A MULTILAYER PERSPECTIVE ON HUMAN COGNITION: A SPECULATIVE OVERVIEW

HENRIQUE SCHÜTZER DEL NERO[1]
JOSÉ ROBERTO CASTILHO PIQUEIRA

Institute for Advanced Studies
University of São Paulo, Brazil

Este artigo contém algumas idéias sobre uma disciplina recentemente criada, que se chama Neuropsiquiatria Cognitiva. As bases principais desta disciplina são as seguintes: (1) Um novo modelo de Psicopatologia que pode ser construído a partir de algumas noções tiradas do conexionismo. (2) A relação entre o desvio quantitativo no nível sináptico e a sensação patológica qualitativa no nível mental, comum em psiquiatria clínica. (3) Arquiteturas conexionistas podem servir como algoritmos interessantes para explicar esta relação. (4) A relação entre o mental e o físico deve ser explicada não em termos das disposições mas em termos do mecanismo que internaliza a psicologia "folk" e as expressões intencionais. (5) Em tal mecanismo acontecem coisas segundo as leis da Física, mas a sensação é que nos níveis mentais e sociológicos superiores coisas como crenças e liberdade realmente existem. Se isto é um erro do ponto de vista físico, é a pré-condição da construção do homem social. Segundo esta visão dos fenômenos conscientes, a individualidade e as categorias mentais não são descrições ou interpretações fora de qualquer sistema cognitivo, mas são conceitos absolutamente necessários que sustentam as estruturas profundas da condição humana.

This paper contains some ideas about a new born discipline called Cognitive Neuropsychiatry. The main foundations of this discipline are the following:

---

[1]HSDN: Graduate in Medicine and Philosophy. Specialist in Psychiatry. M.Sc.in Philosophy. Coordinator of the Group of Cognitive Science (Institute for Advanced Studies-University of São Paulo). JRCP: Graduate, M.Sc. Ph.D in Electronic Engineering. Professor of the Eletronics Department, Polytecnic School-University of São Paulo.
Institute for Advanced Studies-University of São Paulo. Av. Prof. Luciano Gualberto, Travessa J, 374 Cidade Universitária - 05508 São Paulo-SP-BRASIL Phone:(011)2129421 - FAX:(011)2119563 - Email:RABJUNIO2402L3 address: R@ BRUSP (bitnet).

(1) A new model of Psychopathology that can be constructed using some notions from Connectionism. (2) The relation between quantitative deviation at the synaptical level and qualitative pathological sensation at the mental level common in clinical psychiatry. (3) Connectionist architectures may serve as interesting algorithms to explain this relation. (4) The relation between the mental and the physical must be explained not in terms of dispositions but in terms of a mechanism that internalizes folk psychology and intentional idioms. (5) In such a mechanism things happen according to the laws of physics but the sensation is that in the upper mental and sociological levels things such as beliefs and freedom really exist. If this is a mistake from the physical point of view it is a pre-condition of constructing social man. According to this view conscious phenomena, individuality and mental categories are not descriptions or interpretations outside any a cognitive system, but are absolutely necessary concepts that sustain the deep structures of human cognition.

INTRODUCTION

In this paper we shall try to put forward a speculative position that endorses a cognitive program of research by launching the foundations of what we call Cognitive Neuropsychiatry. The central status of Psychiatry in this view is that we have to move toward a model of research that enables us to work with psychopathology with all the suggestive ideas of Cognitive Science, in spite of the fact that some will find such a plan speculative because the concrete objects of Cognitive Science are nowadays much simpler. As a matter of personal belief we have suggested elsewhere that the adoption of a very restrictive object to be modeled can lead to a misunderstanding of the very nature of human cognition. If we do not consider the fact that certain categories, e.g. consciousness, belief, will, etc. are part of our mental life and may have a crucial role in the processing of low levels of cognitive activity, e.g. lower level vision, we will not constitute a real project by moving from the simple to the complex, because this complex "emerges" and feeds back to the lower levels in a way that must suggest to us a kind of combined bottom-up and top-down process. This is what we are going to describe.

1. THE NATURE OF THE MODEL

1.1. What is it to be a model?

There are at least three important senses of model. We are not using the term model in either the mathematical sense or in the logical sense. We are suggesting that a model is a kind of formalism that illuminates the nature of the solution to the problem that is being proposed. If there is a strong or a weak interpretation of the model, that is another issue. As a matter of fact, we shall adopt the kind of algorithm that comes from neural nets (or Parallel Distributed Processing or Connectionism). For all purposes we adopt these three terms as synonyms. We are not claiming that there is a system such as the one that we shall describe, nor even that it might exist. We are merely conjecturing that as a metaphor the model helps us to articulate some ideas about the relation of mind and brain.

1.2. The model

In summary, let us describe our model as a multilayer neural net. Each level, as a matter of semantic interpretation, is an independent net. These will be called $n_1, n_2 \ldots$ and so on. Each level receives input from the lower level and sends output to the higher level. The further up you go in the network, the less concrete it will become, until it becomes just a virtual network – i.e. one in which each individual is a node, and social and cultural rules are the links that keep people together, acting mainly through ordinary language.

It is not time yet to describe in detail how many levels there are and it is also too early to label each level prescribing a type-correspondence between levels in our model and ordinary macrostructures like neurons, brains, and society. Last, but not least: each level sends feed-back to the lower levels by modifying weights in their connections.

The core of the functioning: we adopt the position that there are states in each level that tend to be far from the relaxation point and if

we leave them without external correction of weights they will waste a lot of time in reaching a new point of stability. In this case, we think the immediately superior level pushes the lower level towards stability according parameters and inputs, by modifying the values of the states.

A second class of considerations comes from the fact that complex dynamical systems are non-linear and therefore they will present bifurcations maybe up to chaotic behavior, depending on the parameters. This tends to be crucial because we adopt the view that whenever the system is chaotic at level n, level $n+1$ will detect it and will correct parameters of n in order to allow stability in the space of states.

1.3. Conjectures about the origin of the model

The multilayer nature of our model be explained as follows. Each computational operation performed by an element has a cost function expressed in terms of memory and time. Then a cost functional can be constructed and the system tends to operate minimizing that functional. The one-layer net, i.e. without hidden units, could operate in this optimal situation, but it would be very difficult to reach it due to the high number of different possible topologies to test. This is part of a reductionist strategy we have been adopting. In the limit everything without respect to cost function is a one-layer system and psychological behaviorism is correct in positing that the intermediate level is only a matter of fixed and immutable transformation. But psychological behaviorism is false because under real constraints (environment, real time response, etc) it is necessary to create intermediate levels in order to diminish the time needed to compute the cost function by reducing the number of possible topologies that can be mapped from n parameters of input to n parameters of output.[2]

---

[2] If one imagines a one-layer neural net (a Rosenblatt convergence machine) one will have the number of topologies that must or can be mapped from an n space of input parameters as $2^{n(n-1)/2}$. (Adopting a discrete bi-valued truth table between states of the links). If one had to adopt a probabilistic multi-value link between each node, as seems to happen in real synapses, one would have to complexify

The role of intermediate levels is to diminish the time of computing the cost function allowing the system to create "internal representations" of the problem being handled. We adopt the term between quotation marks because we don't really think the term "representation" is a good one since its historical meanings are mainly attached to propositional attitudes.

In mathematical terms, the function of intermediate levels will be to permit calculating the cost function for a high number of different topologies without loss of operational efficiency. From a more fundamental point of view, in spite of allowing interesting generalizations, we think that the resource of intermediate levels is what shadows the problem of type-reduction allowing the appearance of ghostly guests: emergent predicates[3].

1.4. Type-reduction at the level of the "being" and adaptive emergence at the level of the "senses"

An idea we shall put forward is that type-reduction, or a one to one correspondence between mental-behavioral states and environmental-brain states, is the goal of scientific psychology but is not real regarding the practical problems a human being has to solve. This is why intermediate levels are recruited. They allow a "fuzzyfication" of the correspondence in terms of the identification between mental states and brain states, but they also allow a decrease in the time needed to compute a cost function according to real time constraints. The role of the mental is to be adaptive and not to be a precision machine. This kind of deviation from the reductive character of the mapping, e.g. from the environmental-brain parameters to the mental-behavioral ones, will

---

the equation. For present purposes let us assume a discrete non-continuous liaison between each "neuron" in one's neural net.

[3] For an explanation of type reduction, token reduction, emergent predicates and numerous concepts that are not defined here Cf. DEL NERO, H 1992 (passim) *O problema da mente na Ciência Cognitiva* (manuscript, master's dissertation FFLCH-USP).

force us to construct a fuzzy description-language to designate mental states. This is the time for ordinary-mentalistic language to appear, with intentional idioms such as beliefs, hopes, fears and so on.

The issue here is that the enterprise of reducing mental concepts to physical concepts, or mental states to brain states, is doomed to failure because the correspondence cannot be one-to-one in order to allow the mental to be a quasi-veridical version of the environmental, concrete or abstract world that pushes us on to quick and functional solutions.

It is time to define how we are using the term quasi-veridical. In a nutshell, we think there is no such thing as the real world except for atomic or very simple concepts in each level. For instance, a red object is a red object according to the level where the proper vocabulary contains red as a primitive of color description. Of course, within the limit that the categories of the surrounding world are not those that we name. Our putative movement to build and to label categories is a kind of remodeling and reshaping of a virtual problem-space in order to grasp from it the best inputs and to give the best outputs back. As a matter of constructing the theory every categorization is a contingent one regarding a certain domain. This is why if one defines the domain he (or she) is working with, and if the articulation of the pieces of one's theory is good, every theory will be quasi-veridical or pragmatically veridical in relation to that well-posed domain.[4] The question whether there are or aren't correspondences between our "theories" in each level and the "real" structure of the world is, in our opinion, ill-posed because although human beings process information as animals concerning the concrete world, the mark of humankind is not to process low-grade information of the concrete world, but to process high-grade information about abstract, possible and virtual worlds, e.g. counterfactuals, theoretical terms, ideals of conduct and virtue. This is why language is a powerful tool allowing the upper level of social relation-

---

[4]The notion of pragmatical truth we are using is due to the ideas of Newton Da Costa.

ship to behave as if it were a big neural net, each individual being a node apt to modify his or her behavior according to the links of a rich interpersonal communication device. Moreover, if one asks us if each individual is a node in a localist or in a distributed big neural net, the first being a strict correspondence between node and individual, our answer is, that depends. For some purposes each individual is not a node of a big net, but an emergent result of a distributed representation of individuality, and this is why the sum of the individual's behaviors is not apt to predict or explain certain macro-social predicates. If one thinks about individuals, primitive sensations, each individual is a node, but in order to build a rich interpersonal system some predicates that will define each individual are not represented in a localistic way, and are therefore emergent with respect to the individual level. Certain macro-psychological-social predicates are regions of stability in a n-dimensional hyperspace that is constituted by individuals but that is not explained in terms of the sum of individuals (*pace* Durkheim and Jung). Explanation of mental states occurs in two ways: in the localistic version we explain only very primitive sensations. However, if order to grasp the core of human mentation we have to adopt the distributed representation. Being somebody is not having a body and a certain amount of raw feels. It is now being part of a region of stability in a cultural and contingent hyperspace.

Imagine, for example, that we are trying to make a correspondence between the morals of the Prince (the governor) and the morals of the single individual (*pace* Machiavelli and Shakespeare). There isn't always a correspondence between these two predicates, otherwise the governor could not order a bomb attack for political reasons and at the some time sustain, as an individual, inflexible views about killing. The concept of honesty in individual respects is sometimes different from the concept of honesty in public respects. Why do we understand a Chicago story of provoking a big recession in order to keep the public deficit and inflation at a good level? Because the concept of being fair

in the social net doesn't always map directly onto the domain of being fair at the private individual level. It is quasi-veridical that the virtues are the same. Sometimes, as a matter of level things become different.

Now we are ready to define the space of problems we are dealing with. Is it possible to have an account of the relation between the mental and the physical, or between mental states and brain states or between psychology and neurology? The answer is yes if we correctly divide all the problems involved and define the precise status of the relation between description and being.

In a few words we adopt the position that the further up one goes in the levels of this multilayer network the more the appropriate language for describing the level in question becomes functional, adaptive and not constructed with natural categories. In other words, the primitives of ordinary language are not made to be natural categories but mental-social categories and this is why they cannot be projected in a one-to-one correspondence onto physical categories, i.e. at each level type-reduction becomes more difficult. The reason is that each level has its own appropriate language of description, e.g. for certain levels it will be quantum theory, going up it will be classical physics (Newtonian) and at the level of the mental states and social interactions, ordinary language will become the best tool for the description job (this is Lycan's view of homuncular functional levels of description). But the problem is that: first, we don't have only ordinary language and folk psychology as a correct holistic description of our dispositions and behaviors, but we consciously sense ordinary and mental language categories and not dispositions. Second, it must be stressed that the role of the intermediate levels is to decrease the time wasted in computing a cost function allowing the best match regarding time constraints. But with optimizations, comes also a new language that will describe the new constraints of the level. This is why ordinary mental language is not precise, it is functional and it has an intermediate role of diminishing the time required to do the best regarding all the concrete

limitations. Adaptive mechanisms are not ideal machines but efficient ones. Efficacy is not always synonymous with elegant and reductionistic solutions. This story would be solvable in mathematical terms if we could adopt a philosophical-behavioristic definition of the role that mental-intentional terms play in the system: they describe things that don't really exist at all. Intentional states are but descriptions of internal dispositions. But how do we answer those who say we consciously feel that we are "true believers"? Our answer is that the system has to have a double aspect: it functions at each level according to the general principles of physics, but it "senses", in the upper mental-social levels, as if it were manipulating intentional categories such as beliefs, hopes, fears, and so forth. Therefore, without the notion of being an individual with consciousness one never gets to understand the model in its complete analogical flavor to human cognitive processing.

2. SUMMARY OF THE THEORY

It is of course impossible to give a perfect idea of the big project of research that is subsumed by this multilayer theory. We shall try to give the nitty-gritty of the main idea and discuss it under the possibility of becoming a suggestive foundation for a new discipline: Cognitive Neuropsychiatry. To do so we are going to put forward some blocks of the project.

2.1. Eliminative materialism is right and wrong. It is right when it says that the categories of folk psychology are not natural ones, and hence type-reducing mental states to brain states is impossible. It is wrong in a version that supposes folk categories can be discarded as soon as we know enough about brain states. Folk categories are sensed by the conscious being and have "causal" power relating to action. They are not natural because as a matter of economy they have to construct functional intermediate representations that: (a) enable us to optimize

the cost function; (b) reshape the world under certain categories that are central to the constitution of interpersonal social contracts.

2.2. Consciousness is therefore the privileged locus of inflection from the physical world to the mental-social-intentional world. It is not a matter of the emergence of predicates or duplication of substance. Consciousness is the place where intentional idioms take place and where feelings like individuality, personality, feedom and responsibility are carefully assembled in order to construct social man. Emergence says that translation from macro-mental-predicates to micro-physical predicates is impossible. Everything in our model is a "physical device". Had we had the correct mathematics to predict how a level interacts with all other levels we would have been able to "reduce" or to predict behavior at all levels. However, from the very beginning, we know that prediction in such a case is, for practical purposes, very difficult.

Besides being difficult to predict, there are phenomena like non-linearity, chaos, and so on. While we do not have a correct theory about what is underlying in chaotic behavior, and what laws are behind it (ergodism may be a good way of reinstalling order in chaotic behavior), we'll have to be cautious about laws without exceptions even for the lower levels. Even at the most fundamental physical level there are going to be disruptions in laws if chaos is assumed to exist. This is why we also have to make a correct evaluation of the character of chaos with respect to law and necessity.

2.3. Entities like beliefs and intentional objects like freedom are common in our mental life. Nobody sends to jail a defendant who alleges the crime was performed under certain constraints. Even in civil aspects every contract considers that both parties involved are free of constraint. This is why we must have an instance of intermediate transformation in which the categories are appropriate to be interpreted and sensed in a manner that is compatible with mentalistic-intentional categories.

2.4. The reality of the interpretation is arrived at by a correct discrimination between description and being. The philosophical-behavioristic account of intentionality, i.e. that intentional instances are not really internal but descriptions of an internal process that contains only dispositions is right and wrong. It is right because the dynamics of the multi-layer network behave according to physical laws. The best description of behavior at the stage of being is physical-mathematical. But we already know that, at the moment, the mathematical description of such an interconnected multilayer system is very difficult. This is not a reason for denying that its ultimate structure is in principle a physical mechanism, so far as "physical" is something that means the essence of things (the model is ontologically materialistic). But the behaviorist is wrong because we perceive as if we were acting according to mental-social categories. Even the moral sensation and purposeful feelings of necessity to study a very complicated subject, no matter if it happens to be true or relevant for practical purposes is a conscious state, at least that is what is in our minds at this moment. The physicist in the laboratory may be doing physics but the mental movement that governs him is a mental state of representation in a world overloaded by concepts like freedom, self-justification, urgency, pleasure, interest, and so forth. Our model has, therefore, to demonstrate how dispositions in the behavioristic account are compatible with the physical-mathematical interpretation but at the same time there are true intentional feelings in each subject's consciousness.

2.5. We are dealing with a kind of internalization of language. Each level $n+1$ internalizes as sensation some categories of the language descriptors of the upper level. Suppose, for example, that if we are at the mental-conscious level we'll have only raw feelings before having rich interpersonal contact, fully embedded in ordinary language. If we don't have contact with other people through language we shall not internalize categories like moral judgement. This is why if level

$n+1$ (let us suppose this is the level of social interaction) has certain categories of level n (let us suppose this is the individual's conscious level) it will internalize such a category as a sensation, no matter if the system as a being never behaves according to sensations but only according to its physical rules. Each level is a physical, concrete or virtual, device (remember that at the end of the hierarchy we don't have brains any more but only individuals grouped in social nets). At each level part of the system behaves as a serial machine (*pace* Smolensky) and interprets that level with symbols that are part of the level and part that have been imposed by the upper level. Even if one is constrained to become an alcoholic (due to constraints that come from the lower levels, e.g. genetic predispositions) one doesn't sense it as a compulsion or a necessity but one always has the possibility of being able to stop drinking. In a famous decision of the Supreme Court in the U.S. judges voted regarding this issue that no matter if there are physical *a priori* constraints an alcoholic still has the freedom to stop. This is why the mental individual interprets the constraints that come from the lower level at the sensing subdevice and incorporates a degree of the categories of the upper levels. In this example the court behaves as a supra-individual virtual net. This is why we constitute a physical device that behaves under physical and mathematical pressure but "internalizes" an interpretation that is compatible with the folk-mental-intentional categories. They are part of a larger network that needs such categories.

2.6. In terms of a theory, we can affirm that intentionality is the mark of the mental in terms of the categories it contains. Beliefs, fears, moral judgement, free will, conscious experience and the sensation of being an individual are parts of the blocks we need to build a functional society. We must have an intentional idiom not only to describe dispositions but mainly to feel as if we were free and not only beings that are constrained by physical pressures. If someone asked us if this part of the

story is a cultural development we'd say: everything in natural and cultural categories is made to match environmental pressures. In terms of the complex mathematics that describe the physical implementation of this net we think everything can be explained in terms of physical idioms, but the internal economical representation that decreases cost functions may be operated in a double way: (a) as a matter of time constraints it allows a decrease in cost functioning, (b) as a matter of building the mental-categories it ought to adulterate the physical categories in order to transform, for instance, indeterminacy (at the low levels) into freedom (at the upper levels).

2.7. The part of the research that we call Cognitive Neuropsychiatry has a single fact as its point of departure: in psychiatry we have all the problems involved in contemporary discussions about neural nets.

(a) We know that quantitative deviance of signaling at the synapses qualitatively alters the sensations at the conscious level. One may feel a depressive mood and have ideas of ruin due to a decay in the rate of synaptical firing in certain neural subsystems.

(b) Drugs act mainly at the synaptical level (quantitative) and alter conscious sensations (qualitative mental character).

(c) Some environmental facts push an individual toward depressive states that are called psychodynamical and have nothing to do with the synaptical level. However, it is sometimes difficult to distinguish an endogenous depression from a dynamical one. Furthermore, sometimes a chronic state of environmental depression may trigger an endogenous state.

(d) Psychotherapy acts at the communicative channel and sometimes it can reshape some individuals' conduct. At this point we have to distinguish between support psychotherapy and hermeneutical psychotherapy. The first is directed to one level (conscious pragmatical). The second to more than one (to conscious pragmatical and to mental, not necessarily conscious but effective in the behavioral level). It is

a fact, although many people may deny it, that certain contents are not conscious in spite of having a direct effect on behavior. The relationship between levels, of each level with certain categories and the possibility explaining with a neural net algorithm-metaphor seems rich and unexploited to us. There are a lot of considerations we are not yet able to answer and there are the limitation of the length of this paper.

3. CONCLUSION

Cognitive Science must have in mind that the core of human cognitive phenomena is not recognizing a face. Animals do that well. It is worshipping, punishing, dying, killing, regretting, and beginning again. Briefly, it is building a civilization inspired by the symbol this face represents.

Cognitive scientists may think it is crazy at the moment trying to model such things. However, it is exactly because of deviance in cognitive behavior that we think some concepts from each side, from neural nets to neuropsychiatry, may be suggestive. After all, if Cognitive Science is still a big basket, why not let it be the basket case? Maybe we are the basket weavers?

BIBLIOGRAPHY

Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing.* Cambridge, Mass.: MIT Press.

Del Nero, H. (1991). *O problema da mente na Ciência Cognitiva.* Master dissertation (USP).

Hooker, C. (1981). Towards a General Theory of Reduction in *Dialogue* vol. XX 1-3.

Lycan, W. (1987). *Consciousness.* Cambridge, Mass.: MIT Press.

Nadel, L. et al. (ed) (1990). *Neural Connections, Mental Computation.* Cambridge, Mass.: MIT Press.

Piqueira, J. R. (1991). Estabilidade Estrutural. In Coleção Documentos IEA-USP.

Rumelhart, D. & McClelland, J. (1988). (ed): *Parallel Distributed Processing.* vol. 1 and 2. Cambridge, Mass.: MIT Press.

∏∏∏LOGÍ∏
Revista de Filosofía.