COGNITIVE SCIENCE AND SEMANTIC REPRESENTATIONS

Jean-François LE NY*

ABSTRACT

The main task of Cognitive Science is to construct concepts and models that would be superordinate to knowledge in the various particular cognitive sciences. In particular, one major objective is to formulate a hypergeneral description of representations that could encompass all descriptions given in subordinate domains.

A first basic distinction is between natural and rational representations, i.e. given mental entities and representations that are governed by prescriptive rules coming from logical or scientific thought. In addition, representations must be described in respect to several sources of variability, which are tentatively listed here.

Description of natural representations is based on a distinction between taken representations, which are mental events, and type representations, which are lasting structures registered in memory. The connection between them can be modelled through the concept of activation.

One advantage of activation models is their large compatibilty, not only with experimental evidence in cognitive psychology, but also with facts and hypotheses in neurosciences, and programming modes or requirements in artificial intelligence. Comprehension of natural language is a highly representative domain in this respect, which exemplifies the power of these concepts.

THEORIA - Segunda Época - Año V
Nº 12-13, Noviembre 1990, 85-106

Various, sometimes largely diverging, opinions exist about cognitive science. A major contrast is between supporters of "cognitive sciences" as a plural, i.e. denoting a set of related but distinct areas, which will still remain different in the future anf those of "cognitive science" as a singlar, denoting a single unified domain. The first view presently seems to be easier to advocate: no complete unified body of knowledge has yet emerged across artificial intelligence, cognitive psychology, linguistics, logics, cognitive neurosciences, philosophy of mind, etc., that could both meet the commonly agreed criteria of scientific thought and be considered as concerning a unique well-defined object.

We will here take the risk, as our title indicates, to accept the second view, that of one integrated cognitive science. From this point of view, which is prospective rather than descriptive, the highly abstract concept of "cognitive science" is treated as a general working hypothesis. It is based on an undisputed observation in history of present science: development of current cognitive research shows a high degree of convergence from particular cognitive sciences. This semms to be true both on the theoretical and the empirical levels: not only are concepts often borrowed from a particular domain and introduced into another domain, showing there their theoretical productivity, but also, and perhaps in more convincing way, empirical conclusions obtained in a given science can sometimes be rephrased in the language and on the level of description of another science, thus showing the compatibility between them.

This convergence surely is far from perfect. But discrepancies in cognitive science can also be related with a lack of internal unity in the various cognitive sciences: as a matter of fact, internal differences in a particular domain may sometimes be larger than those between domains.

Thus, researches can in practice focus on either similarities and common ideas between domains or their differences. They can choose to bet on unifications as long as no decisive contrary evidence is affored, or keep to plurality until a threshold they judge significative as been attained. We will in this paper try to show how the unitary hypothesis can be applied to the concept of representation.

Unification does not mean reduction, in the sense of elimination. Among people favoring unity in cognitive science, some go so far to consider the most basic sciences, i.e. neurosciences or neuro-mimetic models, as able to eventually reduce all cognitive phenomena to neural-like explanations. I will in this paper prefer to put cognitive science on the ground of symbolic representation and functional processes.

COGNITIVE AND OTHER SCIENCES AS USING REPRESENTATIONS

Basic common aspects of knowledge aimed in cognitive science stand out if we stop evaluating internal convergences or divergences in it and rather compare this domain as a whole with other large scientific domaims.

Cognitive turns out to have a very peculiar look to its object comparatively to other sciences, in particular empirical ones as physics, chemistry and biology. Whereas those other sciences have as their goal to study parts, aspects or events of the real, physical or living world, and to yield an extensive representation of it, which is potentially correct and "true", each particular cognitive science has as its goal to attend not directly to world objects or phenomena but to representations of them and modes of processing information coming from them. All of them have as their object representations of representations.

This difference, albeit very important, can be relativized indeed, if taken in the conceptual framework open by information theory, realized in computers, and expressed in a theory of extensional identity between mind and matter: i.e. if all representations of the world are considered as being themselves nothing else than parts of the real world. From this point of view, representations untimately are states or events in human brains or computers, i.e. in the natural world. They are based on structures and correspondances that, in spite of their prodigious complexity, are nevertheless constructed from single physical components relatively common in the universe (molecules and their own constituents), communication between those knowledge structures, which is a basic property of their functioning, being also a part of the physical world.

Although this idea seems to be presently acceptable in most scientific non cognitive and cognitive areas, description of representations are actually given in several ways, with different uses and senses of the word "representation" itself.

Study of representations is under the label "representation of knowledge" as a large sub-area of Artificial Intelligence. Its main goal is to yield a set of paper machine or machine memory structures that are easily manipulable in a certain type of processing, and able to conveniently correspond to objects and events in the physical world. The word "representation" in this phrase very often refers to nothing but this: "representation of knowledge" then is a coded and systematized representation of a domain of the real world. Only when the goal explicity is to yield a representation of intensional representations (for example Maida and Shapiro, 1982) is the phrase "representation of knowledge" strictly used.

In this case, the goal is not different from that using cognitive psychology representations: for example, intelligent tutorials may involve the necessity to construct a machine representation of the individual mental representations (in fact, concepts) possessed or acquired by scholars or students, in order to make a conceptual diagnosis of these, and, if necessary, "repair" them (Le Ny, 1989 b and c).

Cognitive psychology gives its own description of representations, which clearly concerns mental representations. The main difficuly is then to correctly tell the individual, idiosyncratic, part of such representations -- which non cognitivists both tend to falsely identify with their phenomenal aspects, and to overestimate -- from the part in them that is stable, and in a certain sense "objective", for example consonant to external reality, consistent with logical criteria, shared with other people, or, in a more restricted way, with expert people, but also has other psychological properties as familiarity, retrievability, Rosch-tipicality, etc.

Such a distinction is very subtle, so that we can entirely understand some authors in the past, who took an anti-psychologistic stance when they tried to characterize conceptual representations. Let us just mention Frege as an example of this, and observe that such antipsychologism was a constant historical characteristic of the whole logics for centuries. Sometimes, some computer scientists -- who appear to not clearly see the differences between clinical and cognitive psychology -- today take the same position and make and extensive use of folk psychology.

As a matter of fact, a general question to cognitive science is now: can we elaborate for this particular type of natural entities, representations, and for their modes of funcioning, an adequate, i.e. both correct or true, and practically usable, hypergeneral description, or hyper-representation, given from a natural science point of view, which correctly subsumes all other kinds of description?

NATURAL AND RATIONAL REPRESENTATIONS

A first step in this direction seems to make a clear distinction between two distinct, albeit not disjunctive, types of representations, here labeled rational and natural.

1. Rational representations, which some authors label "logical", are those submitted to the conventional requirements of scientific and logical thought: firstly consistency, and various inter-conceptual properties as set appartenance, inclusion, partial heritage, etc., and secondly adequacy, i.e. appropiate correspondance with the real world. Given these properties, which may be more or less strictly used, and in interaction with them, rational representations are as a rule shared by a whole human community, often of experts in a given domain: they are expressed through language in books, articles, lectures, etc. Sharing, or common acceptation, is implicitly taken as a major guarantee against subjectivity.

The word "knowledge" in "representation of knowledge" in fact refers to these rational represetations. "Representation" in "representation of knowledge" in addition implies, as said above, that a coded and systematized form has been given to them, that is that the rational requirements of these representations are made explicit, for example in a standard logical way, and strictly applied in order to make them representable to the second degree in a knowledge base, and usable in a computer.

2. Natural representations are by contrast those that reside in human minds, or brains, and are consequently submitted to psychological and neurological laws or rules. Thus "natural" is here an equivalent of "mental", as long as this word encompasses not only phenomenal events or entities in a mind but as well not conscious ones. Ultimately, "natural" is thus an equivalent of "cerebral", even though neurosciences presently have no means to really assess the neural nature and connections of particular representations.

How are these two types of representations interrelated? We will in this paper adopt a radical postulate: that no representation in the world exists but centered in these type 2. representations: i.e. as either a representation-in-a-mind, or a generator of such a representation, or a copy of such a representation.

Pictorial or symbolic representations, for example a photograph, a drawn triangle, a sentence, an equation, all have as their main propety and function to be external generators, in interaction with mnesic long term capabilities, of mental events having the form of representations in humans. Machine representations are first degree copies of representations-in-the-mind of the system's conceptor, and sometimes higher degree copies of other representations, for example a knowledge rule in an expert system is a copy, expressed in a given format, of a copy in a cognitive engineer's mind of a representation in an expert's mind, that is: of a copy of a given relationship in a sub-domain of the world.

It is highly productive, from a scientific cognitive point of view, to consider rational representations, and their machine copies, in this way, i.e. as a derived, rationalized subset from natural representations.

Such a subset is obtained by adding certain properties to natural representations, or, more precisely, by laying down additional restrictions on their natural attributes. For example, in an individual two distinct natural representations in a same semantic field may happen to have anyone of two possible values, positive or negative, on their atribute **consistency**; rational representations always must have the value "positive consistency".

Similary, natural representations that have the role of sub-concepts of a concept may have various values on their attribute **Rosch-tipicality** (or **represetativity**) in respect to their super-concept; for example representations of "sparrow" are more typical of "bird" in most subjects than those of "ostrich". But rational sub-concepts of a concept have no such property as a distinct tipicality value, that is they necessarily have one and only one value for appartenance or inclusion, restricted to 1.

In this view, concepts are particular contents (not forms) of representations, and have no independent existent outside minds; only objects, qualities, events, relationships that are their extensions can exist in such an independent way. In particular, abstracts representations as mathematical ones are by no means copies of "another" category of entities, existing in a different world of ideas. They are a subset of natural representations, formed in minds as a function of both: 1. definite conditions found in the real world, 2. explicit requirements for rationality, accepted by humans. We will here not elaborate this position as concerns the foundations of mathematics.

It may of course happen as another consequence of the previous view that some supposedly rational representations keep, to some degree, natural characteristics that make them incompletely rational.

SOURCES OF VARIABILITY IN REPRESENTATIONS

We can now try to present a model of this that emphasizes sources or modes of variation in representations. Let us take as an example the representation expressed by the phrase: (1) *"the discovery of America by Christopher Columbus"*. Italics here rindicate that we are dealing with the representation, not the fact.

From a linguistic point of view, it is a composed phrase, a nominalization of the sentence: (2) "Christopher Columbus discovered America". This sentence is true, since an event corresponding to it existed in the history of the real world, so representation (1) is adequate. We can say, in a way that is not actually conventional (see Le Ny, 1989 a for a discussion) but apparently acceptable by everybody, that the fact is the **extension of this representation.**

Let us now write representation (1) as: (3) R (e,A,P), in which e is a constant, the extension of the representation as defined above, i.e. the unique fact described in (2). A denotes the set of possessors of this representation, i.e. any cognitive agent in which the representation of e can reside, which we will be concerned about in a moment. P denote the properties of the representation, that will be dealed with in a following section.

We know that representation (1), i.e. with extension e, is possessed by a number of humans. In addition, machine representations can be created in various forms, the semantics of all these being e, "the discovery of America by Christoper Columbus". Thus A, as a variable denoting the entire set of possessors, can be split up in two subsets, that of humans and that of knowledge bases containing this representation.

We are thus led to consider R not as a constant denoting a unique representation, but as a variable denoting a class of them, in this case identified by their common extension e. This class contains a number of instance representations R, that vary in various ways according to their possessors, conditions of formation, times of activation, etc. The possessors seem to be the main source of variation in this class, that is, in addition to the gross difference between men and machine, on the one hand variabilty between humans, or several subsets of them, defined by their ages, geographical origins, levels of instruction, etc., and on the other hand variability between knowledge bases and machine supporting them.

But the interesting question is: how can we appropiately characterize these differences, that is which properties belonging to R, abstractly summarized under P, must we take into consideration in order to appropiately describe alll these instance representations? In other words, which are the common properties and relevant modes of variations of the representations expressed as a function of the nature and identity of their possessors?

Several types of responses to this question are possible, which are deeply distinct. As a matter of fact, each particular cognitive science gives its own description of representations, and their properties, at its own level of description.

It is interesting to firstly examine the use of physico-chemical descriptions: they primarily concern the stuff and basic neurobiological modes of functioning of representations and the contrast they lead to. As a matter of fact, this source of variation can only be now very general: even though we can adopt, at least as a working hypothesis, the view that any representation has its material realization in the functioning of a cell assembly, no method permits to observationally distinguish between a given representation and another representation, both being individualized by a cognitive property, for example their extension.

At the very general level, if we did consider this physico-chemical characterization as important we could add an especial index to our notation, for example C in: R (e, A, C, P) which would denote the material of which representations are made.

Such a source of variation would presently oppose c1 (= made of silicium, including the physical phenomena and properties associated with semi-conductors and integrated circuits), and c2 (= made of carbon macromolecules, including the physico-chemical phenomena present in individual neurons or synapses, the role of specific neuro-transmitters, formation of neural assemblies, etc.). No doubt that we will have in the future to add properties as c3 (= optoelectronic), c4 (= neuromimetic parallel hardware and architecture), and others.

The present characterization generates divergences in cognitive science. Differences in structure and functioning between a human brain and a present computer are known to be very large and to show a "thrilling contrast" (Paillard, 1987): although they did not prevent symbolic Artificial Intelligence to grow, some researches think that the future advances in cognitive science can only be obtained if the structure of computers is made closer to that of the brain.

One principle often focused by advocates of this view is the basic biological principle that "structure determines function"; from this, the idea derives that neurosciences can yield the right level of description of cognitive representation and functioning (see Changeux, 1983), and guide creation of new kinds of processing machines.

Undoubtedly development of new basic components and architecture in computers, in particular parallelism, can have a very large impact on cognitive science. In particular, a very important possibility they can bring is improvement of the capacities of learning in artificial systems, which presumably can best permit elaboration of representations as rich as those in humans.

However, we cannot believe that presently unsolved problems in cognitive science will find a solution in hardware improvement, and in modelling the structure of the brain. Contrary to the general principle quoted above, an extraordinary change substantiated by cognitive science is precisely the use of the inverted principle that, at least in symbolic cognition, "function determines structure".

In other words, since learning and organized memory are the basis for the structure of representations, psychological research shows that the early knowledge structure of the mind is so largely undetermined that its contents can only be formed in personal history after the structures of the world: the basic principle of empiricism (which we can take as only partially valid, leaving a part of innate general structures and processes) must be substituted to the basic biological principle as concerns cognition Assumption that both structural indetermination and capacity of prograssive structuration is also a main property of the higher part of thr brain, which leads to a functionalistic view of cognitive science, seems to be more compatible with the facts.

Then, a functionalistic view of cognitive science, in particular as concerns representations, is far more acceptable that the dream of a neurobiological reductionism, which would replace and eliminate other sciences in this area. We will below examine problems of compatibility of symbolic representations with their neurobiological substrate.

USE OF PRESCRIPTIVE RULES

If we disregard material properties of representations, we are led to their other properties, which can be natural or rational.

It has been argued above that ordinary logics, its use in artificial systems of information processing and knowledge representation, are precisely based on the requirement to diminish variability of natural representations. Many logical enterprises in history were motivated by such a will to supersede variability of the individual beliefs by adopting regular systems of thought.

The means was adoption of prescriptive rules and their application to natural representations. These logical rules are often presented as rules of reasoning, but they often apply to representations themselves. They are of the general type: "check your spontaneous activities of information processing and automatically generated concepts, and, if necessary, restructure them, giving up the inappropiate ones, so as to comply to general principles, historically selected by successive communities of experts to govern rational thought".

The last decades have shown that these rules can be formalized not only for work "on paper", but also for material implementation and operationalization in computers: these make the informational structures of data, assertions and inferences materially manipulable through use of the physical phenomena taking place in the machine.

It seemed also easier in this case to firstly consider these prescriptive rules as "logical", and applied to machine deduction and "reasoning". However, development of "knowledge representation" as a related field shows that prescriptive rules need to be applied to the structuration of information as well as to its processing. Introduction of non-monotonic logics (McDermott, 1982; McDermott, Drew and Doyle 1980), whose role is to permit correction of derived knowledge when the early assertions or beliefs turn out to be inapproppiate, is one more illustration of this interaction between structure of information and reasoning "salva veritate".

DESCRIPTION OF NATURAL REPRESENTATIONS

We can now examine some properties of natural representations. For cognitive psychology, the best view seems to consider that the word "representation" in fact covers two distinct kinds of entities: token representations, which are cognitive events, and long term memory representations, which are cognitive structures. The type/token distinction is borrowed from linguistics, in which it primarily concerns words: the corresponding concepts can be cautiously applied to the present domain.

In linguistics, "token word" is perfectly clear and denotes a single occurrence of a given lexical unit in a context. "Type word" which seems to be immediately understandable, in fact is the product of a double generalization across the various occurrences: the first generalization is across all occurrences in a given text, and the second across all texts or speakers.

If we transpose this distinction to the domain of mental representations in cognitive psychology, we must keep in mind that we use several separate generalizations.

Token representations can also be claearly defined in this framework. They are single mental events taking place in the mind of an individual. A particular token representation, supposed to be individuated by its content $\mathbf{c_i}$, is any mental event in any individual during which this individual thinks of $\mathbf{c_i}$, or image $\mathbf{c_i}$, or speaks about $\mathbf{c_i}$ (i.e. produces speech about this representation, previously present in his or her mind), comprehends $\mathbf{c_i}$ (i.e. constructs this representation in mind as a result of audition of words and appropiate mental activity of comprehension), etc. Such a particular token representation can be written out as $\mathbf{r_i}$ $(\mathbf{c_i}, \mathbf{a_i}, \mathbf{p_i})$, in which $\mathbf{a_i}$ is an individual, and $\mathbf{p_i}$ the particular properties associated with this representation.

What is $\mathbf{c_i}$ in such a description? If we restrict ourselves to the easiest problem, that of representations that are verbalizable, we can use the same example as above, that of the content associated with the phrase *"the discovery of America by Christopher Columbus"*.

However, it is not very satifying for natural representations to have their individuation only based on their extension, as they were up to now. We certainly would prefer to rather individuate them by their "psychological content" or "intention". It seems that this can be best done, in a natural science framework, by resorting to the psychological process of recognition. It is a common behavioral evidence in subjects that they can have a "same" representation again and again: the criterium of "sameness" then clearly is that subjects recognize these events as tokens of a same personal type. Recognition thus is the basis for generalization in a given subject across individual token representations with the particular content $\mathbf{c_i}$. This content, taken as a variable **C**, of course is one of the proprties P mentioned above.

If we now want to generalize across individuals, and say that several persons have a type of representation that is "the same" with respect to its content, we cannot escape all the current philosophical problems concerning reference (i.e. relationship between **e** and **c**, in our notation). However, as we are not here primarily interested in this kind of problems, we will rely on the overall observation that in everyday life a close correlation exists between content and extension.

Our major concern is to caracterize a type mental representation as being a class, **R ($\mathbf{c_i}$, A, P)** -- supposed to correspond to **R (e, A, P)** -- of all mental events that are in all individuals about the content $\mathbf{c_i}$, and correspond to the event **e** this is exemplified by "THE representation of *the discovery of America by Christopher Columbus*", considered in all its realizations in all subjects. In this analysis, when we say: "THE concept of *flower*", we are exactly mentioning such a class. Thus we are now prepared to focus scientific research on the natural properties associated with particular (sub-classes of) representations in subjects.

Which kinds of such properties should we take into consideration? We already mentioned psychological properties as familiarity, accesibility from memory, Rosch-tipicality, etc., that are as a rule applied to type representations, and object of many researches.

We must separately consider other properties as belonging to token representations: for example those created by the context in which a representational content is used.

Let us illustrate this via an example: suppose that the possessor **s** of the representation "the discovery of America by Christopher Columbus" is a scholar attending a lecture about history of American Indians and their peculiarities in the end of the XVth century. Now suppose that s was just hearing a sentence as: "presumably some visitors came from Europa to the American continent before the discovery of America by Christopher Columbus in 1492". We can admit that the representational unit "*America*" will be highly activated in s mind (see below), whereas the unit "*Christopher Columbus*" will be rather less. By contrast, if s was attending a lecture about the life and history of Christopher Columbus himself, and hearing a sentence such as: "the discovery of America by Christopher Columbus was a great success for him, although ...", we can admit that the representational unit "*Christopher Columbus*" will be more activated, whereas that of "*America*" will be less. Experimental evidence, based on measure of response times, support this view (for example Denis and Le Ny, 1982; Le Ny, 1990).

A related important property of token representations must now be mentioned. It can be expressed by the following question: is a particular token representation a replication obtained from a generic matrix representation, or the unique result of a non reproducible construction generated from particular parts? The first alternative involve a "relistic" use of the idea of "type representation" that will be elaborated below.

We deliberately chose to illustrate this paper by a large representation, expressed in a complete phrase *"discovery of America by Christopher Columbus"*, rather than a simple conceptual word as *"flower"*, which could have introduced some unseen bias in our analysis. In addition we chose to use as the extension of this representation an event (discovery) rather than a class of objects (as *"flower"* or *"bird"* might have made). The reason of this choice was that we tried to avoid typical, hackneyed examples that are two often used in cognitive litterature.

Our example representation is clearly composed: linguistically, it is a description, as would be the meaning of "the discovery of the Desolation Islands by Kerguelen". It is initially constructed from parts, i.e. the nominalized predicate "discover-discovery" and its two arguments "Desolation Islands" and "Kerguelen". We will assume in the following that the representation involving America is highly familiar in most subjects, whereas the one about the Kerguelen Islands is very unfamiliar.

Our claim is then that these two representations have distinct natural properties when they are realized as tokens. The first has a counterpart in long term memory, so that any corresponding token of it is an **activation** of such a **long term memory representation,** whereas any token of the second is the result of a semantic **construction involving predication.** An alternative view, which we will not here discuss, is that all token representation are constructed in the same way.

TOKEN REPRESENTATIONS, LONG TERM MEMORY REPRESENTATIONS AND THE NOTION OF ACTIVATION

The view adopted above implies a distinction between "type representations", which is an abstract notion similar to "type word", and "natural basis of token representations", which involves psychological and, possibly, neurobiological assumptions about the "**reality of**" these representations.

On this basis, a modelization of the natural functioning of representations is made possible. The basic assumptions (see for example Anderson, 1983; Le Ny, 1989a, 1990) are:

1. there exist in long term memory lasting representations; the most typical of them are the generic lexico-conceptual representations as "*flower*" or "*bird*", but some others are not less interesting: these can have various forms in language, for example definite descriptions as "*the discovery of America by Christopher Columbus*". But some other may not necessarily involve constructive procedures: an example of this is, in France, the representation generated by the couple definite article + noun "le Débarquement" ("THE Landing"), meaning "the Landing of the Allied in Normandy in June of 1944"; such representations work in the same way as proper names, but their main property seems to be their high familiarity, which makes them easily accesible from memory; other, encyclopedical, knowledge can be lastingly stored with them;

2. a basic property, activability, and a basic process, activation, are associated with these long term memory representations; they make these representations permanently having a variable **level of activation,** which is created by information processing. Above a given level of activation, atoken representation is generated in working memory, which can thus have various levels of super-threshold activation; several behavioral measures (as processing times) can be used as indices of these changes and states of activation (Le Ny, 1990);

3. multiple relationships between representations, which can be modeled in the form of a semantic network, and spread of activation between representations as a function of these relationships are additional assumptions, which will not be elaborated here.

These three main assumptions define the family of models based on symbolic representation and activation, which are among the best used for explanation various cognitive processes, in particular language comprehension. As a matter of fact, these symbolic models are fully compatible with sub-symbolic models, based on parallel distributed processing and use of so called "neural networks" (MacClelland and Rumelhart, 1986; Rumelhart and MacClelland, 1986). The open question between these is whether the concept of symbolic representation is absolutely needed in cognitive science or not: our position in this text is clearly affirmative.

On the contrary, the activation process assumption here advocated is different from the computer-like information processing and working memory models, according to which processing is based on transfer from one memory store to another or several memory stores. We will not elaborate this second family of models either.

One advantage of the present view is that it makes modelization of construction of new complex token representations easy. This can best be illustrated in language comprehension, in reading a sentence, a paragraph or a text. Construction of meaning is then assumed to the based of multiple activations of several representations corresponding to several elementary words: each stimulus word brings about activation of its own meaning. i.e. the semantic representation associated with its mophological representation in long term memory. From this, connection of several meanings in an unique meaning procedes by predication and binding.

Predication is assumed to play a major role in this integration; this is a consequence of the so-called "propositional assumption", which we prefer to label "predicative assumption". According to this, elementary units (corresponding to words) are stored in memory in a form that make them counterparts of logical predicates: basically, these representations own a particular type of associativity with other units.

Let us take again as an example the meaning stored with the word "discovery". The predicative assumption states that this meaning involves: 1. knowledge that there is necessarily a "discovery of (x) by (y)"; 2. complementary knowledge of what x and y possibly denote, for example, for x a set of geographical places, and y a set of humans; 3. more or less clear knowledge of the probablistic distribution of various subsets of x and y, with their bonds.

Comprehension of larger units, as complete sentences, paragraphs or texts, may be modeled on this basis as establishment of other bonds between propositional or super-propositional units or contituents, corresponding to various kinds of schemata in memory (for example Yekovich and Walker, 1986, 1987). Similarly, units of meaning vcan be viewed as representational parts that are not necessarily attached to a whole morphological units in one-to-one way, but as well as comprising "smaller-than-the-world" components this is the basis for a componential view that would escape the criticisms developped against previous feature theories.

Descriptions of such semantics representations subsequent to comprehension can be experimentally tested in an indirect way by comparison with several kinds of empirical data, in particular those collected on response time in experimental standard situations, as self-paced reading, probing or priming. Present evidence is in a not too bad agreement with these models; more results must be collected in order to be sure that they are valid, and which particular model is the best. But internal compatibility, i.e. compatibility between models and facts is unboubtely growing.

CROSS-COMPATIBILITY WITH NEUROBIOLOGY AND ARTIFICIAL INTELLIGENCE

As the family of symbolic models based on the three notions of token representation, long term memory representation and activation has an acceptable validity inside cognitive psychology, the further problem is whether it can be also considered as compatible with concepts and knowledge based on facts collected by other methods, and having a high plausibility in other cognitive sciences. This question derives from the rule to circulate concepts and explanation across areas and methods. It is of particular concern for Neurobiology and Artificial Intelligence.

The idea of activation is itself borrowed from neurobilogy, although its second argument -- "activation of (what?)" -- has a deeply different semantics according as it is applied to the domains of neural or symbolic entities. However, the related idea of token representation can also easily find a neurobiological counterpart when viewed as an event. The psychological model of activation is congruent with everything we know about a category of neural events: change of a neural configuration from an inactive (or few active) to a more active state.

Generally speaking, activation of specific neural assemblies is today the best neurobiological candidate for a support of representational activity, even though neural events that can now be evidenced only concern too small or too large parts of the nervous system, and in no way somathing that could be identified as a specific and relevant neural assembly.

However, there exist experimental data (see for example Bloch and Laroche, 1986) showing concomitance between well specified neural events and definite behaviors, as a rule behavior of anticipation, permitting to infer presence of representations in animals. It is, and certainly will remain, more difficult to say of what these events may be representations. It is interesting to observe that the philosophical problem of "inescrutablity of reference" (Quine, 1961) here takes a new form.

There is more doubt about a possible counterpart of lasting representations. In some versions of parallel and distributed processing, in which the focus is put on "blind" processing, there is no place for such an idea, or even negation of it. But otherwise, it may be easily assumed that the reality of long term reoresentations resides in the sub-network structure of the connexions in a neural pattern. The working hypothesis we use here is thus compatible with the idea of such "neurosymbolic" patterns, which would be specific counterparts of cognitive lasting structures, meaning units or concepts.

In short, a fairly satisfying compatibility can be found between this faminly of psychological models and facts, and those of neurobiology. From this, two way constraints can continue to be imposed on these models by future facts.

A different kind of compatibilty must be looked for the Artificial Intelligence side. We said above that we do not believe it could be found directly between A.I. and neurobiology, at least as concerns high level phenomena: this compatibility must go through cognitive psychology and functional identities at a common level, the symbolic or "knowledge level" advocated by Newell (1982).

This can again be illustrated by problems concerning natural language processing, in particular those that are common to automatic and natural comprehension.

Most current systems of language interpretation have as their goal to construct, from information given as an input text, a final semantic (or syntactico-semantic) output machine representation. This must both: 1. permit subsequent processing, for example production of responses in question answering, summaries or indexing, performing a dialogue, etc.; 2. yield a "representation of the text's meaning" that is fairly acceptable by an expert.

As a final semantic, mental representation is also the result of human language comprehension, "fairly acceptable" in the previous sentence should be taken as referring to a description of this semantic human representation rather than a particular linguistic theory.

It could be shown in more detail than here possible that a set of similar functions must be accomplished, either by the machine automat or by the human mind, to attain this goal. Let us briefly mention:. 1. morphological identification of the successive various units (words) contained in the text; 2. access to (or activation of) the meaning associated with the types (roots and flexions) of these forms, taken into consideration, if necessary, the context of these units; 3. predication into the propositional format, as already mentiones above; 4. syntactic analysis; 5. identification and filling up of the semantic and grammatical schemata stored in long term memory (as frames, scripts, rules of word order, etc.); 6. storage in separate registers of small partial representations (semantic counterparts of constituent); 7. further binding of these meaning parts into larger parts, with progressive transformation of partial representations into a whole representation; 8. decay and partial loss of information, as a function of its importance; 9. use of reprocessing to modify representations and estimation of importance, etc.

The two lastly mentioned functions, which emphasize the role of working memory and selection of information in comprehension, are not often mentioned in litterature, but we consider them as very important (Le Ny, 1990a, b).

The ways these functions are realized in mind and machine are very different, but they can be considered as highly compatible. For example, instanciation is a high-level basic mechanism for language processing, of which activation can be considered to be a natural, whereas copy and adjunction of information is a form in computer.

In short, language comprehension can be viewed as a set of transformations applied to representations, having their end in a final comprehensive representation. A specific fact concerning human comprehension is that this final representation has the very particular property "conscious", whereas either intermediate human or machine representation have not.

CONCLUSION

We tried in this paper to show how an orientation can be taken in cognitive science, in which exchange between concepts, models and facts would permit to mutually enrich current knowledge. Certainly, very much remains to be done before we can construct a "hyper-general", i.e. a high-level, superordinate, description of representations and information processing. But this effort of circulation of ideas is both a necessary and possible goal in the quest of an integrated cognitive science. Anyway, a more complete elucidation of the general concepts in this large area, in particular of "representation", is a valuable means for elaboration of hypotheses and directions of research useful for empirical work.

REFERENCES

ANDERSON, J. (1983) - **The architecture of Cognition**, Cambridge, Mass., Harvard University Press.

BLOCH, V., LAROCHE, S. (1986) - "Facts and hypotheses related to the search for the engram". In G. Lynch, J. L. McGaugh and N. M. Weinberger (Eds.) **Neurobiology of Learning and Memory**, New York, Guilford, pp. 249-260.

CHANGEUX, J. P. (1983) - **L'homme neuronal**, Paris, Fayard.

DENIS, M. & LE NY, J. F. (1989a) - "Centering on figurative features during the comprehension of sentences decribing scenes", **Psychological Research**, 48, 145-152.

LE NY, J. F. (1989a) - **Science cognitive et compréhension du langage**, Paris, Presses Universitaires de France.

LE NY, J. F. (1989b) - "Automatic induction of concept contents from verbal characterizations". In G. Tiberghien (Ed.) **Advances in Cognitive Science**, vol. 2, Chichester, New York, Horwood-Wiley, pp. 165- 177.

LE NY, J.F. (1989c) - "The role of schemata in concept acquisition and diagnosis". In H, Maurer (Ed.) **Computer Assisted Learning** Berlin, Sringer Verlag, pp. 308-321.

LE NY, J. F. (1990a) - "Compréhension du langage, représentations sémantiques en mémoire de travail et processus de reconnaissance", **Communicazioni scientifiche di psicologia generale**, in press.

LE NY, J. F. (1990b) - "Coherence in semantic representations: text comprehension and acquisition of concepts". In G. Denhière & J. P. Rossi (Eds.) **Text and text processing**, Elseiver-North-Holland, in press.

McCLELLAND, J. L., RUMELHART, D. E. & the PDP Research Group (1986) - **Parallel Distributed Processing. Exploration in the Microestructure of Cognition, vol. 2 : Psychological and Biological Models**, Cambridge, Mass., MIT Press.

McDERMOTT, D. (1982) - "Non monotonic logic II - Non monotonic modal theories", **Journal of the ACM**, 29, 33-57.

McDERMOTT, D., DREW, V. & DOYLE, J (1980) - "Non monotonic logic I", **Artificial Intelligence**, 13, 41-72.

MAIDA, A. S., SHAPIRO, S. C. (1982) - "Intensional concepts in propositional semantic networks", **Cognitive Science**, 6, 291-330.

NEWELL, A. (1982) - "The knowledge level", **Artificial Intelligence**, 18, 87-127.

PAILLARD, J. (1987) - L'ordenateur et le cerveau: un contraste saisissant", **AFCET/Interfaces**, 57, 4-9.

QUINE, W. V. (1961) - **From a logical point of view**, Cambridge, Harvard University Press.

RUMELHART, D. E., McCLELLAND, J. L. & the PDP Research Group (1986) - **Parallel Distributed Processing. Explorations in the Microestructure of Cognition, vol. 1 : Foundations**, Cambridge, Mass., MIT Press.

YEKOVICH, F. R., WALKER, C. H. (1986) - "Retrieval of scripted concepts", **Journal of Memory and Language**, 25, 627-644.

YEKOVICH, F. R., WALKER, C. H. (1987) - "The activation and use of scripted knowledge in reading about routine activities". In B. K. Britton (Ed.) **Executive control processes in reading**, Hillsdale, Erlbaum.

***Institut des Sciences Cognitives et de la Communication**
Université de Paris-Sud.