? Subroutines involved in human heuristics seem to be centered on *Gestalten* of various kinds: stereotipical patterns that have been

investigated in AI (as frames, scripts, etc.) are only a parasitic example of basic prototipical schemes. Already in logical analysis of language variational patterns and related thresholds have become manifest in counterfactuals — in view of their *ceteris paribus* clauses. Even limiting ourselves to the logical point of view, research on non-monotonicity (implicit in the diagram above) has shown that it can be dealt with rigorously – or rigorously explained away. Finally, the very applicability of mathematics, with its modular architecture, excludes a globally holistic heuristics. To summarize, the frame problem demands a finer mathematical description of symbolic *Gestalten*, a problem addressed in Peruzzi (1992).

Optimal flexibility does not mean total: this suggests that holism inherits a disguised form of skepticism, specifically coming into collision with psychology of thought and the very feasibility of simulation models (for pattern recognition, etc.), which tell us that reality (knowledge included) cannot be represented as an enormous amoeboid net, ready to support any grid.

Thus, whereas Fodor considers modular systems as totally encapsulated, I do not: consider their built-in coherence. Conversely, I do not take central systems as totally isotropic and Quinean (they are not intrinsically global): consider their built-in localization. In recent studies on the ontology of the common-sense world, a whole set of constraints and *pairing constants* is more and more recognizable to be of an algebraic-topological character, not of a logico-linguistic one. They permit us both to identify an intermediate level, at which we can open cognitive windows onto the shallowness of input systems, and to sufficiently guarantee impermeability to the Neurathian boat of our knowledge.

§ 2.2. Meaning Holism

We have seen that, extending the Chomskian view of language as an innately determined organ, independent from general intelligence, Fodor advances a theory of mind as a net of modules, each operating without interference from the others, supervised by central processes. Moreover, he claims that there are innate universals to which all possible meanings can be traced back, and that his methodological solipsism is a necessary condition for meanings being legitimate objects of scientific investigation. Although Fodor (since 1983) has accepted the holistic view of confirmation, he has denied that meaning holism can be derived from it.

Just this separation between a (holistic) component of belief fixation and a (non-holistic) component of psychosemantics is attacked by Putnam who criticizes the very notion of psychological content. In fact, holism is a multi-faceted view and shows up differently in the work of different philosophers. Here I confine attention to Putnam's

version, in order to specify the rationale of semantics for the Kantian heritage in epistemology.

Against mentalistic semantics, Putnam joins the arguments for meaning holism and his criticism of functionalism (to which he widely contributed in the sixties). Having argued against global holism on belief fixation, I shall argue that it is possible to reject meaning holism too, although for partly different reasons. Both forms of holism rightly suggest a view of the language-world correspondence as *internal* to the bounds of experience, but at the same time, they put aside Kant's foundational project with respect to the phenomenal world. The price for this is that holism wrongly conceives experience as freely floating over an inscrutable ocean of Being, hence *internally* reproducing the classical *external* dualism that leads either to skepticism or to metaphysics. Moreover, the variety of ways in which the world can be captured by theories (elaborated by a mind) would *transcend* the world, in such a way that the freedom of reason is paid for by making it stranger to the world itself. Still, there are significant differences between theories of belief and theories of meaning, and they affect the kind of transcendence involved. Thus, let us proceed first by considering Putnam's criticism of meaning «in the head», and then his appeal to Quine's indeterminacy thesis as a basic support of meaning holism.

From *The Meaning of 'Meaning'*, now in Putnam (1975), many variations have been elaborated on the same theme (the gulf between mental representations and reference) by means of possible worlds semantics. Here, just a sketch of this theme will be given, in order to start the discussion and confine it to its essentials.

We are asked to consider as a concrete possibility that there are two places U and V, subject A lives in U and subject B lives in V, and U- and V-environments are phenomenally indiscernible for A and B. Yet, similar things called 'N' by both A and B are, *really*, examples of the natural kind K in U and K' in V, with $K \neq K'$. So, by using N, A and B have the same mental contents and the same beliefs, while the same sentences containing occurrences of N differ in meaning (for us as external observers, for A and B at later stages of their knowledge, or 'objectively'), since N has a different extension in U and V, and difference in extension implies difference in meaning, thus the meaning of N cannot be what A and B have in mind: *meanings aren't in the head*, Putnam concludes. Rather, they are grounded on sterotypes, as socially shared patterns of reference, contingently adopted within the given linguistic community.

Chomsky, Fodor, and others have replied by defending the strict link of meaning to mental content. However, from a functionalist point of view one has to concede that even when A knows that N refers to something different from what N refers to for B, this fact leaves the reference of N for A unchanged (A has just more information), but then which and how many changes in collateral information are sufficient to change reference? And how can we solve the problem of synonymy for 'mental signs'?

Putnam's proposal is that sameness of stereotype is the only possible condition for synonymy, but then this notion becomes empirical and not completely internal to mind. Should one say that N has a different content for A and B if A believes that N is assigned kind K and B does not believe that? If so, one would simply start what Putnam calls the *infection* problem. It is impossible to list the potentially infinite manifold of contexts and interests pertinent to interpreting N, and taxonomic parsing is supposed to be intrinsically open. According to Putnam (1983), determination of the identity for stereotypes (were it the right strategy for translation) is «nothing like an algorithm», nor would sameness of *perceptual* prototype suffice, since meaning is a function of information that cannot be exclusively reduced to perceptual prototypes; thus, two subjects could assign the same meaning to a given term, and still possess very different perceptual prototypes for its reference.

Chomsky's reaction is to give up the principle according to which intension (Int) determines extension (Ext). So it might be that $Int_U(N) = Int_V(N)$, notwithstanding $Ext_U(N) \neq Ext_V(N)$. Putnam objects that Ext is only one (albeit primary) component of the «meaning vector», and hence the Chomskian proposal follows a wrong path: we would no longer possess a ground for deciding when there has been a change in meaning. Still, the admission that differences in extension are relevant does not mean that we possess an *algorithm* to compare meanings, hence a computational solution to the problem of radical translation, for «until we have started interpreting a language we don't have any idea what the extension of a term is», Putnam (1983) p.149. Echoing Quine, interpretation is essentially a holistic enterprise. It is impossible to formalize our criteria for successful interpretation, for it would entail formalization of *all* humanly intelligible contents. Nonetheless, Putnam thinks that interpretation is a subject of legitimate study: it simply cannot coincide with or be restricted to cognitive psychology. His *exempla* show that even if the functionalist possessed a complete verificational semantics for the language of mind (the so-called «Mentalese»), the problem of reference-preserving translation would remain unsolved: there are irriducibly *global* factors in semantic interpretation and they cannot be governed by rules of computation. Indeed, this is another version of the unsolvability of the frame problem, insofar as non-programmable (and non-monotonic) modifications of any given program are called for by translation.[8]

Putnam adds that any project for describing interpretation (for example, in physicalistic terms, as proposed by Hartry Field, in order to define reference) cannot avoid analysis of the *whole* battery of axiological notions (humanity, rationality, etc.), and there is

---

[8] As we shall see in § 3.4, a similar point has been urged by Hofstadter (1985) in the apparently more restricted domain of letter recognition. The existence of specific cognitive abilities involved in grasping the variations on any prototype implies revision of the thesis of Putnam under discussion, see Peruzzi (1992).

no relevant partial success in such an enterprise (as in physics or psychology) other than a better philosophical consciousness. In sum, it is utopian to project any algorithmic (and more, any mathematical) analysis of global axiological notions — which enter semantics in an essential way.

However sage this invitation to moderating scientistic pretenses might be, it is subject to at least three objections: the holism it embodies passes over i) the existence of objective constraints, in linking perception, language representation and strategies of reasoning, which bound the range of interpretations, ii) the local character of *epoché* (see § 3.1 below) and the related constraints on *ceteris paribus* clauses, iii) the fact that knowledge is made possible by the separability of objects, and systems of objects, in our environment, even if separation is not (and cannot be) complete. To make these points more concrete, let us consider Putnam's criticism of functionalism.

Here is a (slightly adapted) list of his arguments:1) if the human mind were a Turing Machine (TM), its states would be instantaneous, whereas there are states (e.g. jelousy) that are not such; 2) states would be globally defined in a unique way, whereas we often experience different simultaneous and independent activities aimed at even contrasting goals; 3) they would be *i*ndependent from changes in memory, whereas memory units are not static or permanent, being progressively restructured (cf. G. Edelman's theory); 4) any TM is discrete, whereas there are laws such as Hull's laws concerning learning, that specify analytical relations between continuous variables and they might turn out to be correct in their analytical formulation; 5) mental states would constitute a causally closed set, whereas intentional phenomena manifest intrinsic openness. (Additional criticisms, such as that the computational mechanisms involved are deterministic, as well as learning independent, appear by now outmoded by recent developments in computer science. In any case, these arguments do not reduce the strenght of other arguments in favour of functionalism.)

Undoubtedly, functional properties are relevant in human thought, so the famous example of the square peg and the round hole, see Putnam (1975) ch.14, applies to a wide range of cases where what matters is form, not substance. Putnam has successively advocated a many-many theory of the correlation between functional states and mental (intentional) states, instead of the initial one-one theory. That is, just as the same functional states (and programs) correspond to several material instantiations (implementations), so the same thought can correspond to several functional states (programs).

The relevant point is that Putnam adapts the arguments against the thesis that meaning is in the head to a refutation of functionalism: given that propositional attitudes are linguistic manifestations of mental states, if we cannot individuate concepts and beliefs without reference to (changing interactions with changing) *environment*, it follows that mental states are neither computational nor definable in terms of internal parameters of the mind-brain.

This same adaptation is still consistent with an intersubjective view of concepts and beliefs, one which agrees with Putnam in denying that propositional attitudes are computational states of a separated system, cf. Peruzzi (1987). Putnam terms this solution *socio-functionalism* as seeking to describe meaning and reference (and any other intentional notion) as functional relations within a larger system than the single mind, namely within a society of organisms-in-an-environment. He objects that an ascription of meaning to (the linguistic expression of) someone's representations requires interpretation, and to provide this is possible by *discounting differences* : that is, we assign meaning to words for natural kinds, like 'rabbit', 'water', 'gold',, in spite of the fact that other speakers, on the ground of *our* interpretation (by means of some Quinean translation manual), have different beliefs about them, until this disagreement becomes so great as to call for a revision of our interpretation (translation), the only criterion of choice being «reasonableness». So Putnam (1988) says that it is *utopian* to suppose that the latter concept can be formalized in order to define meaning and reference in functionalistic terms. This entire argument rests on two supporting points: indeterminacy of translation and indefinite openness of the mind. Since I have already discussed on the latter in ch.1, the time has come for some comment on the former.

The indeterminacy thesis claims that for any language L and for any sentence $\square \square$ Sent_L (for example 'Here is a rabbit') there exists an open class of functions $f, f', \ldots: \mathrm{Sent}_L \square \mathrm{Sent}_L$ s. t. the interpretation of $f(\square), f'(\square), \ldots$ ('That is an undetached part of rabbit', etc.) is indiscernible from that of $\square$, on the base of the given behavioral evidence e (however large it is). That is $\square \square_e f(\square)$.

In the case of radical translation $\square: \mathrm{Sent}_L \square \mathrm{Sent}_{L'}$ the thesis implies that $\square$is indeterminate; *viz.*

$$\square(\square) \square_e \square(f(\square)) \square_e f'(\square(\square))$$

(in the example, $\square$ is the exotic 'Gavagai').

Quine strengthens this thesis in the «inscrutability of reference» (so, the «darkness» of meanings, as entities defined by abstraction through synonymy, affects reference too): if the extension of a predicate *P* is indeterminable, there is no fact of the matter about what objectively *P* refers to, as much as there was no fact of the matter about which predicates are synonymous with *P*.

This incursion in matters of translation would be unjustified here if Quine had not widely used a parallel between radical translation and language learning.The way in which Quine elaborates this analogy between linguist and child (as 'theoreticians') reveals how deeply goes the inseparability of language from knowledge and cognition: Quine's account of the ontogenesis of reference in the child admits that interpretation of any term comes to

be fixed at most relatively to other terms (by their co-occurrence in sentences of Mommy's language, sentences being the basic units of language), and ultimately to ostensive acts that are no less subject to indeterminacy. But even if they were determined we could not identify precisely the interpretative options adopted by the child in early learning, see LePore (1988).

Yet, the basic relation between words and things is not so open, and inscrutable, once we cease to limit ourselves to behavioral considerations, but use all information at disposal on the local environment and the inner cognitive architecture, to analyze the various components of perceptual and verbal competence of children of different ages, thus trying to do justice to the amount of knowledge that cognitive psychologists have accumulated about those very processes that parallel (or underlie) the development of linguistic capacity. In particular, what behaviorism ignores is that higher learning occurs by deconditionalization: it is necessary to let a whole set of parameters vary — a set wider than that attainable in a S-R mechanism, although it avoids arbitrary variation. Here the phenomenological openness to *Spielraum* and *insight* comes in; if the set of parameters were completely free, no methodological option would remain other than behaviorism, by now discredited in cognitive science. On the other hand, it would indeed be utopian to ask for a science of mind along present-day computational lines.[9] A new perspective is needed, and one of its tests remains language learning.

Confining attention to nouns, we have to distinguish at least four levels in meaning formation: 1) the level of universal traits, basically physico-geometrical (as curvilinear, vertical, closed, bright, hard, etc.), absent from the first lexicon of the child but making it possible; such traits do not exist separately in the external world; 2) the level of hook-nouns (such as bowwow, spoon, water, etc.) that are part of the first lexicon and can be mentally represented by a prototipical scheme, assembled through combinatory operations on traits of level 1); such nouns refer to *basic* kinds; 3) the level of elementary taxonomic hierarchies of sub- and super-ordinate kinds, based on perceptual qualities; 4) the lift-level, that of abstraction from perception to cognition, at which it is possible to objectify relations and give rise to *free* symbolic manipulation *in absentia* (when referents of symbols are lacking).

Levels 1) and 2) *come before* the division of linguistic labour, before translation and synonymy, while level 3) is early influenced by cultural biases. Moreover, the adequacy of any given semantic model depends on which of these three levels is involved. For example,

---

[9] One might think that connectionism resurrects some tenets of behaviorism, and therefore object to my previous appeal to connectionist aspects; the point is that such an appeal neither treats mind as a black box nor reference as inscrutable. Finally, the aspects involved are at the level of performance, although I agree with David Marr in questioning the traditional demarcation line between competence and performance.

it is dubious that level 1) supports the IPM of mind. At level 2) a strong indexical component is manifest, while at level 4) one can find semantic units combinatorially defined by their relation with other units, already lexicalized.

If meaning is a function not of just one sort of variables but of many, then its projection on the purely representational component is essentially *unsaturated*. In particular, the fact that, for any designating expression e_i , Int(e_i) is a function of the actual Ext(e_j), for some other expression e_j , reveals another variant of the 'frame problem': on the one hand a term (as Figure) needs a modal horizon (Background) in order to be projected onto new referents, on the other hand the *relevant* possible worlds (or situations) are those compatible with setting and keeping fixed certain actual objects and their salient properties. What emerges is that, on level 3), we have at the same time

$$\text{Int}\,(e_i) = g\ (\text{Ext}\ (e_j), \boldsymbol{u}) \qquad \text{and} \qquad \text{Ext}\,(e_i) = g^{\wedge}\,(\text{Int}\ (e_k)\ , \boldsymbol{s})$$

where $\boldsymbol{u}$ stands for some list of prototypes, and $\boldsymbol{s}$ for some list of indexically anchored sortals, g and $g^{\wedge}$ being suitable functions, codifying patterns of internalized procedures for tracing identity of individuals and kinds (the case $i = k$ is revealing)[10].

When holists tell us that what is at stake is the *interpretation* of any text as a whole (translation concerns interpretation of texts in a different language), either they are saying something obvious (any novel, fairy tale or scientific work has always required such an activity: understanding is not point-like) or their suggestion puts us in the condition of being unable to distinguish the referential background of a sentence like 'It is warm' and that of a complicated theory or of a mythology. This indistinguishibility points in the direction of hermeneutics. Unfortunately, that would prevent us from achieving the aim with which we started: to explain how it is possible for the child to refer and understand others' referential expressions. The child's task is *finitary*, and its explanation has to be so too, whereas the information required by the holist is not. (It is not haphazard that skeptical consequences emerge as a corollary of holism.) If seeing Mommy coming near and referring to a ball are assimilated to seeing a fossil in a stone and in motley of blobs a foliage from which a gun is emeerging (as in Antonioni's *Blow up*), then hermeneutics is coextensive *tout court* with everyday's experience, and not only with that of Sherlock Holmes, literary critics, historians and jurists. In this assimilation more than a democratic glimpse is given at interpretation, there is also the desperate effort to appease the nostalgia for the mediaeval world: a forest of symbols crafted by God. On that view the world had a sense, since it was the book written by God, where everything has its definite meaning. In

---

[10] For simplicity's sake, I have omitted the possibility that g ($g^{\wedge}$) depends on a set of extensions (resp., intensions).$^{\wedge}$

our times, unless one says with Heidegger that man is 'spoken' by Being (as a depersonified *logos*), the death of God compels man to take the same task on his own shoulders. What better way of granting success to such enterprise than by saying that nothing can be said to exist outside the net of human languages? Perhaps, man is no longer the master of nature but on this view he is still the owner of its sense. Perhaps, Kant's orphans have lost his power over truth but still are the owners of the key to disclose its sense. Davidson offers to free us from the dogma of the conceptual scheme, a dogma of empiricism no less than of Cartesian rationalism and Kantian idealism; but it seems that we have to embrace the mutual dependence of everything: language, mind, perception and reality are what they are only within one and the same intersubjective mega-frame, in which communication and interpretation are the central activity.

In making this proposal, however, Davidson is strangely aligning his position with that of the contemporary desubjectivized version of idealism made familiar through Continental European structuralism. The (repressed) Romantic impulse towards fusion with the Whole has simply lost the optimism and the Subject-driven character of the original. Unfortunately, the outcome of European structuralism has been no less aporetic than that empiricism from whose third dogma we had to be set free; by itself, the shift from subjectivity to intersubjectivity (be it *de facto* or *de jure*) has never changed the substance of any epistemological problem.

In Europe, similar views have widely exploited an attracting device, the so-called «hermeneutical circle»: understanding the parts of X is possible only thanks to knowing something about X as a whole, and vice versa. But granted the ubiquity of language, any X may be reduced to a text. So hermeneutics connects analytic philosophy via the back door. (More on the same topic in §§ 3.5-3.6.)

If, in order to specify the meaning of any sentence as «The cat is on the mat», the holist holds that it is necessary to make explicit the context of all presuppositions of all *possible* sentences, either he/she is renouncing rationality or he/she is advancing a request that *obviously* cannot be satisfied, unless by God. In either case, the holist is theorizing the unsolvability of the «infection» problem, instead of seeking a solution to it. The Kantian world has fragmented into a multiplicity of world views, becoming another forest of symbols, but one that is now either suspended over an inscrutable ocean of Being, or transcending any given (finite) symbolic act: no path through the forest is orientable. Hence the conclusion that there is no way to identify the literal meaning of any sentence — if this meaning exists, it is unknowable. In principle, understanding each other is simply participating to the same game, in fact it is an illusion fostered by habit or an unending process: from Kantian certainty we are back to Hume's skepticism. Furthermore, holism leads to consequences that make our daily linguistic practice impossible. Indeed, the meaning of any sentence □ is intrinsically *by default*, and what is relevant to the

interpretation of □ is, on each occasion of use, a finite segment of background information, which is not provided by sets of verbalized beliefs of a baroque complexity, but rather by prototypical schemes of language-world coordinations. These coordinations are strongly linked to features of the local macro-physical environment and thus are not of linguistic origin. The case of rival cosmological theories, differing in their global hypotheses, does not alter the point: any test for them is locally focused.[11]

Moreover, if the meaning of □ is to be regarded as indeterminate on the ground of information codified by □, one can object that no intentional horizon or bio-cultural background would be able to fix it in a unique way or even approximately, for the simple reason that the sequel of increments to a larger and larger range of possibilities may *augment*, rather than reduce, semantic indeterminacy; for at each step one might have to sum □'s indeterminacy to □'s, and so on. Therefore the holist has to demonstrate that it is not so, *viz.* that over a certain threshold, the complexity of the network reduces indeterminacy sufficiently to make it compatible with our linguistic practice: meaning becomes identifiable with functional role only if the critical mass of the semiotic system is reached. Even if such a proof were available, I doubt that it would support holism: the reason why such progressive constraints are effective it that default assignments are the mark of over-, not under-, determination of meaning. Consequently, the explanation has to be sought in a completely different direction.

To begin with, we need to draw a basic distinction between cultural stereotypes and perceptual prototypes of natural kinds. Instead, Putnam collapses the two notions and insists that stereotypes are not preserved in translation: e.g., in old Siam, Siamese cats are just cats. But if we enter the argument for holism he develops by appealing to this fact, we find that language is conceived as an enormous stock of predicates, floating above the world and freely combinable from the very start: this is the latest effect of Humean acid, of structuralism in linguistics and anthropology, and of formalism in logical analysis. On this view, it is as if, in order to mirror the supposed unlimited freedom of nature, there were no place left for *any structure* itself rooted in nature, in terms of which to account for «associations of ideas». Yet, we are able to understand that a given stereotype has been altered in translation, and this is possible because of a stable background of other perceptually fixed parameters. Undetectable changes presuppose a separation of theory and experience that the holist denies. But, to detect change in stereotypes, constancy of prototypes is required.The holist is not entitled to separate Figure and Background in any such intrinsic way. But if this separation is merely a matter of convenience, how can the

---

[11] Here the need of clarifying the deep significance of the philosophy of geometry for the philosophy of language becomes manifest, since a theory of coordinative definitions, stressing their non-linguistic character, as suggested by Parrini (1976), would have helped in avoiding such skeptical consequences.

holist avoid assimilation of this view of objects and their properties to the Humean view of events and their causes? Indeed, in a globally holistic world, there cannot be causes, since every event E that occurs depends on everything else that happens anywhere — or in an apparently more prudent formulation: we could never say where the border between the events affecting E and those not affecting E lies. Still, since holism denies any border between matters of language and matters of fact, the holist is obliged to apply the doctrine not just to language and cognition, but also to the physical world. Thus this view *extends* Mach's Principle. Then how can holism be made consistent with the theory of relativity? Moreover, if causes cannot be said to exist, neither can things, as macro-physical units, since the notion of such units presupposes stability of a given set of properties relative to a specific background. This would permit, for any given thing, a division between events allowing the preservation of the properties in question and events preventing such preservation. Finally, even apart from the lack of a precise examination of the dialectics of local and global involved in macro-physical interactions, a further basic defect common to holistic and skeptical arguments is that they are not *smooth*, since they fail to take into account any differential-geometrical constraint: this is what makes relativism differ from relativity, and what distinguishes the (scientific) explanation of action, by means of local contact, from the (magical) pseudo-explanation, by means of trans-local analogy.

§ 2.3. Against reductions of the mental?

We have seen that, according to functionalism, mental states cannot be identified with physico-chemical states of the brain. Then we have seen that for the more recent Putnam neither can they be identified with computational states, because of the intrinsic plasticity of mental states (however they might be said emergent and supervenient on causal/computational states). The same mental state can be manifested in ways that possess different computational structures. Therefore, thoughts *cannot* be programs — or defined in terms of programs.

To reach this conclusion, Putnam appeals to meaning holism, by contesting the separability of any semantic domain from general intelligence: the notions of 'same meaning' and 'same reference' are the product of the whole network of our beliefs. Yet, the above reservations on meaning holism re-inforce a different view, according to which semantic universals are of a gestaltic nature and the way in which they act as *role-schemes* in language can best be described in terms of category theory (since predicates are not all on the same footing, there is a basic level in typing 'individuals', and basic *actions* on these types are essential). Without entering the technical details of such a description, I wish to

stress here is that semantic universals pertaining to levels 1) and 2) of p. 39 embody constraints on the Figure-Background pairing, and these are not confined to visual perception. *How* in cognitive development these universals are activated, and *which* cognitively privileged structures are involved, are problems to be dealt with by observations, experiments, and computer-aided models (simulations).

Meanwhile, we can recognize that the architecture resulting from such universals paves the way to a geometry of mind, not to a *lingua mentis*, consistently with their geometric root in kinetic experience. This perspective is precluded to Fodor, who, at least until recently, was appealing to holism on belief fixation, at the same time claiming that *all* meanings are innate. True, the existence of severe obstacles in relating examples of changes in linguistic meaning to such presumed language of thought, as well the fact that meaning is interactive, depend «not just on what is in our head but also on what there is in our environment and on how we interact with that environment», Putnam (1988) p. 18. Still these facts are not obstacles to the view I propose, for it needs neither the hypothesis that meanings are innately determined nor the hypothesis that mental representation determines reference. Therefore, Putnam's arguments against functionalism do not apply to the semantic geometry I regard as providing the route out of the aporias of holism.

As to the cryptographic model of meaning according to which to understand a proposition □ equals decoding □ in the language M of mind (M for «Mentalese»), Putnam reminds us that, in order to avoid the usual *homuncularis* regress, it has to imply that in M sign and meaning coincide: semantics = M-syntax. Then, if e.g. «rational animal» and «featherless biped» express different concepts, there is no need to bring into question modalities and possible worlds in order to distinghish them. If any two concepts *c* and *c*' are different signs in M, they automatically differ in meaning and reference: M is really the ideal language searched for by Frege! Yet, if such a universal representational system exists, it does not seem to be purely logical, and anyway, one can doubt that it exists *as a language*. [12]

If meaning has other dimensions (extralinguistic and social) irriducible to those of mental representations, it may be the case that different representations contribute in forming the same meaning, under the relevant respects. Meaning (in the head) does not determine reference, so much the less does reference determine meaning, but since language learning (as reference fixation, *in primis*) is an essentially intersubjective experience, no meaning is ever purely private, and since stability of reference requires intermodular coordination, the source of meaning is distributed. For Putnam this

---

[12] Language learning does not *create* concepts, rather it latches on stable clusters of protoconcepts in order to manipulate them. Formation and stability of the lexicon demands, so to speak, an horizontal cartographer, not a vertical criptographer.

distribution is maximal, whereas I hold that it is minimal, but not atomistic, so that meaning and reference are mutually constrained in definite ways.

Let us consider a last concession to the functionalist project of defining the basic notions of semantics: such a project might recognize that *in practice* we are unable to achieve the goal of the definition: still it could be done *in principle*. Putnam replies that the appeal to 'in principle' arguments, largely used in scientific *reduction* (e.g., temperature defined as mean molecular energy) is not satisfying in the present case since any proposed definition of the given kind would be limited to *contingent* physical-anthropological-computational features: it might be false in other possible worlds.

I have mentioned this 'in principle' way out to make clear that it is *not* the solution proposed here. For I am not saying that, although we are in fact unable to *define* reference, co-reference and synonymy in functionalistic terms, such a definition is still possible in principle. Of course any theory needs idealizations, but they have to provide *smoothly* perfectible approximations to data. That we cannot hope to possess the «Master Algorithm for Interpretation» could simply mean that the algorithmic idealization is not the right one to describe the mastering of semantic notions — nevertheless, it remains profitable to develop computer simulations for *some* features of this mastering, as well as different mathematical models of other features. After all, contemporary mathematics is not exhausted by logic, set theory and computing!

Moreover, to ascribe, with Putnam (1988) p.92, to the functionalist the project of «making sense of the propositional attitudes as *possible psychological* states of any physically possible organism» is to impute to her/him with an obviously excessive pretense, though one motivated by the aim of avoiding any species-chauvinism. Rejecting the recourse to folk psychology as an unsatisfactory explanation of propositional attitudes does not entail the renunciation of any theory of universals for the common-sense world, with their plasticity sufficient to support a rich combinatorics, involving also belief, seeing, etc. Were their nature non-computational, still it would be useful to evaluate the degree of non-recursive character of mental states, approximating mental states as suitable quotients defined by an equivalence relation between computational states. The open-endedness of intentionality has to be bounded in some way, otherwise the combinatory explosion of mental-cum-environmental states would stifle any process of belief fixation at its birth. Just these bounds make openness efficient. True, in almost any real-life situation we proceed «discounting differences» (we make quotients, that are kept as long as they do not turn out to be too fine or too rough, depending on the actual experience), however, in forming these quotients, our goals are not free-floating over nature, and this is not because nature itself is a construct of mind. For example, suppose we have two Turing machines M_1 and M_2 having different number of states and different transition rules; furthermore, suppose there is no homomorphism between M_1 and M_2. Still, we can consider substructures of states-cum-rules

that give rise to concording exits, as equivalent. The question is reduced to the evaluation (and improvement) of any given, formally modelled, quotient, and this depends on embedding the implementation of both M_1 and M_2 in an environment extremely rich in information.

On the contrary, by elaborating Quinean arguments for semantic indeterminacy, and by appealing to the Löwenheim-Skolem Theorem, Putnam (1983) p.IX extracts from the impossibility to characterize reference in a categorical way new evidence for rejecting naturalism :

> no matter what operational and theoretical constraints our practice may impose on our use of a language, there are always *infinitely many different relations* (different 'satisfaction relations', in the sense of formal semantics, or different *correspondences*) which satisfy all of the constraints.

Therefore, any pretense of knowing what is *the* reference of nouns like 'table', 'cat', etc., is grounded on an unsatisfiable premiss: we should have proved categoricity for the relation 'refers to', but we provably cannot. Perhaps, if reference is not fixed by anything mental, could it be fixed by something extramental? But no, since the idea that «*nature itself* determines what our words stand for» is totally unintelligible (*ib.*, p.XII). For it would mean that our signs are natural (bringing us back to mediaeval essentialism), whereas «there isn't a ready-made world».

Yet, it does not seem so unintelligible that there should be constraints on possible signs, built into nature, once attention is shifted from campus communication to vital, basic, prototypical concepts of forms and actions, whose stable complexes are represented in a child's lexicon, and to the processes involved in their acquisition by living beings on the planet Earth. In realizing this there is no harcking back to mediaeval essentialism, rather the anticipation of a non reductive naturalism, that focuses on physical, perceptual and conceptual attractors, some of which are already investigated by synergetics, chaos theory, thermodynamics of dissipative systems, catastrophe theory. But if in many kinds of natural structures there are *objective* attractors whose manifestation depends on boundary conditions, and they can be rigorously dealt with, then it is not clear at all why meaning, as a natural structure, should elude such analysis. Thus one is no longer obliged to conceive of holism as the only descriptive pattern complementary to reductionism — emergence can be smoothly interpolated.

If, on the contrary, the language-mind correspondence is not univocally fixed by anything and there is no formal or phenomenological constraint that at least restricts the range of possible interpretations, how did we arrive at our common experience unless by appealing to some pre-established harmony, to some continuously renewing miracle, or to

some deceiving devil? The only thing one can say in such hypothesis is that we live in an extraordinary world! At this point philosophy can stop: as with Wittgenstein's stairs... After all, our mind could turn out to be too superficial to explain so deep a universe. Hence we should be less pretentious. But imagining that the limits of scientific adventure stem from the theory of reference is also highly pretentious.

It must be added that Putnam rejects the Quinean doctrine of ontological relativity, since it would make the notion of object something metaphysical. The question then becomes what can one substitute for that notion after having appealed to the Löwenheim-Skolem Theorem as blocking any realist semantics. A way of internally constraining reference (and truth) is grounded on conditions of assertibility (an axiological notion); that is, on the possibility of determining the correctness (or not) of any sentence in specified conditions. Michael Dummett's has projected a constructive and systematic analytic philosophy, apt to solve the problem of identifying conditions of justification for any given sentence of natural language. However, the by now classical objections against verificationism spring up once again. Note, by the way, that one could repeat the *-argument as for 'table', 'reference', etc., also in the case of 'correctness' and the related *evaluative* (logical and methodological) notions: and this would be the end of the story, for then we would ignore what we are talking about in philosophy.

In all Putnamian *exempla,* so masterfully elaborated, some notion is allowed to vary, while keeping *constant* some other notion; but, in conformity with holism, the strategy can be reversed, and then applied also to the philosophical notions used in describing the examples. If any of my words might be understood by you in an observationally indiscernible sense, but still in a *different* *-sense, then the same should apply to the whole logical and epistemological vocabulary within which Putnam's arguments are formulated, in particular the one concerning his *description* of examples such as table/table* etc. Once more, the gain of such a holistic strategy is dubious. In order to be understood, meaning holism has to put in parenthesis its very assertibility. Therefore, how can it be true? (Functionalism can still turn out to be false, but on grounds other than holistic ones.)

The point established by holists is that no philosophical ism or anti-ism is immune to progress in the formal machinery, in the experimental devices, in new theories and data. But this claim applies to holism too. After all, isn't it contrary to the positive claim of holism to pretend that we can discuss science, meaning and mind only on philosophical grounds, and to pretend to reach firm conclusions in this way?

In the light of his view of rationality as historically evolving, referential internalism, and use of the Löwenheim-Skolem Theorem, Putnam (1983) admits that the equation truth = idealized justification is vague, but it is «as vague, interest relative and context-sensitive as *we* are» — although Putnam recognizes that «our notion of rationality cannot be quite as flexible as Quine suggests» (*ib.*, p.138). Hence it is impossible to specify ideal (well-defined

and uniform) conditions of justification. Nevertheless Putnam wants to transform this stalemate into a victory, and more than that, in the supreme victory, that of man's freedom over any attempt at reductionism and regimentation: if these are impossible in semantics, formalization of intelligence is impossible too.

So, even if unready to be captured in any given formal system, ideal assertibility remains a meaningful notion. We would finally have a realism by the human face, an «internal realism».[13] The notions of truth and reason, so understood, are meaningful as Kantian *ideas* (by their transcendence and their regulative power); however, having abandoned with Quine any *a priori* notion, those *ideas* rest on a vacuum, until we adopt a naturalistic conception of the transcendental level as well as of the guiding values involved. Though this move can be interpreted by traditional epistemologists as missing the sense of the transcendental project, in its original form, only such a conception can account for ecological and structural constraints on the *formation* of concepts and their semantic role. This applies also to the concept of truth. Belief, rationality and the other intentional notions are irreducible to non-intentional ones, in the sense that they *emerge* out of causal/computational levels through structures of perceptual coordination — rather as laser light emerges by phase coordination. Then, assertibility and truth are no longer seen as being of a radically different character from that of any natural fact.

Here one could raise an objection: on one side I accept the emergence of intentional notions, on the other I seem to commit the naturalistic fallacy with respect to axiological notions. The point is that I am contesting *a certain idea of nature*, *viz.* that of a fixed and closed system, incapable of autoproduction of order, oscillating between mere chance, blind determinism and pre-established harmony, in any case a nature from which man is a separable entity (unless the structure of the system is a byproduct of thought) and has to be such, in order to preserve the gulf between the features of rationality (free will, etc.) and

---

[13] Contrary to what might appear on the basis of the identification of truth with idealized rational acceptability in *Reason, Truth and History*, «the suggestion is simply that truth and rational acceptability are *interdependent* notions», Putnam (1988) p.115. By his admission, in the 1981 volume Putnam gave examples only of the dependence of truth on rational acceptability, while «the dependence goes both ways: whether an epistemic situation is any good or not typically depends on whether many different statements are *true* » (*ib.*). This is exactly one of the theses I advanced in *Noema,* ch.4, as an answer to the one-sidedness of positions such as verificationism (in philosophy of science) and constructivism (in philosophy of mathematics). Thus my present disagreement with Putnam is not rooted in an acceptance of the reducibility of truth to other epistemic notions. Rather, it stems from the fact that Putnam's internal realism tells us that truth does not transcend use, whereas on my view it is unlikely that any use can exhaust the open-texture of truth, whose interdependence but not interdefinability with rational acceptability suggests precisely that the notion of truth, once emerged, will appear as internally transcendent — of course, it cannot transcend the set of (in principle) possible uses, but such a modal clause is empty, since this totality is completely indeterminate. Once more, the appeal to use (and its historicization) cannot provide any real explanation.

those of any contingent mechanism (be it biological or artificial). But if we succeed in grasping how the hexagon of a snowflake can be generated without intervention of a geometer, why should we not succeed in grasping how intentional notions may be understood as properties, no less natural, of the man-environment pairing?

Now, let us consider how internal realism is linked to an anti-functionalistic argument based on Gödel's Incompleteness Theorem.

Putnam interprets Gödel's Theorem as showing that «we cannot formalize our own mathematical capacity *because it is part of that mathematical capacity itself that can go beyond whatever it can formalize»* (*ib.,* p.118). More generally, reason is destined to transcend its self-image: a theme central to Gentile's *actualism*. But meaning holism holds that reference is a matter of interpretation, and this eludes complete formalization. So the reason why functionalism (and naturalism) fails is already encapsulated in Gödel's Theorem. Formal systems in mathematics, and programs in AI, are useful (and some are better than others), but none of them can *define* what is truth, justification, synonymy. Is this argument really victorious? At least three comments are relevant here.

First, it is to be noted that the initial argument against the computational reduction of intentional notions has turned into an argument against any formal definition of axiological notions, even if it has nowhere been shown how to prove that the range of a non-computational mathematical structure is empty. Actually, this is an intelligible notion, that provides many concepts, such as real number, choice function, compact cardinal, truth-in-a-language, etc. By this, I am not necessarily committing myself to either platonism or some constructive philosophy of mathematics, wider than a recursion-theoretic one but stricter than a classical one, even though constructivity remains an important aspect of the phenomenology of meaning. Nor am I suggesting that we rest content with an informal notion of mathematical use, as wider than any given computational device. Though «meaning is use» is not to be taken as a definition, being rather a sage *aperçue* pointing in the opposite direction from that of much analytic philosophy, 'use', if it is not to be a mere noise, always refers to something particular and actual, since a use transcending use is contradictory. We can study mathematically the costructions used by mathematicians, in the same way as we can study 'limitative theorems': as objects of rational investigation. Would it not be curious to say, *à la* Hegel, that the story of an historically evolving reason ended forever in 1931 with Gödel's result?

Second, if the formalization of intentional notions is impossible, why should the formalization of *velocity* be possible, given the multiplicity of velocities, and the multiple ways of specifying it in relation to the different sorts of phenomena we find in nature? That is, how much is Putnam's not to be seen as an argument for the impossibility of physics, and more generally of any science? If knowledge has to be considered as an undivided whole, the presence of undefinable ingredients in some part of it makes any

definition of any notion whatever in any other part illusory, something accepted just for practical convenience, unless any notion is definable by the whole system (once again the thesis of meaning as global functional role) although such definition cannot be finitely specified. In both cases we are left with a pragmatic fictionalism, as unpalatable in itself as it is hard to reconcile with internal realism.

Third, if socio-functionalism is contradicted by Gödel, since it would assume that there is a larger system (organisms-cum-environments) giving some ultimate account of all intentional notions, why should we hesitate to say that cosmology is impossible as a science? Analogously, shouldn't we apply the same *reductio* against the holistic view of the presumedly amoebiform field of epistemic forces.

May be *some* of the many possible versions of holism *can* find a way out of this impasse.With these three comments I only wish to stress that more caution is needed in appealing to logical incompleteness as a paradigm for a theory of reason. The notion of «All the Objects There Are» is questionable (though not 'empty'), but on different grounds: any totality can only be locally specified and may be expanded only by means of qualifications motivated by actual needs of life and science, otherwise the identification of any universe of discourse would be questionable too. The holist *could not* be content with that emptiness, since one cannot invite us to abandon our limited perspective on any given subject matter and then blame us for following the invitation too far. The holist never explained how it is possible for a child, during the first years of life, to develop the basic referential capacity and to master intentional notions; if there is such a global open-texture to every notion, explanation is hardly possible: once more, we should just be reduced to saying that we live in an extraordinary world. My hypothesis is, on the contrary, that there are intrinsic constraints on the structure of subject-object pairing, and that these constraints are natural in origin, hence of a different character from idealistic readings of Gödel's result and from the consequences inferred according to a formalistic conception of mathematics, seen as separated from the mind and its location in *this* world.

In sum , Putnam establishes a four-fold connection among: 1) the non-algorithmic character of meaning (holism); 2) the non-computational character of mental states (anti-functionalism); 3) the non-criterial character of rationality (historicist openness of intentionality); 4) the non-descriptive character of value (anti-naturalism). Items 1) and 2) have already been discussed; now, putting some of my preceding arguments together, I wish to conclude with some remarks on 3), leaving an analysis of 4) to ch.3.

In XXth century philosophy, the most influencing bearer of a criterial conception of rationality has been the logico-empiricist Principle of Verification, condensed in the slogan «The meaning of a sentence is its method of verification». A Niagara of ink has been shed on it, and so many criticisms have been raised against it that one wonders why anyone

should bother again. Putnam answers: the objections to this principle are objections to *any* criterial conception of rationality.

The view underlying the principle was: every statement is either analytic or empirically testable or nonsense. But by this same view, the principle of verification is self-refuting, since it is a criterion that is neither analytical nor empirically testable. Therefore, it should be nonsensical. Now, exactly the same gambit can be applied to any other overall semantic criterion. (Logical empiricists conceded that the principle of verification was a proposal, not a declarative assertion, but they *argued* for it; it was not a primitive rule.) The limitation it shares with any other criterial view of rationality is that it embodies the idea of philosophy as grown out of some methodology. But also another thesis, and one bound up with a radical criticism of logical empiricism, namely that of *incommensurability* (originally formulated in relation to competing scientific theories) is likewise self-refuting. Were it true, no translation would be possible between languages L and L' (or historical stages of one and the same language). Neither, by Quine and Davidson's arguments, would we understand foreigners as speakers: «To tell us that Galileo had 'incommensurable' notions *and then to go on to describe them at length* is totally incoherent», Putnam (1983) p.193. As far as meaning and reference are concerned, only admission of a possible translation (however inexact and partial) makes others' speech intelligible — the suggestion is: adopt a principle of charity to maximise, *à la* Vico, the humanity of our interlocutor. It is the possibility of translation which allows us to make sense of disputes about divergencies of *Weltanschauung*, both in comparative anthropology and in history of science. What can provide the ground for such possibility? My contention is that it is grounded in extra-linguistic facts of perceptual and ecological nature, not on a liberal meta-linguistic standpoint. Putnam's position is different. He argues that behind the radical relativism in vogue in the 197O's, after Kuhn and Feyerabend, a scientistic conception of rationality persists, for the latter was supposed to be modelled on an ideal computer program: whether we follow the rules associated with experiments in the physical sciences, or the rules of a given culture, the scheme is the same. Attacking this scientistic conception means displacing logic and methodology of science from their privileged position in philosophy. If this displacement contributes to recognizing how many different cognitive resources are involved in any form of knowledge, over and above those consisting in following pre-established rules and mechanically combining verbal symbols, then that is all to the good. But then why should we consider as the core arguments of philosophy those which rest heavily on areas of contemporary (standard) logic, such as the indeterminacy of translation, the non-categoricity of reference and the non-mechanizability of mathematical thinking? Isn't this the mark of an even narrower form of scientism, one assigning an excessive privilege to a highly technical domain over any other research field within the scientific enterprise? Setting aside the controversial assessment of the place of logical analysis within

our image of rationality, Putnam can be seen here as simply linking Quine's holism with Moore's thesis of indefinability of 'good'.

In effect, Putnam does not accept the way out that consists in saying that if neither norms of culture nor norms of rational acceptability determine whether a belief is rational, then an *ideal theory* of rationality covering all possible worlds (or stages of knowledge) would succeed. Such a theory, if it has to preserve the meaningfulness of the notion of rationality *for us now*, is not in view. He concludes that the problem of yielding any definition of intentional concepts like meaning, truth, etc. (all involved in rationality) «has no solution — and that *is* the solution», Putnam (1983) p.2O2. For there is no external, cosmic, point of view from which to state the required definition and establish its correctness. Still, we are not left with a position of unbridled relativism. Neurath's boat becomes a *fleet*, providing a metaphor not only for science but for all of culture, each of its vessels being dependent on others. Asking for more is merely symptomatic of «our old and unsatisfiable yearning for Absolutes».

This appealing argument is not new, however. It was a *Leitmotiv* of the historicist turn of Neo-Kantianism, and particularly recurrent in the works of Giulio Preti, who also subscribed, although with some significant provisos, to the criticism Moore raised about naturalistic views in axiological theories, as is widely attested by Preti (1976) and Preti (1986). The doubts expressed by Peruzzi (199O) about Preti's solution to the problem of reconciling rational, emotional and pragmatic elements of ethics, apply also in the case of Putnam: if the previous refutation of any criterial rationality is right, then neither general holism nor the indeterminacy thesis supporting it can be right, since they make essential use of logico-linguistic criteria as defining the range of what could or should be rational.

If this recourse to a non-criterial rationality has any convincing motivation, it is because it corresponds to the way we *are*. Were reason completely separable from our nature and from the nature of the physical environment where our existence became possible at a certain stage of cosmological evolution, how could *we* understand 'rational', 'indefinable', 'ontologically irriducible', etc.? At the same time this is far from leading to some absolute contingency. Instead of admitting that the world is miracolous or is the *sensorium Dei*, we have to acknowledge that rationality is not monolithic because the architecture of the mind is not either. The output of its modularity is flexible: we can let one component of the self play against another. Rationality manifests itself as a byproduct of the variety of evolutionary steps coexistent in ourselves, and this variety is inseparable from the way order inhabits nature and makes possible a system as complex as human mind.

Chapter 3. Phenomenology and Its Bad Company

Contemporary functionalism is grounded on the silicon proportion: mind is to body as software is to hardware. As the same program can be implemented on computers of different physical structure, so the features defining human thought cannot be reduced to neurophysiology. We have already examined some drawbacks of such view, and we have also confronted ourselves with Putnam's arguments not so much against the resulting dualism as against the attempt at preserving any criterial standard of human reason; therefore multiple realizability had to be freed from computationalism too. So, Putnam tries to lay the groundwork for a philosophical view of central notions such as reference and truth within a holistic perspective, enabling us to avoid any form of reduction of intentionality to natural facts. In chapter 2 such a project was criticized, though not in order to repropose the sort of functionalism involved in standard cognitivism and in classical AI. Now, I intend to link the concept, presented earlier, of the self as a sheaf of locally connected modules, with the inseparability of meaning from perceptual anchorage to topological universals. This link will be exploited to criticize the recent trend to join the pragmatist conception of epistemology and semantics to the historicist tradition. What will emerge in the end is a view of intentionality and reason as features of the architecture of complex natural systems. Albeit only occasionally explicit, the leading intuition underlying the chapter and conferring unity to the whole design of the essay borrows the treatment of variable structures and their 'gluing' from category theory, in the idea that categories provide the proper language for embedding semantics and epistemology into a new *philosophia naturalis* .

§ 3.1. Local *epoché*

In Greek philosophy, the term '*epoché* ' signifies suspension of judgment, that is abstention from assent (or dissent) to perception and belief, but it can also mean abstention

from assigning usual referential force to singular and general terms, thus affecting the assignment of truth-values to sentences, by bracketing the pre-judgements of common sense. *Epoché* was allocated the role of a basic *method* by Husserl in order to individuate meanings, even if he overweighted *epoché* with features and aims that will be skipped over here. The question is: can this suspension be global? Namely, can it affect the totality of language, giving us a self-sufficient domain of intensional entities?

Well, there is a sense of *epoché* that makes it global, but trivially so: when language is reduced to its syntax, or when a (natural or artificial) cognitive system works formally on its representations, taking these as computational units subject to symbolic manipulations independently from meaning. This is *not* the sense for which *epoché* is considered by Husserl,[14] for this sort of suspension is intended precisely as a device for reaching the level of pure mental *contents*. Thus: is a global but not purely formal or syntactic *epoché* possible? My answer is negative. Consistently with the previous rejection of Kant's dualism of form and content, this implies rejection of Husserl's «pure» phenomenology (located *above* every regional ontology and its associated material logic) as well as the functionalism that such purity underpins. Reduced to its core, the argument may be divided in its two parts, a) and b).

a) In the first volume of *Ideen* (1913), Husserl presents a paradigmatic example: the tree in the garden can burn, but the tree as such, the content of 'tree' cannot. Unfortunately, this remark, aimed at supporting the attainment, through *epoché*, of a domain of pure meanings, boils down to a fallacy. Were 'to burn' be relativized from real physical objects to their representations, the statement «The tree as such does not burn» would involve a categorial error, since, outside metaphor, burning can no more be ascribed *or denied* to mental contents, than colours can be ascribed/denied to numbers. The fact that the above statement is nevertheless intelligible (and not for merely syntactic reasons) means that for any representational content we separate, suspending it from its referential/indexical anchorage, some other representation has to remain anchored in order to endow judgments about the given suspended entity with sense. Revealing tests which support this point are provided by modal discourse, descriptions of fictitious entities, and counterfactuals. In these three cases, disactivation of some link between language and world requires that the activation of other links is preserved, as in the case of Archimedes Principle in hydrodynamics. In fact, my remarks can be summed up in a *Principle of Invariance for Referential Potential* (PIRP):

---

[14] For greater exactness it would be necessary to take into account different phases in Husserl's philosophical development. See Peruzzi (1988).

PIRP. For any sentence Φ, if some expression in Φ is lacking or intentionally deprived of reference, some other expression in Φ preserves or increases its role of referential anchor to concrete (perceptual) experience, in order that intelligibility of Φ be preserved.

If PIRP is correct, as the evidence provided by the tests mentioned above suggests, the noematic level aimed at by Husserl is not completely autonomous. Hence global *epoché* is impossible.

b) The chess example is canonical in demonstrating that one and the same form (program, system of rules, and so on) can be realized in a plurality of ways. Chess pieces can differ in shape, materials, dimension, colour, etc., and still the game remains what it is. This is usually taken as validating the inference: since for any given physical feature X of pieces (and board), X is not necessary to play chess, then no physical feature is necessary to play chess. But this is fallacious: there is no chess playing in an empty universe, or in one lacking stable particles. The inference confuses local and global suspension. Chess playing presupposes objective identifiability of pieces or their names (be they light spots on a computer screen, mental symbols, or anything else), an *extended* manifold (even discrete, as for the tape of a TM) sufficient to contain all the different positions (squares), and finally univocal occupation of any position. An immediate objection might be that these conditions are perfectly formalizable in *additional* 'background' rules for chess: after all, chess computers exist, without embodying any geometric program, and even configuration analysis, supposed to be the winning weapon of human players, might in principle be mastered in the next computer generation. Still the objection can be countered as follows: first, the existence of computers presupposes matter and thus space; second, computers need humans to interpret their ability as the running of a chess program. One could even add that present desktop computers are not playing chess, for the simple reason that they aren't *playing*. The ethological (and psychotherapeutical) aspects that define a game are lacking in any chess computer. Were the next generation of computers to manifest an autonomous game-playing behavior, one would have to say that the standard computational paradigm, underlying both cognitivism and AI, has been overcome. At this point, one might ask: is chess-playing realizable in any possible world sharing physical laws with the actual world? Or, does it require some specific *symbolic* feature realized in the actual world? In every case the separation of form and content cannot be so ultimate as functionalists suppose. Once more, the mere appeal to possible worlds is neither necessary nor sufficient.

Among the various lines of arguing which features of the Universe are required, the most interesting proposal consists in the Anthropic Principle, as elaborated by Barrow and

Tipler (1986).[15] Applied to the very foundations of semantics, this principle (in its *weak* form) offers a naturalistic support for the claim that there can only be local *epoché*. In effect, once we realize how fine the tuning of the fundamental forces of physics is, and the way in which it gives rise to the delicate balance that affects cosmological evolution at all orders of magnitude and levels of complexity, then the further step of employing the Anthropic Principle in semantics is at hand. If perception and cognition are natural features of a living system, and if the presence of life demands satisfaction of definite physical conditions, we should not expect to observe a universe much different from the one we are living in: this is the kind of world in which we *can* live. And more than that, this is the world in which language is possible: the cosmological window within which modalities themselves are intelligible is itself not modally indeterminate. The actual roots of reference impose strong constraints on possible worlds, and the cosmological window within which we exist constrains possible cognitive architectures. As Paul Davies admits, appeal to the Anthropic Principle is not an explanation of *why* the world has the properties it has, but provides a strong heuristic argument for determining in which way our own existence (and, I add, our own *semantic* ability) is a constraint on the range of the possible structure and evolution of the Universe up to now; conversely, mathematical models of the growth of complexity for physical structures suggest natural constraints on the conditions of possibility for meaning and the whole battery of intentional notions. If the Universe differed greatly from the way it is, epistemology and semantics would not be possible; and were the terrestrial environment radically different, they would undergo deep-seated alterations which a cosmological context for the study of the evolution of 'knowers' would be needed to investigate. But once knowledge and meaning do exist, and are as they are, not every categorization of phenomena (formalization of theories, etc.) is possible. It is in such scenario that, ultimately, the limitations of computability and provability find their rationale. All this applies, in particular, to the notion of the *Ego* (as a self-referential system): as a representation, should it not be subjected to *epoché* ? On the grounds of the weighted modularity of cognitive functions advocated in ch.1, the answer is that for any given module M_i, it is possible to suspend M_i - outputs, but this cannot be done for variable i, for all modules simultaneously. This serves to remind us that a completely modular model of the mind is neither an indefinitely flexible anarchy nor a rigid hierarchy, but rather a constrained eterarchy of modules that, at higher levels of integration, compensates the loss of rapidity resulting from openness to other modules by increasing parallel constraints.

If a) and b) point away from any idealistic hypostatization of a Subject, they also support a more concrete conception of transcendental investigation, as a *scientific* activity.

---

[15] For a clear updated survey of the different meanings of such a principle, see Balashov (1991).

(Of course, to the ears of Kantians and many analytic philosophers, this sounds like a *contradictio in adjecto*.) It follows that the self cannot be completely isolated from interactions with the actual physical environment (in which meaning is rooted), and the concept of the 'I' keeps track of such an interface with external systems.[16]

Suspension of specific semantic links affects *existence* as well as aspects of the *structure* of reality. If this lends support to the local character of *epoché*, at the same time a systematic application of the method of suspension discloses *other* aspects of our inner reality, one usually undetected. In particular, by concentrating reflection on the very activity of thought, in its fluency, we arrive at a consciousness of the possibility of not separating the world into marble-like individual things, and specifically we experience a chance of avoiding the cut between the self and the external world. Buddhist meditation, for example, leads to a suspension of the emotions while not leaving us estranged from the world; instead of making us perfect models of functionalistic theses, it enables us to reach a deeper participation in the world. (By the way, this does commit us to the adoption of Buddhist metaphisics.) On the contrary, Cartesian doubt is a version of global suspension and is for that very reason thereby misleading; for it is supposed i) to secure rational certainty as its residue, in contrast to the unreliability of the senses, ii) to achieve an absolutely autonomous ground of judgment. Local suspension, on the other hand, is not committed either to i) or to ii), and so it can avoid the recurrent tensions of modern epistemology in its desperate attempts to transfer the certainty apparently secured in the island of rational self-transparent thought, to the foundations of our knowledge of the external world. The last great attempt of this sort was Husserl's; it required the abandonment of the «natural attitude»: *epoché* had to be global. And this, besides not being compelling, violates PIRP, while a) and b) testify against such a violation.

However, in terms of the late Husserl, meaning is *constituted* through interactions within the «life-world». So PIRP corresponds to the fact that the wholeness of these interactions cannot be suspended.[17] And more than that, natural and phenomenological attitudes are mutually intertwined: the first does not need to be free from self-consciousness, the second does not need to be meta-natural. Whatever meaning is, conscious reflection on

---

[16] This has long-range effects even on ethics and philosophy of logic, where it stands in contrast to both subjectivism and conventionalism. A particular application of this argument is that elaborated in Peruzzi (1989a) in opposition to the 'free logic approach' to definite descriptions, and it is part of the evidence I mentioned under a) above.

[17] There is also another aspect of such a principle, one dual to that cited above. The idea of local *epoché* can be seen as linked to David Bohm's thesis of an *implicate order* (in the physical world) that cannot be made explicit *all at once*: there are several stages of 'involvement' and each time we unfold one of them, another becomes more involved. The connections between my local holism and the debate on local/non-local theories in relation to Bell's inequalities will be investigated in a forthcoming paper — the sequel of Peruzzi (1991).

it is part of a child's development and of common linguistic practice inasmuch as it is 'natural' for us to use self-reference in order to establish truths about the external world. Self-reference keeps track of intersubjective mediation (always internalizable) and is intrinsically bounded. So, there is a wide epistemic *braid* at work. Even in the case of deictic adverbs like 'here', understanding is possible because one is able to *commute* from another's location to one's own and vice versa. Analogously, phenomenological parentheses are opened and closed more often than usually thought; it is, however, a highly distributed and stratified dialectical process. At last, as the *bourgeois gentilhomme* has always spoken in prose without knowing it, we can become conscious that all of us are phenomenologists in training.

This point is relevant also to another Husserlian theme: what he calls the «pure, determinable X» as the referential hook «in abstraction from all predicates» (*Ideen I,* §131) is not so pure, for at the phenomenological level there are no such extrasortal entities X, as Gestalt psychologists realized in the domain of perception. Recently, Macnamara and Reyes (1990) have revived a similar approach in psycholinguistics, joining the framework of categorial semantics for type theories with Eleanor Rosch's theory of prototypes for basic-level kinds. Here mathematics meets phenomenology: any such X acts as a *pointer,* but the kind of entities referred to is not arbitrary: without a Background, no Figure, and without a Figure-maker, no terminating pointer procedure, see also Sajama, Kamppinen (1987) p.1O9. Moreover, Figure and Background have to possess the right relative dimension, just as the representational patterns have to satisfy various constraints, manifested by the child through the various stages of language learning. Once a sufficiently rich space of concepts, one basically anchored to actual topological experience, is established, it becomes possible for the child to operate on representations, inverting, in suitable contexts, even the roles of Figure and Background, or changing a variable into a constant and vice versa: the local character of suspension imposes that in the course of this process we cannot turn *every* variable into a constant or vice versa, and exactly this dialectics lends itself to description by means of the tools offered by category theory, through notions such as universal and adjoint, see Peruzzi (1991b).

In this same way we are able to imagine fictional objects such as winged horses, etc. This procedure is not just algebraic, as it might appear if we were to confine attention only to language. For instance, consider what happens in the case of the winged horse. Starting from our acquaintance with actual horses and birds, the dynamical 3D-schemes of a typical horse and a typical bird emerge; then we *separate* the two wings of the bird, *dilate* them in the right proportion to a horse's dimensions, and finally *paste* the wings on its flanks. The procedure is both algebraic and topological in character, and in particular the dilation is analogous to the functorial transformation of E after A into E after A. It is in consistently satisfying similar proportionalities, rather than in rules for modifying the logical behavior

of material conditional, that counterfactuals manifest their central philosophical importancer ; they are primarily related to transformations of local character within a topological setting, those necessary to recognize reference under any suitable combinatory manipulation of both Figure and Background.

§ 3.2. Skepticism naturalized

The naturalistic 'braid' is modally manifested in grounding the possible within the actual and locally identifying the actual through the possible. The form of this view outlined above was intended as appealing to more constraints than are present in fictional activity. For example, in the previous section we neglected the constraints raised by mechanics and evolutionary biology about the consistency of a winged ruminant: weight, morphology, opposite function of wing and hoof, etc. In imagination and fiction, we are involved in a combinatorial exercise; transferring the outcomes of this exercise back to reality, we might find ourselves involved in an unbounded or unconstrained sequence of semantic transformations, potentially changing the reference of any word: if the sequence is unbounded or unconstrained the transfer is either unintelligible or irrelevant. If this does not happen, it is because *epoché* is necessarily local: the macro-world we live in is structured in relatively autonomous regions and relatively stable layers of objects and relations.

In contrast with this view of the naturalized roots of the possible, Hume and later empiricists admit limits to the range of the possible only in the domain of logic and mathematics, supposed to be the realm of reason. In the Treatise Hume claims that necessity is something existing in the mind, not in objects. (Hence, for Hume, mind is foreign to nature, in apparent contradiction with his claim that there is no bridge between such necessity and the necessity whereby, he says, nature brings us to reasoning as to breathing!) It is as if knowledge could reach certainty on matters other than purely formal only if we have access to a globally transparent point of view, which we have not, and here stemmed Kant's project of internalising such point of view. As should now be clear, I entertain deep reservations regarding the separability of 'reason' from the structures of perception, and one can trust even more in direct sense experience than in deliverances of merely formal (logico-linguistic) rationality. (Would I *prove* to you that this page is red, would you believe it?) As is clear from both classical and quantum mechanics (minimum princples, virtual particles), the possible permeates our understanding of the real world, and this provides independent support for PIRP: the understanding of meaning is modal and cannot be separated from the range of its perceptual roots. Hence the entire debate about necessity and contingency need to be re-cast, and distanced from the legacy of Hume (and Kant). For

any given problem, we can *circumscribe*, within the bounds of skillfully examined experience, the kind and the range of externality demanded, which thereby comes to be localized. The truth of any proposition can thus be conceived not just in terms of internal validity, which the 'classical' epistemic approach to truth suggests as the optimum strategy to avoid the aporias of metaphysical realism. Here, the dialectical intertwining of local/global with internal/external aspects becomes once again apparent and deserves much more attention than it has receive until now.[18]

The point is that such static dualism between form and content has dogged the course of epistemology at its roots. Holism might at first seem to be as a way of overcoming such dualism. But it isn't. For it shares with Hume a basic tension between two features: the skeptic and the naturalist. They diverge, and hereby typical difficulties of empiricism stem. Quine renounces any form of localistic epistemology insofar as it aimed at transcending experience, while Putnam repudiates any form of naturalism as fallacious because he sees it as carrying to reductionist consequences. And yet both sorts of holism only threaten to substitute an even *deeper* skepticism, inherent in the *nature* of human knowledge.[19]

So, let us now consider some skeptical arguments, in order to test the import of local *epoché*.

Skepticism has been nourished by two sources: sensory illusions and paradoxical results of reason. Here I skip the discussion of paradoxes of reason (Zeno, Bayle, Kant, contemporary logic), since their phenomenological analysis would involve entering upon questions of the philosophy of mathematics inessential to the point I presently wish to stress. Attention will be restricted to sensory illusions, pausing at the issues raised by that form of skepticism related to the deceiving devil and the Dream.

A wide variety of skeptical arguments has appealed to subjective divergencies as to the sweet or bitter flavour of foods, as to non-idiosyncratic optical illusions: the stick that, immersed in water, bends, the pain felt in an amputed limb, up to the more sophisticated experimental artifices of contemporary psychology. In all these cases the skeptic emphasizes facts showing the possibility of an irreducible lack of agreement among mental modules, and an irriducible gap between appearence and reality. On the weighted modularity view proposed here, there would be even more occasions of (local) skepticism, but such fracturing dichotomies as that between Sense and Reason, or Mind and External Reality are no longer possible. Furthermore, if we accept a naturalistic stance in

---

[18] On the logical side, a contribution to provide a suitable framework for the analysis of such immanent/transcendent duality is given by the ILEG project — see Peruzzi (1991).

[19] Humean naturalism has been correctly linked to Wittgenstein's appeal to our «forms of life» in explaining epistemic and semantic questions, cf. Strawson (1985). In fact, the nature of such forms of life is culturally determined. So this naturalistic foundation is at most a variant of *pragmatism*, relevant reservations about which are entered in § 3.6, where a more coherent naturalism is opposed to it.

epistemology, skepticism turns out to be a *theory* about nature, a theory that collides with the mosaic of theories that compose contemporary evolutionary cosmology.

Let us go back to one of Hume's examples, one of enormous importance because it has been advanced again and again for more than two centuries in several variants: we do not perceive the real table, since its size remains constant as we move away from it, whereas we see the table diminishing in size. So, what we see is the *image* of the table (the table-phenomenon), not the real table. On the eterarchical view of mind, such argument would at most provide an example of how the multiple sources of information do not necessarily agree with each other. Taken as ground for a general epistemological thesis, it is easily contested. T. Reid already realized that this argument is a *non sequitur*. Hume's (incorrect) argument can be confronted with the 'constancy of size', which Gestalt psychology has widely investigated. The persistence of the argument can perhaps be interpreted in the light of the widespread tendency of philosophers to jettison the «natural attitude» of common sense. In fact, the argument tacitly leans on that same demarcation line between inscrutable physical objects and epistemic representations that it sets out to demonstrate.

The above discussion of local *epoché* shows that skeptical hypotheses are intelligible only as long as they persist in keeping ontologically fixed some referential component of what they are about. It is just as a dualized form of hyper-rationalism that skepticism boils down to the presumption that global suspension, *i.e.* complete bracketing of actual reference and judgment, is possible. Were the skeptic to deny that, he/she should subscribe to some variant of the coherence theory of truth, one with a holistic and pragmatist flavour; but we have already encountered the aporias stemming from holism and will encounter those related to pragmatism in the following. Hence, is it really possible to *describe* a perception without any commitment to the correctness (under any respect) of other perceptions? On the view being advanced here, the answer is negative. This answer is independent from the thesis that perception (and cognitive experience in general) is theory-laden, but rather turns on the claim that the structures of intentionality are *existence-laden* : it appeals to the *real* presuppositions of the possibility that something like a table-perceiver exists. Herein lies the position I term *entwined naturalism* :

THE STRUCTURE OF MIND PROVIDES THE POSSIBILITY OF EXPERIENCE OF REALITY, THE STRUCTURE OF REALITY PROVIDES THE POSSIBILITY OF EXISTENCE OF THE MIND.

In order to show how such a braid, together with the local character of *epoché*, act in relation to skeptical test cases, let us consider now Descartes' argument based on the possibility of a deceiving devil.

In this argument, the hypothesis of an actual knowledge of the realm of mathematical truth is contrasted with the hypothesis of an entity of infinite power , since it has to exceed

any arbitrarily large multiplicity of information we could reach for checking whether proofs are correct, etc. A fortiori, the devil can deceive our senses, so he can undermine our trust in deduction as well as induction (though Descartes avoids appealing to the devil in the case of the external world). On the ground of PIRP, either the description of such meta-rational devil isn't really intelligible, or it is simply false. The objection generated by braided naturalism to Descartes' deceiving devil is analogous to that against Maxwell's devil. These two devils violate conservation principles of a semantic and physical character, respectively. In the case of the evil genius we appeal to a particular instance of PIRP, rather than to information theory.

But, isn't this interpretation misleading? The point is not that did the deceiving devil exist, we could not know it, but rather that we could not *understand* its action, since, in order to understand it, we would have to possess a far larger amount of *knowledge* than we really have — a strange presupposition on the skeptic's part! And since *everything* would be thereby contaminated, even the words used to express the argument would be deprived of their intended meaning. So, to reduce the devil's hypothesis to absurdity, we can adapt the parable devised by Putnam (1981) against the hypothesis that we could be brains in a vat, kept alive by some galactic superscientist.[20] In practice, such a hypothesis has absolutely no basis within contemporary cosmology.[21]

A depersonalized variant of the argument in question is the hypothesis of experience as a Dream. In *Noema*, I adapted Wittgenstein's strategy in relation to 'doubt': the term 'dream' would no longer retain its usual meaning, and nothing prevents from extending this change to any other expression. Once the meanings of all our words are changed, 'dream', like 'deceive', is not immune from meaning change. So the argument becomes empty. But then it cannot be claimed that we could not realize we have reached truth, even if we really have reached it. This is not so much because a more general capacity for discerning the gap between actually reached stage and goal (in any matter whatsoever, from problem solving to athletic competition), would disappear by this admission; it is rather because either truth would be made independent from the way of arriving at it, or the external would be rendered unattainable via the internal. What Minsky calls «difference motors» may disfunction, of course, but truth is on the same footing as our daily negotiation of street traffic jungle without accidents, and since truth is a notion pervading human experience, to call into question our mastery of truth assignments is a self-paralizing move. If *some* difference motors are suspended, some *other* other component of our

---

[20] The parable must be treated with caution, since it doesn't bring to the conclusion Putnam derives from it. I will not enter here into the question whether this sort of arguments is indeed consistent with "internal" realism.

[21] Treat this observation as a homeopatic dose of coherentism, to be taken at this point in order to avoid a long *detour* in the development of the argument.

cognitive architecture has to remain active in order to understand the very hypothesis of an untransparent truth or a hidden disfunction.[22] So, the local character of suspension proves to be a *universal* tool for dissolving the examples adduced by skeptics concerning the nature of the interface between mind and external world.

§ 3.3. Intentionality

Briefly stated, the problem of intentionality is: how can thoughts refer to something? The discussion of this classical problem has recently received a strong impetus through the debate originated by Searle (1980). He devised a by now famous example, the «Chinese Room», to refute the project of 'strong' AI, consisting of the thesis that a suitably programmed computer is a mind, in the sense that it *understands* (and has other cognitive states); Searle admits only 'weak' AI, according to which intelligent programs are useful tools for testing theories on the functioning of the human mind, by simulation of some of its performances.

There is not enough space here for an exhaustive survey of the wide discussion of this subject that took place in the last decade, especially in the forum provided by the journal *Brain and Behavioral Sciences*. This and the following section will exploit only some aspects of that discussion, allowing us to show the relationship of braided naturalism with other widespread views of intentionality.

Strong AI is strictly associated with functionalism, since it involves the same dualism between form and content, as based on the silicon proportion mind:brain = software:hardware. Accordingly, in order to possess intelligence, it suffices to have the right programs. What matters is a suitable manipulation of symbols, however complicated it may be, the nature of their physical support being irrelevant. The Chinese Room is designed to show that mere manipulation of symbols lack 'aboutness', whereas mind is in fact more than formal; it has a semantics. Therefore, as any computer has only syntax, no computer is a mind. Understanding of language requires mental states, which in turn require meaning (reference), and thus no computer can *understand* a language — at least as long as the architecture of computer programs is designed and developed in the way of classical AI.

---

[22] As to the absoluteness of the Cartesian Doubt, Peirce reminds us that the search for a foundation, from which to recover our usual certainties, is «as useless a preliminary as going to the North Pole would be in order to get to Constantinopole by coming down upon a meridian» (*Collected Papers,* 5.265).

The Chinese Argument argument has attracted many comments and objections. The most relevant can be summed up in the following items: 1) It is necessary to look at some larger system, e.g. the whole 'Room'-system, in order to ascribe language understanding. 2) The computing device would have to be equipped with a robot (a 'body'), allowed to freely interacti with the world. 3) In order to achieve reference, the system must possess causal connections with the environment. 4) The Chinese Room argument passes over many orders of system complexity: it is the hierarchical architecture of the system that allows reference to emerge.

Searle developed various responses to these objections, whose details I shall omit here. Anyway, objections 1, 2, and 3 have already been discussed in the previous sections, in a way that (partially) meets Searle's reply to them . The point of his arguments is that formal procedures are intrinsically unable to make the leap from syntax to semantics, while generic causal connections are unexplanatory. Mental states are biological phenomena, and intentionality has to be explained through (and only through) the causal powers characteristic of that biological system which is the human brain.

In this manner what is undermined is a central claim of cognitive science, according to which intelligence results from a hierarchically organized system of formal rules for processing information, as exemplified by generative grammars with respect to language, by strategies for face recognition with respect to vision, etc. When humans follow a rule, they are guided by its formal aspect but also by its *content,* while computers do not know that in calculating 6+6, '6' stands for the *number* 6 and '+' stands for *addition*. *A fortiori,* we must keep distinct computer simulations of complicated capacities of mind from these same, strictly biological, capacities .

A different approach to intentionality has been proposed by Daniel Dennett: the attribution of the predicate 'intentional' to a system S does not describe some intrinsic property of S; rather it has to be relativized to the needs of explaining and foreseeing the behavior of S. So, the intentionality of S is a function of properties of the system S' reporting on S, as in the case of the consistency for formal systems: it is a *relative* phenomenon. The descriptive utility afforded by intentional vocabulary (reference, belief, wish, etc.) could be exploited by computers too, but intentional folk-psychology could turn out to be *empty* at the end, cf. Dennett (1971).

However, residual doubts remain as to whether we should consequently adopt an *als-ob* philosophy of mind, or we should interpret the situation as similar to the measurement problem in quantum mechanics. One could say that self-reference (S = S') implies intentionality *absolutely*, unless the attribution of intentionality to a system is just a fictionalistic label. Besides, in view of the pervasive presence of intentional discourse, what could be guaranteed not to be a fiction? In any case, the question can neither be decided at a desk, nor by leaving implicit those hypotheses regarding the systems' environment E(S,S')

that specify the causal interactions of S and S'. But then the door is open to an ecological semantics: meaning comes from outside, from the external background, but only in definite ways; every time it is brought inside (the given cognitive system) and considered in isolation from its genesis, it becomes a play-meaning. Yet, our previous explanation of the braid linking cognition to the physical environment makes it clear that such an ecologically anchored view captures only one half of the truth.

By way of contrast, if the semantically relevant layer is just the internally projected world (as the soft Neo-Kantianism of Putnam and Jackendoff suggests), then meaning is reduced to a simply wider play-meaning — the option between a more holistic or a more constructivistic perspective being secondary.

Aside from the old-fashioned insistence on the autonomy of epistemology from psychology, a common reply is that we can always make this 'externality' more remote, by progressively enlarging the system's range. At the extreme, the entire Universe might be counted as *the* system (we could call 'the Cosmic Room'). The trouble is that we could no longer appeal to anything external in order to inject meaning into the formal wheels of the mega-system. As a corollary, it would be impossible to confer 'meaning' on the Universe itself, unless meaning is already intrinsic in the net of representational structures that categorize possible experience. Either such an inflationary structuralism, be it produced by transcendental wheels or not, goes hand in hand with convictions as to the absurdity of life (once one maintains the externality of meaning in the absence of God), or the System itself is the ultimate source of meaning. Thus, real = internally justifiable = rational. So, if you inject into the System the dynamic character of the above enlargement and the holistic nature of the resulting changes, you have a new route from Kant to Hegel.

Let us suspend judgment on this aspect of the problem for a while, and come back to a more parochial answer: *we* provide semantics for any computer, as well as for our own symbols. Of course, such an answer is less parochial if some form of idealism is assumed. So, one could say that it is the Subject who creates meaning and confers it on the Object — moreover, both activities conform to internal rules: say, logico-mathematical principles. Unfortunately, this is far from being convincing. It is neither a question of referential causal chains nor a question of protoplasmic powers of the brain, inherently designed to secrete meaning, for meaning is not created at all: any master *homunculus* (as the internalization of the Subject) labelling symbols with their meaning stands in need of explanation as much as such a mysterious quality of the grey substance. Rather, we have to understand how intentionality *emerge*s from the interaction of cognitive subsystems with the environment, and thus with each other. Their physical support cannot be completely separated from functional architecture, and vice versa. Indeed, any system endowed with semantic access is a highly specific physical system. Angels lacj intentionality since they lack any interface between function and substance: they thereby resemble computers more than humans. At

the same time, nothing prevents intentional capacities from being multiply realized provided the physical environment is sufficiently rich in structure.

If there is more than syntax in the mind, our internalization of the world by means of formal systems might well turn out to be incomplete, since it embodies some *trace* of the existence of relations irreducible to manipulation of discrete symbols: true, any given semantic relation could be internalized within a higher level syntax, but then some other semantic relation has to remain unformalized (this is in perfect concordance with the thesis of local *epoché* and PIRP). Anything can be a symbol, as long as something is not a symbol. Thus, we cannot say once and for all where the limits of formalization lie. So, the range of what is expressible in language and the range of what is not, are *inseparable* from each other.

In this way, we have the tools to assess the Cosmic Room argument. That argument is unacceptable on phenomenological grounds: internal reference is not reference at all, internal truth is not truth at all. They are just secondary byproducts. But then we have to be careful to keep the internal distinct from the local, whereas previous debates on intentionality neglect the central role of the dialectical pair local/global: global inseparability between sets is, in terms of category theory, only a particular form of intensionality, see Peruzzi (1988a), and this property is a manifestation (on the verbal level) of intentionality; thus, if the entire argument above can be codified in S (and for S = the human mind, such a condition is satisfied), then S reveals itself to be *absolutely* intentional. Therefore, neither may the *als-ob* way out be preserved any more.

By considerations related to objections 1, 2 and 3, we find both Searle's solution and Dennett's way out unsatisfactory. The fourth objection still remains to be discussed.

§ 3.4. Emergence out of the Boolean Dream

One of the most interesting replies to the challenge issued by the Chinese Room is that semantics is a function of the complexity of a system, rather than of its causal links to the physical environment. In a naturalistic perspective, the role of complexity is especially relevant when we come to consider our understanding of self-reference, but even if complexity doesn't suffice, the emphasis on it converges with the recognition of its genetic links to ecological properties. That is, A) complexity is possible only under conditions of the stability of matter, and specifically under well-determined physico-chemical constraints; B) the richness of the internal landscape, sufficient to let meaning emerge out of a «bubbling broth» of interactions, has to be *smoothly* linked to the richness of the external

landscape. Exactly this smoothness is missing in the Chinese Room case, but as far as the consideration of A) is missing, even objection 4 is inconclusive.

The innumerable parallel processes of millisecond scale, out of which intelligence emerges, are inaccessible to introspection: they are encapsulated in Fodor's sense, but in contrast with Fodor, it is possible to conjecture that precisely in this region of opacity the interface between syntactic and semantic categories occurs. This interface permits the emergence of concepts and the capacity for creative behavior, as resulting from an interconnected architecture, whether neurophysiological or artificial. According to Hofstadter (1985) p.528, the basic character required of something in order for it to be a concept is «the way it is integrated into memory [...] the property of being a concept is a property of connectivity». Thus it involves both algebraic properties concerning the way the given structure relates one component to another, and topological features of a global character, to be ascribed to the underlying network of computational spaghetti.

From rigidly programmed behavior to higher and higher flexibility in performance, there is a continuum along which *sensitivity to patterns* grows, as the distinguishing property of creativity, 'concept-hood' and consciousness itself, that is, «an ability to spot patterns of unanticipated types in unanticipated places at unanticipated times in unanticipated media" (*ib*. p.531). Yet here we find not a plain continuum, but rather a *punctuated* one, since the grid into which pattern recognition distributes data is not a homogeneous space: there are singularities, structuring the syntax-semantics interface above, and these singularities keep track of an enormous amount of evolution *common* to mind and its environment. In order to make this point more perspicuous, and more firmly linked with the previous appeal to Gestalt constraints on cognition, let us focus attention on the example of letters and fonts.

The apparent simplicity of the capacity to recognize letterforms is far from being easily simulated. The open manifold of fonts in Western, Chinese, Arabic alphabets, provides an argument that expert systems in specialized domains can turn out to be an obstacle, instead of an aid, to more penetrating programs in the simulation of our typical ability in pattern recognition. Variation in shape of one and the same letter in different typefaces is sufficient to raise the general *analogy problem* in the most clearcut way.[23]

---

[23] As Hofstadter (1985) p. 633 suggests, the problem of recognizing letterforms can be taken as really paradigmatic:

> The problem of intelligence, as I see it, is to understand the fluid nature of mental categories, to understand the invariant cores of percepts such as your mother's face, to understand the strangely flexible yet strong boundaries of concepts such as "chair" or the letter 'a'. Years ago, long before computers, Wittgenstein had already recognized the centrality of such questions, in his celebrated discussion of the nonpindownability of the meaning of the word "game". To emphasize this and make the point as starkly as I can, I hereby make the following claim:

Optical reading machines for blind people already exist with satisfying results, but their successful behavior depends on a strong reduction in complexity with respect to the general problem: they are only able to determine which, out of a *fixed* storage of characters, any given letter-instance most resembles. The performances of OCR software are by now accessible to a wide audience through the diffusion of scanners, but they suffer from the same limitation (actually, a virtue too) of expert systems in language understanding. It is not through a similar comparison (though one of larger scope) that we are typically able to identify a particular individual or to classify it as a member of a kind. Once again, possible quotients are so numerous that only ecologically anchored constraints can give account of such abilities in humans. This remark is also related to the importance of counterfactuals in cognition, as they are involved in filling sentences like «*A* is to *B* , as A is to ...». One could speak with Husserl of *eidetic variation* as the central theme of such a renewed phenomenological analysis. It is not a hyper-complex language or logic that is required to achieve 'good' simulations of common sense. Rather, it is through parallel and analogical architectures attuned to natural wholes, that we can fill the gap between perception and cognition. The same tuning points to giving up the functionalist dualism of form and content, thus also undermining the IPM in psychology. Understanding is not produced by manipulation of formal tokens. The difference from Searle lies in that formal tokens are viewd by him and most of his opponents as 'passive', whereas we have to focus on *active symbols*, as suggestively described in the «Ant Fugue» by Hofstadter (1979). So, the present point is not simply that concepts can emerge holistically at a suitable global level while being absent at lower ones, but that symbols are not not just *passive* objects manipulated by some overlying program. Reference and other intentional notions emerge at the global, collective, level through active and self-stabilizing symbols; the property of being a symbol is not *designed* as computational, even though it can be seen as emerging out of a computational 'bubbling broth'. The reasons of such activity and self-stabilization have to be related to A) and B) above.

We arrived at this position by facing the problem of understanding how our capacity of pattern recognition is possible in the case of letterforms, but this same capacity is manifested in the use of common nouns in ordinary language, corresponding to the *fluidity* of concept representation.

Another paradigmatic example is presented by Hofstadter (1985) p. 643:

The central problem of AI is: *What are 'a' and 'i'?*

> By making these claims, I am suggesting that, for any program to handle letterforms with the flexibility the human beings do, it would have to possess full-scale general intelligence.

> Does the fact that taxis stop at red lights mean that traffic jams stop at red lights? One should not confuse the properties of objects with the properties of statistical ensembles of those objects. In this analogy, traffic jams play the role of thoughts and taxis play the role of neurons or neuron-firings. It is not meant to be a deep analogy, only one that emphasizes that what you see at the top level need not have anything to do with the underlying swarm of activities bringing it into existence. In particular, *something can be computational at one level, but not at another one.*

Here too the local/global duality is not absolute and does not concern only one level, because the road network of a city can be internalized as a local object for town-planning; moreover, the example shows that it is necessary to complete this duality with the internal/external one: while the driver considers as traffic what he/she encounters locally (say, between two consecutive traffic lights), the external observer looks at the simultaneous traffic patterns of the whole city. Analogously, the way out consisting of the mere relativization of 'computational' is unsatisfactory: I have already argued against a similar move for 'true', 'real' and 'intentional'. The transition from passive to active symbols has to be amalgamated with the two dualities above into a naturalistic dialectics, whose theoretical framework is already at hand in the form of category theory — in particular I exploit here Lawvere's view of the multiple adjunction between discretes, points and chaotics, cf. Lawvere (1989).

What differentiates the appeal to active symbols from the IPM and most of current AI is that, in order to flow, they do not need any central program to manipulate them, since, once the proper level of complexity is reached, symbols behave as self-stabilizing. In conformity with the view proposed in ch.1, there is no longer any need for a Central Manipulator, if we recognize an autonomous dynamics of symbols. Instead of elaborating more and more complicated projects for simulating competence in idealized domains (*i.e.*, trying to grow computational flowers without natural stems) conceptual fluidity suggests we look rather to the parallel activity of analogical character that permeates the brain wetware.

The idea is that, once a certain complexity threshold for parallel, mutually cooperating, units (in analogical code) has been crossed, the output of such activity can be sequentially mastered (in digital code), and hence meaning can emerge, and actually does emerge provided that the system is able to interact autonomously with the physical objects that are the targets of its symbols, and at the same time the continuous source of suitable feedbacks. The Boolean dream of finding the «laws of thought» in purely formal rules for manipulating propositional/class symbols has to be abandoned: the essence of cognition is missed if it is identified with its algebraic and computational end-products only. Neither does it suffice to appeal to causal interactions (via sensory connectors) with the outside world, since *any* actual hook can be suspended (by local *epoché*) without thereby losing the meaning of symbols; nor would emergence and stability of the semantic system be

simply products of mere statistical assemblages. Space-like constraints are essential to the dynamics of such growth in complexity: *invariance*, as described by algebraic topology, and *stability*, as described by differential topology, here meet and reinforce each other. Meanings become associated with stable paths in a generalized space where the matching of language and perception takes place. Here, by the intrinsic presence of attractors we can account for the fact that even in fiction and counterfactuals meanings remain active and stable and can thus be referentially anchored to some 'world-slice' model M , not necessarily actual, but internally coherent (in such a way as to support mental imagery), given *actual* boundary conditions.[24]

Thus complex symbols are more than epiphenomenal over *both* computational Brownian motion *and* imprinting phenomenological gestalts: they are built into the windows among different modules. The interference of many threshold effects accounts for undecidable, or impredictable, properties of thought, and hence for its global non-programmable character. The import of this is far from blocking AI projects or undervaluing the opportunities offered by new mathematical tools. Just the opposite. New research can be motivated by the problem of explaining semiogony, once we are no longer content with the simple appeal to «causal powers» of the mind-brain, with the classical IPM and with such a prodigiously stable statistical emergence of concepts as connectionism claim to provide. If standard functionalist approaches to mind appear to revive mediaeval theological themes of the hierarchy of angelic intelligences (a bodyless society of free-floating programs), on the other hand the exclusive anchorage of intelligence to causal interactions with a fixed environment would justify the charge of 'hardware chauvinism'. Fortunately, this is not an *aut-aut*, once topological constraints are taken into account, recognizing the prime importance of algebraic and differential topology, as intrinsic to basic mental constructions rooted in features of the macro-physical world.

Up to now, only a few of software-hardware-environment pairings have begun to be investigated by AI, and all of them are of low phenomenological complexity. If the question becomes, where does intentionality enter such pairings, there are two lines of research in prospects: one more theoretical, one more experimental, but the philosophical moral of the above argument is that, as far as the structure of cognition is concerned, the two lines cannot be mutually extraneous. We have to move back and forth between them in order to see the mental slippability of active symbols as linked with aboutness, showing how

---

[24] These 'world-slice' models can be variously termed: situations, possible worlds, contexts, frames, etc. They are made possible by local *epoché,* individuate cognitively natural chunks of referential activity and are the bearers of mental manipulations, by varying in imagination some of their combinatorial features, or filling some of their slots in different ways from what actually occurs. Yet their meaningfulness is due to features locally held rigid.

it is possible for semantic fluidity to emerge through a complicated eterarchy of features, suitably stratified. Topological constraints are primary in identifying the attractors of the subcognitive dynamics at the level of prototype perception, and this is the level at which concept-hood and intentionality first manifest themselves to awareness. The unity underlying this two-sided approach can be grasped in the framework provided by notions of category theory, because of the central place occupied by adjoints in the dialectical interface of mind and world.[25]

At this point one may ask: how can such an approach explain that our capacity for self-reference does not generate an unlimited hierarchy of internal (and external) observers, each codifying the preceding ones?

Actually, my initial interest in this problem stems not from AI or the philosophy of mind but from the need to avoid the traditional regress in the foundations of mathematics: from any theory *T* to its meta-theory *MT,* then to *MMT* and so on. In *Noema* I remarked that the problem of understanding a system S, be it formal or cognitive, is *one* problem in two forms, thus one and the same phenomenological approach should apply in both cases. The idea was, first, that the 'standard of measure' for understanding any concept is not a formal system, but rather the genetic operations (constructions) through which the given concept emerges (is abstracted) from dynamical interactions with macro-objects in physical space; and second, that local *epoché* provides the needed flexibility for manipulating concepts (and formal systems).

Instead, traditionally, one takes the object-system *S*, associating it with the ordinal *O*, and goes on to the *n*-th meta-system; then, one passes to the □□limit and so forth: $S_O,...,S_n,...,S_{\square},...$. Suitable ordinals provide the opportunity to prove consistency for lower-stage systems. Analogously, in order to be sure that a given program *P* is correct, one needs another program *P'* examining *P*. And the situation recurs, so one would need a *P"*, and so on. In some cases this strategy is profitable, but in general the testing meta-program cannot terminate, since the stronger a program is the less easy it is to control. It can be argued in parallel with these considerations that self-consciousness cannot be *reduced*, the activity of the transcendental subject cannot be objectified, etc. If the mind/software analogy is accepted, this is just a version of the general unsolvability of the halting problem, related in its turn to the failure of Hilbert's finitistic meta-program. If, at certain ordinals □, with □<□, we can represent $S_{\square}$ into $S_{\square}$, we have self-reference for $S_{\square}$ and thus, if not the liar paradox, at least incompleteness: a warrant that the certainty searched for cannot be warranted by progressively shifting from weaker to stronger systems.

---

[25] A precise symptom of the unifying power of categorial notions in the present context is offered by the meaning that the notions of completeness and incompleteness of formal systems receive in terms of adjoints, cf. Peruzzi (1988a) and Pavlovic (1989).

At this juncture, the above consideration of active symbols and the sheaf-model of the mind can be jointly exploited. In fact I doubt that conservative, or inert, self-reference (*i.e.*, leaving the original system, in its original language, unmodified) is the faithful *explanans* of self-reference. In a 'sufficiently powerful' living system *S*, there is dynamic interference between *S* and *MS*. If we confine attention to consistency and truth-in-a-model for the usual, static, formal systems and models, Gödel's impasse is insuperable. But this is reminiscent of the myth of an all-embracing intelligence which remains a pure Eye — thus, after all, a limited form of intelligence. Still, that human intelligence is irriducible to such an Eye does not tell what the dynamic diagonalization characterizing us consists of. For instance, perturbations of attention during self-observation are sometimes not strong enough to affect any activity, and sometimes affect the simplest task. (Recall the story of the millipede who was paralized when asked how walked.) Indeed, Jean Piaget and Seymour Papert already faced the question of the effects of the auto-epistemology adopted by the child: what do children think of *learning*, why and how far does their auto-epistemology influence their performance? A child who considers him/herself scarcely intelligent will fail in learning certain tasks, while a know-all will tend to interpret any problem of life in a bookish style.Conversely, how is it that in a system *S* dealing with a domain *D*, once *S* is enriched with global self-reference, it can lose its hold on *D* ? Should the working mathematician be conscious of the whole range of mental acts performed in problem solving, mathematics would not exist. Only when foundations become a *local* subject they can be dealt with mathematically. Another example of the same situation, with respect to ethics, is examined by Preti (1986).

Moreover, the theory/meta-theory interference affects not only *T* but also *MT*, and in definite ways, suggesting that the adequacy of any cognitively relevant model of self-reference is of strictly local range: its adequacy grows in direct proportion to the narrowness of the domain of the self-represented theory, since it is better isolated from perturbations deriving from other domains accessible to the representing system.[26]

We can by now recognize that i) in order to formalize these various cases of dialectics, ordinary logical/set-theoretical approaches to self-reference are inadequate, ii) the tools provided by category theory improve the formalization substantially since they give us a general setting to deal with localization of meaning and internalization of truth, iii) by means of such tools, the sheaf model makes room for localized self-reference. We lack only the further step needed to relate i)-iii) to the emergence of symbolic *activity*.

---

[26] This remark would need many qualifications, for cosmology is an example of a science with the largest possible domain, one that in principle embodies a description of the mind that conceives cosmological theories.

Gödel's Theorem may be interpreted as saying that the complete fusion of an observing and an observed system is impossible (starting from a system in which observation boils down to a recursive encoding), but the point is that the systems cognitive science is interested in are not simply observers (codifiers), but self-modifying *agents*. So, in order to reach an empirically relevant representation of the resulting dialectics, one has to consider features of what is *real* self-reference for systems interacting with their environment in various ways. This consideration, too, can profit from the use of *adjoints*, through which we can begin to see the outlines of a dynamic formalization of the active interference between theory and meta-theory. In Peruzzi (1989) a similar dialectics is presented in relation to the interaction of descriptive and constructive aspects in the phenomenology of logic. In this context, the presence of a suitable adjunction can be taken as the indication that symbos have become *active*. Gödel's Theorem finally lends itself to interpretation as a particular case (for static systems) of a more general *duality*, one capturable in terms of adjoints — cf. footnote 24. I am perfectly conscious that much work is needed to achieve solid conclusions on these topics. It is nonetheless important to focus attention on a line of research that has been neglected for too long, namely a view of semantics and epistemology as branches of a natural philosophy, to be developed not only or mainly in terms of linguistic analysis, but by means of the best resources provided by contemporary mathematics.

§ 3.5. The Historicist Exit

Putnam's reading of Gödel's Theorem points to a historicist interpretation of rationality, and my appeal to a dialectical conception of semantics and epistemology might seem to confirm such a line. Yet, this is not at all the view I maintain. Thus the reasons for repudiating the historicist way out of the danger of 'instant' rationality must be pointed out, making clear that the related pragmatist conception of meaning and truth can also be avoided. Finally, having contrasted Searle's yapproach to intentionality with my own position, I have to show how the divergence from historicism and pragmatism forms part of a naturalistic perspective. This and the following section will be devoted to these points.

Only *some* aspects of historicism and pragmatism will be considered here. My claim is that these come close to a naturalistic conception of semantics and epistemology, but, for different reasons, do not succeed in reaching the goal of providing a foundation or a liberation from any such foundational project. (From the very beginning, the aim of the present essay has not been completeness or formal exactness, but an analysis intended to break up the existing philosophical subsoil and to free the ground from the encumbrance

of archaic, though subtelty mantled, positions, and thus to make possible the growth of a fresh theory that fills the gap between the domain of 'pure' mathematics and of 'pure' epistemology.)

Two forms of historicism can be distinguished, say a *soft* and a *hard* one. Cultural relativism and Hegel's absolute idealism represent their best known manifestations, respectively.[27] Historicism in both forms is seeking an overarching philosophy in the wrong direction, for it leads either to skepticism or to the genetic fallacy. There are other ways of approaching philosophical problems without such consequences. This conclusion, for which I am now going to argue would not deprive us of the richness of insights stemming from some historicist attempts. Such richness has left its trace in the dynamic naturalism I advocate in opposition to historicism.

According to soft historicists, any conceptual framework has to be seen as immersed in the stream of history, which, being *ex hypothesi* neither guided by external forces that determine its course, nor endowed with autonomous laws, is a *contingent* result of the becoming of social practices; these are organized in wholes that undergo unpredictable changes. Since language is just one of such practices, meaning and truth do not transcend each given culture — say, during a certain period of time of sufficient semantic stability (contemporary philosophy of science is replete with variants of this thesis).

The hard version of historicism claims that appeal to history is a necessary step to the completion of semantics and epistemology: the laws of history represent an *intrinsic* guide for understanding cultural changes, that affect not only particular technologies and institutions, but even categories of thought.

In its hard form, historicism claims that there exist laws of history, *i.e.*, regularities in the variations to which cultural events (though apparently unique phenomena) obey. Therefore, the variations involved are, in effect, not 'free', and the uniqueness of historical facts does not preclude lawlike behavior of collective phenomena, at the 'right' level of analysis, which differs, however, from the level of naturalistic investigation. Subjectivity is not excluded from the range of these laws, rather it becomes an (usually unaware) instrument of an inbuilt historical dynamics — the Hegelian theme of the «cunning of reason» has undergone innumerable variations in contemporary philosophy of the social sciences, even within methodological traditions far removed from any direct link to Hegel.

Trouble arises when historicization applies to science itself, and granted that historiography is in principle a science, to historical knowledge of historical laws concerning this same knowledge. For, instead of regarding such a circularity as indicating the need to refine our awareness of the multiplicity of natural roots underlying axiological

---

[27] Actually, several other variants have spread in XXth century but their examination is not deserved by my present aims.

notions, and of their subtle dynamics, very often we face a leap in the dark, prepared by a sequel of aporias. In the heyday of the logical-empiricist methodology of science, these could be thought to concern only historical knowledge (and the way it enters humanities), but later, through the impulse of the contributions by Kuhn *et al.*, they have come to be spread all over (the image of) science and transformed from vicious circles into virtuous matters of fact.

It seems that we are confronted here with a problem: how might (hard) historicism be refused, after having proposed the embedding of epistemology itself into the largest possible history, that provided by cosmology?

The arguments traced so far in this essay suggests a twofold answer: i) the usual historicist view of scientific knowledge involves an uncoupling between the 'human' world and nature, that a naturalized phenomenology avoids; ii) if there are laws governing cultural phenomena, they are, rather than historical laws, laws associated to the emergence of structural complexity in any system, and so are not specific to human cultures. The conjunction of i) and ii) permits a peculiar exploitation of Popper's criticism of hard historicist views, as they reveal themselves in the holistic doctrines of contemporary social sciences, without at the same time closing the door to any rational theorization of the dynamics which self-referring systems exemplify.

One might object that, even thus conceived, regularities of socio-historical phenomena would preserve the *inevitability* feature characteristic of natural regularities. This objection, however, ignores the myriads of threshold effects diffused in social phenomena just as much as in nature: this affects prediction in both cases. (A related objection diagnoses a flaw in the argument akin to that involved in the naturalistic fallacy: the collapse of the possibility of evaluation attendant on the naturalistic fallacy would be a corollary of hard historicism too; this objection has already been touched on in relation to Putnam's anti-reductionism. It will be discussed again later.)

In this regard, Popper (1957) particularly insisted on the Oedipus Effect, *i.e.,* the influence of prediction on the realization of the predicted event: many typical phenomena of individual and social psychology testify to the pervasiveness of the effect, for instance Stock Exchange fluctuations. For the same reason, predictions are, where possible at all, very arduous in the social sciences. The interaction between observer and observed (in which self-reference is involved) may produce changes in any course of action, making this appear indeterministic, whereas it is by now clear that many physical systems are inpredictable, albeit deterministic.

In contrast with the holistic conception of society as an undivided whole and as the minimal stable overall context for describing and explaining human behavior, Popper's proposal of a «piecemeal technology» applied to society offers another example of a *modular* view: that is consistent with his «bold conjectures», intended to maximize the risk

of falsification, since otherwise there would be no identifiable element to be declared falsified. Certainly, the holist will contest this point, but we have seen in ch.2 that such an objection is flawed. In the same manner we can also make an effective criticism of the globally holistic character of historicism, implying that society is a unique whole (that only global revolutions can change).

From this modular 'engineering' perspective, one of the objectives is to make clear what cannot occur *under given conditions*. In this sense there exist laws of economic behavior, etc. (e.g. it is excluded that a centrally planned economy realizes a pricing system adequate to the role played by market prices, as it is excluded that a 'pure' capitalistic economy can assure social equity); and the existence of these laws may already be regarded as confirmed by the inauspicious effects of holistic centralization in the former Soviet Bloc, as it is by the lasting effects of economic lobbies on Western governments.[28] One can also hypothesize an acentric holism, according to which any disturbance in the network of socio-economical relations is propagated everywhere, as in a generalized butterfly effect. But this mandates careful examination of a whole battery of boundary conditions and of *specific* affordances of the systems involved. *Chaotic* models are not of direct application to such highly structured dynamical systems. In particular, Neurath's boat has watertight bulkheads. It is a delicate task to explain their precise structure and function (and their bursting points).

The admission that reason itself should be 'historicized' calls for a higher-level use of reason, in the task of working out such a historicization. The same problem recurs at the next stage, just as for the sequence of *homunculi* in the system of mind (and of meta-theories in the hierarchy of formal systems). Unless there is a Meta-historical Rationality, which, however, would be by definition unattainable by our history-laden minds. Yet, neither infinite regress nor metaphysics exhaust the range of possibilities in front of us. When reason is criticized, this is possible because what is criticized is not the totality of rational procedures, some of which are used in the criticism itself — in fact, the same totality can be arranged in different ways. Reason is neither monolithic nor *given* all at once. So, there is no need to posit a historically variable reason, since our reasoning activity already intrinsically embodies a multiplicity of perspectives. By a process exactly like local *epoché*, we test certain aspects of rationality by appealing to others (or to a different combination of the same aspects). This is why we can agree with Galileo when (in the letter to Liceti, Sept. 15, 164O) he says that, if Aristotle came back to life, he would agree with Galileo's way of reasoning, and not with that of Aristotelians.

So, is there no place for the dynamic character of reason? No. It simply is no longer to be sought in a micro-description of the actual historical becoming of knowledge, since

---

[28] Notice how our epistemological view also provides a model for socio-economical relationships.

that would be an empty search: in history there are too many interactions with contingent features of intersubjective life, producing unpredictable butterfly and Oedipic effects in almost any sphere of human activity. Nor is it to be sought in some sort of eschatology. The dynamic character under investigation is far more accessible within the structures of mental development, and by now some of these structures can be mathematically described. I would add that consciousness of how a given concept or theory has been historically developed over a given period is neither necessary nor sufficient for us to understand it: rather, one needs to run over again, in the first person, a definite sequence of cognitive steps. To have an idea of what historically took place provides an aid, but one has to take a suitable schematization, as a pedagogical artifact. And if we remove pedagogical considerations altogether from the study of history, what remains would no longer be of such philosophical importance. This applies particularly to history of philosophy: it is simply another instance of the way the possible and the actual (no less than norm and fact) are to be viewed as braided.

A representative form of soft historicism has been recently offered by Richard Rorty. In the Introduction to Rorty (1979) it is explicitly affirmed that the message common to Wittgenstein, Dewey and Heidegger is historicistic, since it has to be seen as set against the absolutization of problems and solutions in philosophy (or anywhere else), through their embedding into the historical reality of culture.

Rorty sees a strong bulwark of this historicism in the arguments adduced by Quine and Davidson, respectively refuting the dualism of language and fact, and of scheme and content. By means of these arguments, the problem of how language *latches onto* the world vanishes, hence the classical realism/antirealism problematic vanishes too: meaning and reference change together with statements historically accepted as true, and there is no definite border separating theory-change, or a change of subject, from a change of conceptual scheme; there is no longer any need to maintain the Kantian requirement of an all-embracing, neutral and permanent matrix of rationality, within which to frame human knowledge. Rorty appeals to Wilfrid Sellars' view of science: its ‘rationality’ consists in being a self-regulated enterprise each part of which is open to revision, as the metaphor of Neurath's boat emphasized .

Albeit ‘soft’, this historicist view would not fall into the flaws of traditional cultural relativism. The claim that no representation has a privileged relation with nature would not herald relativism, for neither foundational attempts nor their negations can transcend social practices — of language, first of all. Yet, Wittgenstein's anthropocentric equation of facts with a matter of language use is denied by Rorty, since nature and convention are terms that stand or fall together. If there are only *linguistic* facts, the notion of ‘fact’ loses its semantic value and utility: we are simply relabelling old facts as linguistic, and in this way confusing the Quinean thesis that there is no definite and stable point where language and

world are separated, with the thesis that there is no separation at all, not even internally to the field of epistemic forces.

So, here we see in action a correct phenomenological response operating against a disguised form of reductionism, *viz*. the collapse of the variety of our experience into linguistic practice. Unfortunately, Rorty's world tends to be, in the end, a world of texts (the claim that text and world differ, and the difference can be but linguistic is too cheap a reply) and so he fails to develop the above objection consistently: this is evident when he denies the existence of any way of getting from the analysis of an organism/world interface to the critical analysis of an organism's beliefs about the world. Thus, we can assign epistemology only a descriptive task, without guarantees of success, releasing it from the onus of achieving knowledge of truth and truth about knowledge. There is no guarantee of success in *any* human enterprise, much less one in which the issues ramify into such a delicate tissue of interdependencies as in epistemology.

Rorty opposes *any* project of identifying a source of necessity, be it in things (Aristotle), the mind (Kant), or even language (as logical empiricists and analytical philosophers have claimed). In this vein, the late position of Wittgenstein is assimilated to that of Dewey,[29] with positive implications on our way of doing philosophy, liberating it from 'academic' tradition: philosophers should finally give up the idea of the mind as a mirror containing representations to be investigated by 'pure' methods, and the idea of a culture guided by overall principles or by hard facts. Culture would not become, thereby, *decadent* .

Still, an immediate objection is at hand: everybody can find and use successful criteria and methods to obtain the required degree of certainty in any given subject. We have already seen how skepticism and metaphysical realism are faces of the same coin. And this recognition is to the point here, in opposition to the brand of soft historicism that nowadays is making many proselytes within the camp of «continental philosophy». I agree with the renunciation of the mirror paradigm, but, if we really do not want to fall into a decadent project, the way out of traditional speculative philosophy is nearer to the sort of naturalistic epistemology I argue for, with respect to the concerns of mind and language, and to the sort

---

[29] Here I shall avoid discussion of the difficulties of merging the legacies left of such different philosophers as Wittgenstein, Dewey, Heidegger and Davidson, within one well-defined consistent perspective, difficulties accentuated when the perspective is also intended to reconcile the main features of pragmatism and holism. Davidson is certainly right in being embarassed at his inclusion in the pragmatist camp, and Dewey's transactional model appears to require a world that is independent of language. Whereas, if the common orientation of these thinkers is towards a more or less historicized pragmatics, other difficulties come back — as made clear, in relation to language, by Husserl's and Chomsky's arguments concerning the autonomy of expressive structures from communication, and as regards logic, by the failure of the New Rethoric emphasized by Preti (1984).

of realism which, with respect to general themes of philosophy of science, has been defended by Hacking (1983). Both lines of reasoning are incompatible with Rorty's historicist pragmatism. Yet they do not fall under his objections to classical epistemological programs. In particular, the place I assign to epistemology no longer requires i) that we know our own mind better than anything else, and ii) that we can know everything about mind even if we know nothing about anything else.

Who will not wholeheartedly subscribe to Rorty's polemic against armchair philosophy? Transcendentalism and linguistic analysis were alike designed to be immune from any disturbing empirical discovery. Moreover, both positions reduce philosophy to just a *Fach* among other academic subjects. Doubts arise rather to whether we should take Rorty's exit out of the *Fach*, since scientific research has to be taken seriously, that is as something more than a technological «conversation». Rorty easily gets consent when he recalls the well-known words of William James on Wundt as a paradigmatic indictment of the all-embracing attitude typical of the armchair philosopher. Nevertheless, the line Rorty is proposing to follow turns out to be similar, *mutatis mutandis*, to the line of Wundt.

This is also the paradoxical destiny of many of Wittgenstein's followers, *i.e.*, of the philosopher who fought most strongly against the traditional search for a foundation stone, hidden inside the philosopher's armchair. The paradox lies in the fact that Wittgenstein himself, in making fun of the idea of mirroring the world (and suggesting that there is nothing to 'explain'), was making room for another version of the traditional metaphysical thinker: *nihil sub sole novi* as far as the 'great' problems are concerned. Thus, no surprise that in continental Europe the impact of Wittgenstein on the fertile ground of Nietzsche's nihilism and anti-scientific reaction has turned out to furnish the best possible alibi for avoiding serious confrontation with the *specific* problems left unsolved by the crisis of the logico-empiricist tradition and the stalemate of the Chomskian universalism in semantics.[30]

Furthermore, Rorty's appeal to the Wittgensteinian «forms of life» already finds several of its ingredients in early twentieth century discussions by Wilhelm Dilthey and Georg Simmel on the interinvolvement of culture and nature. In both these thinkers, the dynamic character of Hegelian *objektiver Geist* emerges in much more breadth than in contemporary pragmatism. For the *objektiver Geist* features the coagulation of a world of meanings that intrinsically tends to forget the constitutive activity, which gave rise to them,

---

[30] The lasting contribution of Wittgenstein's philosophizing remains associated to his maieutical brainstorming, just the role of a Zen master. And in fact, many of his arguments-by-questions may be taken as *koans*. If we want to preserve this contribution, we have to move in a diametrically opposite direction from that of the philosopher as a «synoptic intellectual», proposed by Rorty: otherwise the philosopher indeed remains once more a meta-speaker without contacts with live research. To such a «conversational» appeasement I prefer the one-stroke painting of Zen tradition. But there is no reason why this bifurcation, here referred to Wittgenstein's legacy, should be exhaustive of the paths philosophy can take.

and ends by sucking the domain of 'things' and 'thoughts' into the domain of common (institutionalized) use, while at the same time providing the Subject with a pre-structured horizon of possibilities. The tracks of such historicist dialectics were until recently scarcely acknowledged in the pragmatist camp.

Dilthey and Simmel share the design of a transcendental historicism, one designed to ground the possibility of scientific investigation, though *sui generis*, of the many forms of *Verstehen*. This wing of the hermeneutic movement found its best realization in the «objectivism» of Emilio Betti. On the opposite wing Hans Georg Gadamer is in agreement with Heidegger on the ontological pervasiveness of language, but with an additional emphasis on the historical character of comprehension/interpretation. Since I have not the space to enter the detail of Gadamer's argument, cf. Peruzzi (1992), I shall confine myself to note that, according to this second line of thought, hermeneutic analysis is impossible as an objective science, of the kind developed in the investigation of the natural world, since it would presuppose some neutral position of the *interpretans*, whereas such neutrality is excluded by the embedding of any interpretation in the context of the tradition to which the *interpretandum* belongs. Yet, if the fore of interpretation is real, material, existence, neither could we begin to search for objective theories about the physical world. Here, the ultimate outcomes of holism, soft historicism and such hermeneutical position converge. Once holism becomes diacronic, instead of syncronic, and becomes tied with historicism, meaning (as far as it remains a 'meaningful' notion) can be determined only through some form of eschatology. The ideal of clear thinking sinks in the swamps disclosed by the hermeneutical circle. One wonders whether this unhappy end is indeed inevitable. An alternative was offered by Simmel.

Simmel's attempt at reconciling Kant and Hegel brought him to distil the positive legacy of Kantian epistemology and to insert it in a dynamic framework of relations among man, history and culture. His contribution was vast and, although it involved reference mainly to specific problems of the social sciences and seems to agree with Dilthey on the methodological dualism of *Geistes-* and *Naturwissenschaften*, is even now of undoubted relevance as an example of the transition from historicism to naturalism that is possible while remaining within the background of the Kantian tradition.

First of all, Simmel is *not* an ontological relativist. His main tenets fit in with the project of a *philosophy of life*, one that, liberated from late romantic rethoric, finds its realization in recent naturalistic epistemology. It is true that Simmel, in *Die Probleme der Geschichts Philosophie* (1892), writes that «Kant has freed us from naturalism» but this claim should be interpreted in the context of the polemic of the period against a widespread psycho-sociological trend in looking at questions of epistemology, and, in any case, that claim is rejected by his later *Lebensanschauung* (1918): thus the naturalism to which he previously opposed has to be referred to various reductionist views with respect to *values*

operative in nineteenth century German methodologies of history. The cited passage, indeed, continues: «what matters is to become free from historicism»! This is in agreement with the criticism I have until already levelled against similar routes out the landscape left by the crisis of Kantianism; this sort of exit remains widely diffused in the recent symbiosis of certain aporetic outcomes of analytic philosophy with hermeneutics and pragmatism. (As is to be expected, there are analytic philosophers who reject such symbiosis, trying to recover the originally phenomenological sense of 'analysis'; but then they encounter difficulty with the privileged position accorded to language.)

Simmel intends to emphasize the founding character of semantic presuppositions (rooted within the basic situations of real life) for all generalizations that express the unity of traits of any historical subject. These generalizations are thereby neither purely empirical causal laws, nor formal laws concerning pure mental contents. The historical nature of man is something deeper than the actual contingency of socio-cultural facts: it consists of an essential tendency to overcome human products themselves; but this essence is put in danger by the alienation (Simmel says «objectification») depersonifying man in modern times, as Simmel illustrates in his *Philosophie del Geldes* (1900). From this stems also the sense of loneliness of contemporary man. Cultural creations can revolt against their creators since the created objects (be they abstract or concrete) are subject to specific autonomous laws — an anti-psychologistic theme, endowed by Simmel with a decadent significance: such laws contrast the 'laws' of life. By now, technological and abstract objects constitute so vast and complicated universes (with so many peculiar local structural properties) that the individual is trapped by them, so that people risk surrendering their lives as a tribute to the pursuit of specialist aims, whereas the realization of human potential lies in the integration of a multiplicity of possible objectifications. Life, to assert itself, has to rebel against its sedimented forms that negate its freedom.[31] And from this effort new forms originate, in turn. The last chapter of Preti (1968), while not itself focusing on Simmel, contains one of the best elaborations of this Simmelian approach to the intertwining of culture and life.

Thus we are faced with a crisp dialectical vision of relations between subject (one no longer seen as a «mirroring essence») and object. Though conscious of the idealization it embodies, I should like to represent the underlying pattern diagramatically, as displayed in the following adjunction-like picture:

---

[31] The dualism of *Geist* and *Leben* found in Scheler one of its most dramatic expressions. But its presuppostions lie in the classical claim that the 'I', by itself, is constituted by forms: it has no content. Cassirer's attempt at mediating the poles through his theory of "symbolic forms" was aimed at bringing the dualism *within* the range of the *Geist,* whereas my project points to internalizing the tension within the autopoiesis of nature itself.

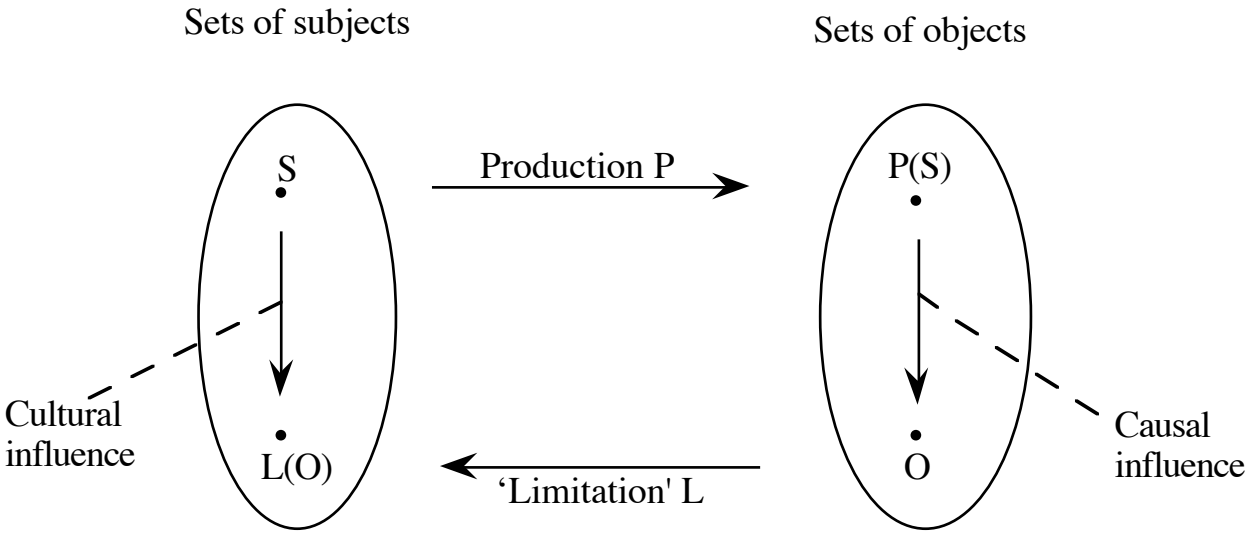


Figure 4.

Still, a globally irrationalist orientation seems to be given to the 'philosophy of life'; this would be the price paid in order to develop a more refined naturalism, one free from the dogma of a pre-established Synthesis. I consider this appearence to be misleading, and the price not to be paid. What remains relevant in Simmel and Preti is their appeal to natural guiding forces, cohesively intertwined, that mould both cultural systems and their byproducts on the large scale. Thus I gather further support for the entwined naturalism introduced above at § 3.2. From the same standpoint we can interpret Husserl's final appeal to the *Lebenswelt* (the Life-world), since its «antepredicative» sphere is the level of prelinguistic categorization, a level rich in parallel structures (such as pattern recognition, sensorimotor capacities, ethology of affectivity) providing the semantic ground for sequential symbolic representation, as realized in language.

On that same ground Heidegger tried to erect a surface-phenomenology.[32] Yet he issued only in a general instrumentalism with respect to any cultural product. Heidegger's existential analytics converges with ordinary language philosophy to the primacy of pragmatics. The assimilation of both positions within the pragmatist camp demands their supplementation by other hypotheses. With or without such supplementary hypotheses, one must remain doubtful about what further is brought to a systematic and rigorous phenomenology of the common-sense world by such approaches, cf. Smith (1992). In

[32] The admission of at least one sort of *epoché* is what basically prevents agreement with Heidegger's confinment of phenomenological analysis to the *Dasein* level, with its consequent instrumentalist emphasis. Still, I agree with the central role of *schematism*. So, local suspension calls for being completed with an understanding of the *logical* import of imagination. Such completion is presented in Peruzzi(1992a).

either approach the chance to exploit the variety of tools that cognitive sciences and AI provide us, once they are framed within a naturalistic conception of semantics, is lost. Thus we are led, instead of into «the abyss of *Dasein* », back to the cosmological roots of the pairing between the macroworld and the representational structures of living beings, in conformity with entwined naturalism.

Moreover, if we conceive science and reason in an instrumentalist vein, we are exposed to Horkheimer's charge in his famous *Eclypse of Reason* : the desease of reason lies in its arising from the human need to dominate nature, in the glance of the first man who saw the world as prey. The point is that the oppressive effect of what Husserl called the scientific-technological *Ideenkleid*, the tissue of ideas with which man has covered the natural environment, is strictly tied to the lack of comprehension of the original epistemic operations within their original context. The will to be rational — to which Husserl appeals in *Krisis* (1929), opposing tiredness (*Müdigkeit*) which he sees as the greatest danger for the European spirit — is inseparable from the drive to calculate and control the course of nature. So, it must be balanced with an understanding of the cosmological web of connections of man with nature, yielding a Taoist-like feeling of participation that does not lead to resignation or to abandonment.[33]

In view of all this, one cannot classify cultural relativism, in its multiple variants, as a particular case of naturalistic epistemology, insofar this ultimately traces (epistemic) values back to facts (about human nature). For, then, either nature itself is a cultural construction, or we have to follow the structural anthropologist in conceiving of cultures as systems of values, extraneous to the physical world. In both cases relativism becomes the by-product of a subject-centered, anti-naturalistic, philosophy. More specifically, in the former case it boils down to a version of idealism (as is manifest in C. Lévi-Strauss); in the latter case, we must sharply distinguish between social and physical nature, coming back to the dualism of *Geisteswissenschaften* and *Naturwissenschaften,* so that their respective methodologies would mark the boundaries of subject matters of radically differing ontological status. This is a well-known move, but one that has appealed to an impoverished stereotype of nature: one which hinders a mathematical understanding of the dialectics of being and becoming, as it crystallizes in the emergence of meaning.

Here, we have to stress the validity of the basic objection raised against relativism: 1) understanding the standards of cultural contextualization presupposes reason; 2) reason is

---

[33] *I.e.,* to the *Gelassenheit* theoricized by Heidegger. This does not imply a diminution of the ethical import of abandonment, which remains among the highest values, but we should not be induced to give it a dominant role, denying the importance that the ability to prove results of rational investigation and the intuition of truth have in our relationship with nature and culture. We simply ought to live our search for certainty with a smile, the smile of a being facing a participating Universe, instead of a weary refusal of the scientific adventure.

«both immanent (not to be found outside of concrete language games and institutions) and transcendent (a regulative idea that we use to criticize the conduct of *all* activities and institutions», Putnam(1983) p.234. The point is, if 1) and 2) are not suspended in an ontological vacuum, and at the same time are taken as not in need of some metatheoretical foundation, then they become legitimate themes of a naturalistic study of cognition, for they express two *facts*, not two commandments of an introjected God, nor the condemnation to the labour of Sisyphus.

Putnam claims that cultural relativism is simply *inconsistent*. But this is an excessive claim. Consider the statement

> REL. For every sentence $\alpha \in L$, α is true (rightly assertible) if and only if it conforms to the norms of the culture using L.

By substitution of REL to α, Putnam obtains that REL implies not-REL, hence (by *consequentia mirabilis*) not-REL follows. Therefore, relativism is self-contradictory. Such inference is no longer convincing, not much because the norms of our culture are fuzzy (as assertibility conditions in many cases are), but rather because the relativist could always make the rejoinder that the logical standards by which we infer not-REL from REL are in question too, and cannot be assumed as independently warranted; *in saying this*, the relativist is using these very standards, and although that dramatizes her/his situation, no one is obliged to hold that cultural norms constitute a consistent totality — a possible refuge is the type-relativization of skeptic doubts, as already imagined by Whitehead and Russell in the First Introduction to *Principia Mathematica*, and it is much harder to prove the inconsistency of the resulting typed skepticism, at least as difficult as it is to achieve a convincing foundation of mathematics through the shift from T to MT etc., as discussed above. At all events, the relativist position, whether in its typed or untyped version, is not easily refuted, by means of merely logical arguments.

Also, the use of transcendental arguments leads to incorrect conclusions unless such arguments are rooted in real properties of real subjects in the real world, and not in some appeal to abstractly 'possible' worlds — even if their multiplicity gives a sense of power to the logically-minded metaphysician who employs them.

If our culture is neither totalitarian nor theocratic, and thus «does not have norms which decide *philosophical* questions» (*ib.*, p.239), the undecidability of such questions with respect to natural sciences, as physics or neurophysiology, cannot be seen as an eternal 'success', or something independent of other specific features of our culture — but this is

just the point of the relativist. So, I doubt that such independence can be proved by a priori arguments.[34]

Two comments are to the point. First, the holist is in the least favourable position to appeal to logic as independent from cultural contingency. Second, it is curious how philosophers tend to argue against any given thesis in terms of its inconsistency, rather than those of its falsity: a reflection of their false modesty since it seems to ask for less than science, whereas it implies commitment to the certainty of a wide set of boundary-facts, and it is not at all clear why introspection or ordinary language should provide better sources of argument than the relevant bodies of scientific knowledge.

One alternative to relativism would lie in cultural imperialism: *our* own culture is to be treated as the universal standard. But this culture, being other than a monolith, allows the re-articulation as internal to it of those oppositions supposed to be intercultural. So, the onus of the proof passes back to the rationalist. And we would be back where we started in chapter 1. Why does this battle of frogs and mice seem to lead nowhere?

Briefly, the difficulty is twofold:

i) the opposing stances have been made *absolute*, *i.e.*, disengaged from any definite question, only in relation to which it would be useful to start the motor of logical reasoning, which idles if isolated. This does not mean that rationality is informal, intracultural or confined to *actual* praxis. Conversely, we have also seen that it cannot be completely formal, transcultural and supernatural. Rationality can idle and is not monolithic. If we presume to understand all this, we have to make explicit the natural roots of thought. Surely, we are able to reason about possible worlds, but they are worlds constrained by the structures of the Universe we actually exist in. Our modal *bricolage* is heavily indebted to the materials from which our most abstract tools have emerged. The philosophical game that opposes these two globalizing features, the Relativist and the Rationalist, played out for more than two millennia, has at last revealed its emptiness It has proved neither necessary nor explanatory in the solution of any concrete problem. (Of course the game has had and retains an ideological role: one whose evaluation is not in question here.)

ii) Both the relativist and the rationalist *share* the basic option of confining the philosophical debate within the bounds of purely meta-linguistic epistemology, as if their discourse could preserve meaningfulness, once separated from extralinguistic facts and the scientific problems that gave rise to the debate: in other words, they tacitly adopt global *epoché*, thus flying in the face of semantic reality.

The 'contamination' of value with fact, rather than to be avoided, is *vital* to culture, and it is neither a prelude to the elimination of value nor a step towards skepticism. The

---

[34] Similar doubts about (the abuse of) transcendental arguments, aimed at showing either the unintelligibility or the paradoxical selfreference of skepticism, have already been raised by others, cf. Stroud (1979).

opening function of values would remain unchanged even after we showed their natural structure and genesis: a system does not become less chaotic after we have established the deterministic dynamics of its single components; knowledge, on a future chemist's part, that erotic emotion is induced by a certain reaction, named *E-□□6,* would not eliminate emotion. Analogously, mutual dependence of the extension (range) of knowledge and its criteria gives another way, under the name of «the Cartesian Circle», of attacking both the foundationalist's and the reductionist's projects as fallacious. Alas, recent discussions on this circle have re-surrected typical XVII century questions, together with the recurrent epistemological stalemate they produced. The insertion of local *epoché* within the bounds of entwined naturalism escapes metatheoretical circularities like this, not by the abdication of rationality or a renunciation of the search for truth, but rather through a genetic phenomenology of meaning. Contrary to what is often held, such a phenomenology is in no way antithetic to the use of tools from AI in approaching the conditions of knowledge by means of variational methods, implemented in suitable computational devices, *provided* the resulting thought-experiments are no longer disjoint from actual environmental constraints.

The degree of interference of value and fact may range from the minimum in most perceptual phenomena, as the incorrigible character of optical illusions reveals, to the maximum in sophisticated meta-cognitive constructions, as I argued in relation to the theory/meta-theory loop. However, understanding the natural roots of values leaves their axiological function unchanged. So here too there is room for a wide phenomenology of reason, extending well beyond its linguistic manifestation.

Only after all these provisos have been made can one subscribe to the conclusion drawn by Putnam (1983) p.47:

> If reason is both transcendent and immanent, then philosophy, as culture-bound reflection and argument about eternal questions, is both in time and eternity. We don't have an Archimedean point; we always speak the language of a time and place; but rightness and wrongness of what we say is not *just* for a time and a place.

Indeed, rather than a conclusion, it is the statement of a great open problem. And as for every open problem, it has to be further clarified in order to be solved: the present section was intended to serve just this aim; its role should be seen as breaking up the problem grid and presenting us with an alternative starting point for specific investigations, aimed at describing the way our world and our cognitive resources form a tightly coupled self-selective and self-transcending system.

§ 3.5. Varieties of Pragmatism and the Presocratic Entrance

Since the reader might have gained the impression that the position I am delineating nevertheless remains, on some points, consistent with the pragmatist perspective, it is important to make explicit the basic differences between entwined naturalism and pragmatism.

Some American pragmatists such as James and Dewey contributed to a refreshing contact of philosophy with concrete problems of extra-academic life. Still, pragmatist attempts at avoiding any sort of foundational philosophy start with the wrong step: the reduction of truth to an *instrument* for achieving some practical goal X — where X can be left indeterminate here.[35] A well-known criticism of such a conception of truth is that given by Bertrand Russell: the basic limitation of pragmatism consists in its being a philosophy in which the human world occupies the entire fore; the existence of humans on a small planet within a large cosmos seems to be extraneous to the pragmatist's perspective — and Russell stigmatizes also the moral damage deriving from such an attitude toward nature.[36]

More recently, the neo-pragmatist Rorty tells us that there is no way of knowing when truth has been reached or approximated. Such a professing of humble skepticism is not 'practical' during the actual course of any enterprise and is hardly characteristic of every form of pragmatism.[37] Nor does it help in accepting the contingency of all starting points.

---

[35] We should add that, in a letter to C.A.Strong, James clearly stated his realism. However, as soon as utility is something *objective*, it cannot remain the ultimate canon of truth.

[36] Russell's charge on pragmatism was reproposed by Popper in *Conjectures and Refutations*, as embodied in his opposition to any form of instrumentalistic reduction of scientific knowledge.

[37] Actually, truth was also defined, in pragmatist terms, as the belief on which all researchers will converge at last. Indeed, this is hardly consistent with the view that characterizes pragmatist interpretation of truth as dependent on the *effects* of any belief. One could say that what James and Dewey were trying to formulate was not another theory of truth, but rather an escape route from any attempt to capture truth, through the denial that truth has any essence at all to capture. There are no constraints on investigation other than those contextually determined by one's own interlocutors, says Rorty. Unfortunately, his attempt to revive pragmatism in this anti-positivist sense makes his arguments turn into a rubber ring suspended over a void (recall Magritte's castle on the rock): they float on a sea of facts to be viewed only in quotation marks, but then these quotation marks must in turn be treated likewise, and so on. Why on earth should contextual intersubjective constraints be effective (in the way they actually are, or are not), if not through being part of larger regularities? Why shouldn't there be *facts* concerning intersubjectivity, made possible by the architecture of those complex systems into which matter is organized during a given stage of cosmological evolution?

Even the most mechanicist of Newtonian cosmologists could concede the contingency of original boundary conditions and their being only imprecisely knowable. Moreover, accepting contingency is neither necessary nor sufficient to assure that we renounce becoming a correctly programmed machine, contrary to what is claimed in Rorty (1982). We all might be contingent machines and still possess truth as a guiding value, or intrinsic attractor, rather than as an instrument.

So, although having escaped both the Kantian engineer's pride (in having dutifully founded the System of Knowledge) and the decadent, eternally ephemeral, conversation of those who find inhuman the attraction for the extra-human realm, there still persists another antithesis, one which we have to *aufheben* in Hegelian spirit: the antithesis between enlightenment and pragmatism.

Rorty presents the appealing face of pragmatism when he says that pragmatists consider the foundational project as a wheel that has no part in the mechanism, contrasting the Platonic/Kantian philosopher who assumes a given product *p* of a season of creative thought (in any field), then transposes *p* to an abstract level, and finally invents a vocabulary in which to translate *p*, while proclaming to have *founded* it. Fortunately, our reliance on mathematics, physics, and democracy is not dictated by our achievement of such 'ultimate foundations' for them. Indeed, the foundational project has caused a growing isolation of philosophers from the life of their culture. The alternative Rorty is proposing consists of practicing philosophy in a synoptic, holistically inspired, and edification-centered, vein: something that possesses the same defects as the foundational project, simply looked at in reverse. A purely «conversational» philosophy would not convince (and possibly not help) anyone: all of its claims should be understood as being in quotation marks, once again as if global *epoché* were possible. Nor is eclectism nearer to actual (scientific, artistic, etc.) investigation than the foundational project. Moreover, holistic views have generally provided more obstacles than promotion to research in any specific domain. Finally, if pragmatic criteria have to be context-driven, they cannot be globally holistic.

Faced with the problem of distinguishing the different lines of thought coexistent under the heading of 'pragmatism', Rorty sides with Dewey, against Peirce's legacy. From the logical point of view (Peirce's contributions as a forerunner *vs* Dewey's misinterpretation of contemporary logic) and in the light of the consonance of Peirce's semiotics with aspects of Husserlian phenomenology, the reader might be prompted to take the opposite view. Yet, even if we reject the pragmatist way out of endless epistemological debates, it would be unfair to deny that Dewey has opened the way to a framework in which to integrate certain elements of Hegel's legacy with naturalism — this was also the positive assessment of

Dewey's contribution accepted by Preti. Dewey also tried to overcome the traditional gap between psycho-pedagogical and epistemological problems, and to give a naturalistic reading to «transactions» with the environment, linking subjective and extra-subjective structures in an essential way. If such a transactional project is to be pursued (overcoming the difficulties into which Dewey's attempt to deal successfully with the basic question of truth ran), it requires a deeper phenomenological understanding of axiological notions, not their *reduction* to pragmatist bases.

Dewey's search for a synthesis of this return to *praxis* with a historicist view influenced the thought of Preti, who fused (Neo-)Kantian elements with others of an empiricist character in a very distinctive way. More specifically, Preti integrated this influence within a Simmelian approach to the dialectics of life and culture, paving the way to the recognition that this strategy becomes fruitful *only* in a naturalistic setting: it cannot be faithfully embodied in the pragmatist framework. Yet isn't naturalism nonetheless ultimately *reductionist* ? To answer in the negative, I shall now discuss the objections raised by Putnam against «naturalized» epistemology, as reducing the notions of justified belief and rational acceptability to *descriptive* notions. Among Putnam's targets there is also Quine's claim that the best choice we have for naturalizing epistemology is to lean on psychology, which can at most teach us «how science is in fact».

In the past, some attempts toward the naturalization of epistemology were advanced along Darwinian lines: as any other capacity, reasoning too has to be explained in terms of survival value. Yet the dinosaurs' extinction as well as the effects of a global nuclear war might easily shed doubts on progress along the evolutionary scale, and on rationality, as promoting survival, not to mention that humans had never approached the reproductive success of insects, though they consider themselves much more intelligent than ants and bees. This, however, is too simple an argument. First, the same cases might provide confirmation of the naturalist's stance on the inseparability of cognition and reason from other factors, related to the evolution of the environment as a large integrated system, not to be examined once for all nor in an entirely non-modalized way — here, this resort to modal discourse is intended to avoid a widespread criticism levelled to I. Lakatos' dynamical methodology. Second, the arguments devised by Boltzmann, Piaget, Lorenz and Popper (among others) in support of a broadly evolutionary epistemology differ from each other, to the point of leading to somewhat diverging theses, and anyhow they are more articulated than the (circular) claim that a belief is rational if it is arrived at by the exercise of the capacity of reasoning, as if reason were a monolith instead of the result of a modular eterarchy, whose components are strictly coupled with the set of macro-features of the local natural environment.

A further objection raised against the project of naturalizing epistemology states that this project assumes a realist notion of truth as correspondence — a notion whose very

consistency has been questioned. It is not needed, however, to enter a criticism to Putnam's «internal realism», since, even admitting such a refutation of metaphysical realism, it is not granted that naturalism requires such a resort to a metaphysical notion of truth, just as it is not committed to the Darwinian theory of evolution. The main objection is rather another one, even if it is not new, being an epistemic variant of Moore's refutation of the naturalistic fallacy in ethics. Putnam (1983) mentions colours as a paradigmatic example: colours are physically undefinable, not because of their subjective character, but rather because they depend on «background conditions, edge effects, etc.», briefly on «normal observers and normal conditions» (*ib.* p.230). This implicitly appeals to *counterfactual* clauses, therefore not subjectivity, but subjunctivity is involved in blocking any reductive definition of colours. And of course, the same strategy is intended to be applied to any other substantive component of our referential resources.

The point is that this very intervention of inferential, and logically sophisticated, factors in perception is disputable. Doubtless, 'normal' observational conditions are different in defining as red a star or a shirt — and this difference has systematic features. Still, such different conditions can be (and actually are) *compared*, thus making possible our use of one and the same term such as 'red' unambiguous; second, this comparison is not a cut-and-dried matter, e.g, in the child's learning about colours, the optical system stabilizes (independently from language, and thus from subjunctive discourse) the relevant differences among red, green and other colours, suitably compensating for variable conditions of light in a way which does not involve the drawing of (unconscious) inferences; and even if culture-driven lexical parsing confers different preference weights to different thresholds in the chromatic spectrum, such a complicated process of mastery is not guided by beliefs and remains modular.

The link between the last two objections lies in a circularity argument: once truth is conceived by the naturalist in a non-metaphysical way, the resulting equation «truth = rational acceptability» leads towards a definition of reason as «a capacity for discovering what is (or would be) rationally accepted», says Putnam. And this is an obvious circle. Therefore naturalized epistemology is vacuous.

There is a double reply to this argument: i) the charge of circularity neglects the theory/meta-theory interference, which generates a dialectical *spiral* (in Piaget's terms), instead of a static vicious circle; ii) it is a *fact* that we are able to test our statements, progressively recognizing errors as well as new evidence. If rationality were an instant slot machine, or a compact monolyte, that would be impossible. But, contrary to Lakatos' methodology, this does not mean diluting judgments indefinitely. If we can pertinently inject, into the discussion of any problem, other reasons, balancing those dictated by previous reflection, it is because rationality is neither punctiform nor globally holistic. It appeals and cannot but appeal to *present* resources — otherwise such a balance would be

unintelligible. But the 'present' is intentionality-laden. So, there is no need for an omniscient Eye capable of surveying the future. Yet, the range of possible resources is not indeterminate.

How could we understand justifiability in principle if not on the ground of information concerning what *now* counts as 'in principle'? But what does count as 'in principle' cannot be understood without taking into account its objective genesis. Once again, in this recognition there is no circularity: a symptom that the classical dichotomy between a realist notion of truth and an internal, idealized, rational acceptability is not exhaustive. It is a *fact* that we are able to grasp the difference between a fact and what we know about it, as well as the difference between actual knowledge and knowledge in principle. The hypothesis that the notion of 'in principle' is independent from our actual constitution is, ultimately, in contrast with the hypothesis that the world can be rationally understood. Yet, if rational understanding has *something* to do with *how* we actually are, a large amount of research has to be pursued in order to specify the 'something' and 'how'.[38]

As to the Moorean character of Putnam's strongest objection to naturalized epistemology, a more specific reply has to be adduced. For naturalistic interpretation of any key notion in the epistemological vocabulary (such as cause, explanation, reference, etc.) is thereby called into question. The argument says that none of these concepts is definable in physical, psychological or any other, factual, descriptive, terms, because the same notions might be present also in a possible world very different from ours in respect of any physical, psychological,..., features used in the definition. Therefore, the defining notions must be 'formal', in order to avoid dependence on such contingent features. The trouble is that no purely formal notion can be intentional in the light of Putnam's anti-functionalist arguments, and Searle's Chinese Room. But then what is left other than some pre-established harmony of the natural (content) and the supernatural (form)?

However, Putnam has two more arguments, the second of which finally seems to point in the right direction.

The first argument is model-theoretic. The Löwenheim-Skolem Theorem is interpreted as proving the plurality of interpretations of any intentional notion: no operational or theoretical constraint would suffice to cause the collapse of this plurality. Hence the presumed natural relations corresponding to 'reference', 'explanation', etc. are independent of any rational constraints. Which does not seem to be the case.

---

[38] I am perfectly aware that many claims I advance depend on theoretical and empirical research in progress. Unfortunately, many philosophers argue either as if all possible research had already been completed, or as if no research would be pertinent: in neither case (unsurprisingly) could any new result affect their claims.

The argument has received many technical rejoinders, that make appeal to other model-theoretic considerations. My point is simpler. By a line of thought already exploited in ch.2, the argument turns out to be self-defeating: what is Putnam speaking about? If there is a remote possibility of an answer to this question and thus a way to justify the intelligibility of the argument, this possibility appeals ultimately to natural relations, numerous, realized in a variety of ways, and subject to being placed in parentheses (local *epoché*). It is just through such relations that understanding occurs. Incidentally, if the holist is right, he/she cannot presume his/her words will be understood by readers not already sharing the holist's entire system of beliefs, including modal views on reference. Which does not seem to be the case.

The second argument tells us that epistemic notions are essentially intentional. In order to succeed in defining them in naturalistic terms, they would have to turn out to be *identical* with inbuilt natural properties, which would render them *unintelligible*. Putnam, on the contrary, stresses the 'interpretative' character and the dependence on values, both intrinsic to our epistemic-intentional concepts, which are irreducible to a purely descriptive language, as the naturalistic one would seem to be. This unintelligibility argument is revealing. For it is a way of pointing to a phenomenologically basic situation: that of a *complementarity* of two kinds of (interchangeable) *roles* assumed by active symbols in our semantic and epistemic practices. In this sense the argument lacks the force it is supposed to have, yet it correctly identifies two functions of our concepts, and these two functions are natural-property-laden .

So, via the connection of values and norms, an objection is raised against natural epistemology (in any of its various forms): to describe how we develop beliefs, make inferences etc., is all well and good but cannot carry any consequences for how we should infer, what we should believe etc. For there is no consequence from *quid facti* to *quid juris*. So the way is prepared for a return to square one: Kant's transcendentalism as an answer to Hume. Yet the only reasons for so objecting are based on *facts* : facts such as our errors and our ability to acknowledge them, facts such as our failures consequent to the adoption of fallacious schemes, facts as our successful *insights* in problem solving. Facing the further fact that we follow conventions, one cannot be content with saying: their being conventions means that physical and biological laws do not oblige us to behave in this, instead of that, way, when we are driving, arguing, playing. Actually, as suggested by David Lewis, any convention is a solution to some practical problem of intersubjective coordination. But then a theory of conventions is hardly based on conventions: it is a branch of ethology, not of linguistic analysis. It describes regularities in the behavior of humans *and* in the architecture of mind. So it would be at least curious for a holist to resort to the above anti-reductionist argument: behind the dualism of fact and value there can once again be seen the untenable dualism of language and fact, as well as the more recent functionalist use of the silicon

proportion. Then, it is necessary to ask: how and why do we follow conventions, and couldn't any arbitrary convention be chosen? The answer is not provided by a priori considerations, but only from our patient study of nature — including the way nature is 'coded' within the mind.

Likewise, the normative character of certain schemes of thought is justified only on natural grounds. We can no longer expect to solve the classical problems of epistemology, by remaining in our armchair for the exercise of meta-reflection. Shouldn't the philosopher frequent the world with the same curiosity (and caution) as a scout? Neither imagination, logical rigour nor experimental accuracy should be missing from the scout's knapsack. The same holds for the philosopher exploring the landscape of cognition and valuation. In this way philosophy comes again to be an integral part of the scientific adventure, in a renewed Presocratic spirit.

What thereby remains of transcendental analysis? First, it becomes an investigation into the *project* of Neurath's boat, and since there exist other species endowed with cognitive *apparati*, it then becomes the investigation of all these projects. Second, if biological evolution can be seen as an essentially epistemological process, inasmuch as epistemology cannot be uprooted from the growth in complexity of biological structure, there is no longer any gulf separating such comparative transcendentalism from the empirical research of scientists. Third, it becomes the study of possible (also artificially designed) projects, once parametric boundary conditions are fixed, in order to control variational procedures. Present AI is a first step in experimental epistemology. Which mathematical notions should be used in clarifying the relationship between any such project and its variable realizations, as well as the relationship between different projects and their equivalent output? Here enters the notion of *universality*, as elaborated in category theory, makes its appearence. It is this which makes possible a unified approach to that complementarity of plasticity and schemes, required by both natural and artificial intelligence. (Adapting an idea of Bill Lawvere, my conjecture is that the growth of complexity levels for natural structures can be represented by a sequence of 'downward' functors having left and right adjoints.) With this unification we avoid a simply descriptive evolutionary epistemology no less than an epistemology without the knowing subject.

In such a perspective, the classical dichotomies innate/learned, a priori/a posteriori, rationalism/empiricism, come to lose their traditional absoluteness. They will remain only as approximations to different architectures in a more complex dynamics of modules, differently realized in different system/environment pairings. Kant could not admit the natural character of *a priori* components, for it would imply inserting them within the same phenomenal world that they were intended to found. I have tried to show that the involved dualism of form and content and the naturalistic fallacy are both neutralized by a dialectical conception of epistemology, linked to the local character of *epoché*, and framed by the

principle that the Universe self-selects the property of being symbolically represented — although it does not determine uniquely the ways of such representation: they depend on local features and constraints imposed by the complexity of any given system, no less than on the specific pairing with its changing environment.

Since the 1950's James Gibson has underlined the need for studying outside of the laboratory the way humans see *as they move* and causally interact with the whole manifold of objects present in the natural environment where information abounds, so that senses have only to pick it up. The movement of «ecological psychology» that he initiated, has rejected many basic tenets of mainstream psychology. For Gibson denied information processing, representations, interpretative models, inferential strategies: it would suffice to explore the world and to discover the specific «affordances» for any given organism, in order to eliminate ambiguity in any perceptual scene.

Since on several occasions the thrust of my naturalistic proposal may appear to be closedly linked with Gibson's methodology, it's timely to make it clear that the dialectical notion of epistemic braid to which I have often referred prevents agreement with the view that invariance comes only from external reality and that thus it does not need any epistemic contribution. Although for any given object O and situation □, there are structural eco-constraints on the potential activities (of a specified organism S) inherent in O or □, these constraints are insufficient to reduce to *one* predetermined set the range of affordances available for the given organism, once the representing system of S is sufficiently complex to provide for the emergence of active symbols. (In category-theoretic terms, cognition is *fibered* over perception.) Herein lies the basic openness by which interpretation and combinatorial manipulation of thoughts enter the picture. So naturalism cannot be globally extensional either. With Dretske (1981) I agree in claiming that the content of any mental representation is a function of its origins (and environmental boundary conditions), but this function also contains *other* parameters, whose tuning permits local *epoché*, abstraction, eidetic variation and finally philosophy itself. On the one hand, the interactions with the environment are internalized as a basic condition for conceptual *stability*, on the other hand cognition always requires some *ceteris paribus* clause, but the class of such clauses varies in a not arbitrary fashion. This permits the enrichment in structure of the initial cognitive situation. In the same spirit, observability can be related to the (weak) Anthropic Principle, involving deep hypotheses on the structure of cosmic becoming: an arbitrary universe would not afford the conditions necessary for the existence of observing beings. Conversely, given the existence of cohesive natural totalities, the possibility of representation of their components by any system itself embedded in such totalities is constrained by laws of emergence. The data-types of a human being can be limited and their organization can be recognized as subject to error (on the ground of *other* information accessible to the mammal who recognizes the situation as one of breakdown),

but they remain extraordinarily rich and structured, as vision, motion and manipulation in human beings indeed are. It is from these ingredients that any proposition takes its constituents of meaning.

References

Y. Balashov (1991) The Anthropic Principle, *American Journal of Physics* **59**.

J. Barrow, F. Tipler (1986) *The Anthropic Cosmological Principle*, Oxford University Press, Oxford.

J. Bruner (1986) *Actual Minds, Possible Worlds*, Harvard University Press, Cambridge (Mass.).

D. Dennett (1971) Intentional Systems, *Journal of Philosophy* **68**, now in Haugeland ed.(1981).

F. Dretske (1981) *Knowledge and the Flow of Information*, Bradford books, Cambrdige (Mass.).

J. Fodor (198O) Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology, *The Behavioral and Brain Sciences* **3**, now in Haugeland ed.(1981).

J. Fodor (1983) *The Modularity of Mind*, MIT Press, Cambridge (Mass.).

I. Hacking (1983) *Representing and Intervening,* Cambridge University Press, Cambridge.

J. Haugeland ed.(1981) *Mind Design*, MIT Press, Cambridge (Mass.).

D. Hofstadter (1979) *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books, New York.

D. Hofstadter (1985) *Metamagical Themas*, Viking, New York.

R. Jackendoff (1983) *Semantics and Cognition,* MIT Press, Cambridge (Mass.).

P. Johnson-Laird (1983) *Mental Models*, Cambridge University Press, Cambridge (Mass.)

F.W.Lawvere (1989) Qualitative Distinctions Between Some Toposes of Generalized Graphs, in J. Gray, A. Scedrov eds., *Categories in Computer Science and Logic,* American Mathematical Society, Providence (RI).

E. LePore (1988) The Indeterminability of Determinants: Quine on the Ontogenesis of Reference, *Lingua e Stile* **23**.

S. MacLane (1986) *Mathematics: Form and Function*, Springer, Berlin.

J. Macnamara, G. Reyes (1990) The Learning of Proper Names and Count Nouns. Foundational and Empirical Issues, to appear in Id. eds.(1993).

J. Macnamara, G. Reyes eds. (1993) *Logical Foundations of Cognition,* Oxford University Press, Oxford, in press.

C. McLarty(1992) *Elementary Categories, Elementary Toposes,* Oxford University Press, Oxford 1992.

M. Minsky (1985) *The Society of Mind,* MIT Press, Cambridge (Mass.).

U. Neisser (1976) *Cognition and Reality. Principles ans Implications of Cognitive Science,* Freeman, San Francisco.

P. Parrini (1976) *Linguaggio e Teoria*, La Nuova Italia, Firenze.

D. Pavlovic (1989) On Structures of Paradoxes, # **57O**, Mathematisch Instituut der Rijksuniversiteit, Utrecht.

A. Peruzzi (1987) Il Soggetto della Logica, *Teoria* **7**, II.

A. Peruzzi (1988) *Noema*, F. Angeli, Milano.

A. Peruzzi (1988a) Forms of Extensionality in Topos Theory, in C. Cellucci, G. Sambin eds., *Temi e Prospettive della Logica e della Filosofia della Scienza Contemporanee*, vol.I, CLUEB, Bologna.

A. Peruzzi (1989) Towards a Real Phenomenology of Logic, *Husserl Studies* **6.**

A. Peruzzi (1989a) The Theory of Descriptions Revisited, *Notre Dame Journal of Formal Logic,* **3O**.

A. Peruzzi (1990), E' morale la filosofia della morale?, in *Il Pensiero di Giulio Preti nella Cultura Filosofica del Novecento*, F. Minazzi ed., F. Angeli, Milano.

A. Peruzzi (1991) Meaning and Truth: the ILEG Project, in E. Wolf, G. Nencioni, H. Tscherdanzeva eds., *Functional Semantics, Structure of Meaning and Pragmatics*, Academy of Sciences, Moscow.

A. Peruzzi (1991a) Categories and Logic, in G. Usberti (ed.), *Problemi Fondazionali nella Teoria del Significato,* Olschki, Firenze.

A. Peruzzi (1991b) Prolegomena to a Theory of Kinds, in J. Macnamara, G. Reyes eds.(1993).

A. Peruzzi (1992) *La Mente In-esistente*. Note del Seminario su Semantica e Cognizione, Università di Firenze, Firenze.

J. Petitot (1983) *Morphogenèse du Sens*, Presses Universitaires de France, Paris.

K. Popper (1957) *The Poverty of Historicism*, Routledge and Kegan Paul, London.

G. Preti (1968) *Retorica e Logica*, Einaudi, Torino.

G. Preti (1973) *Umanismo e Strutturalismo,* Liviana, Padova.

G. Preti (1976) *Saggi Filosofici*, vol.I, M. Dal Pra ed., La Nuova Italia, Firenze.

G. Preti (1984) *Logica e Filosofia*, A. Peruzzi ed., F. Angeli, Milano.

G. Preti (1986) *Il Problema dei Valori,* A. Peruzzi ed., F. Angeli, Milano.

H. Putnam (1975) *Philosophical Papers,* vol.II, Cambridge University Press, Cambridge.

H. Putnam (1981) *Reason, Truth and History*, Cambridge University Press, Cambridge.

H. Putnam (1983) *Philosophical Papers,* vol.III, Cambridge University Press, Cambridge.

H. Putnam (1988) *Representation and Reality*, Bradford Books, Cambridge (Mass.).

R. Rorty (1979) *Philosophy and the Mirror of Nature*, Princeton University Press, Princeton.

R. Rorty (1982) *Consequences of Pragmatism*, University of Minnesota, Minneapolis.

S. Sajama, M. Kamppinen (1987) *A Historical Introduction to Phenomenology*, Croom Helm, London.

J. Searle (198O) Minds, Brains and Programs, *Behavioral and Brain Sciences* **3**, now in Haugeland ed.(1981).

B. Smith (1992) La struttura del mondo del senso comune, *Iride* **9**.

P. Strawson (1985) *Skepticism and Naturalism*, Methuen, London.

B. Stroud (1979) The Significance of Skepticism, in P.Bieri, R. Horstmann, L. Kruger eds., *Transcendental Arguments and Science*, Reidel, Dordrecht.

R. Thom (198O) *Modèles Mathématiques de la Morphogenèse*, Bourgois, Paris.