P. R. MEDINA-MARTINS*

Emergence, Diachronism and Machines

It is argued that emergence [a term to be interpreted as dealing with 'the autonomous or operant responses that a system may exhibit' (Pask, 1975)] is *ultimately* a social phenomenon. As such, in addition to the descriptive and temporal 'dimensions' pertaining to the classical mechanistic framework, emergence also requires as a mandatory condition a 'dimension' of the attribution of meaning. The inclusion of this last condition implies, in turn, a reappraisal of most of the approaches on which emergence as been based so far, as well as further extensions of the concept itself. Amongst these, the distinction between 'outer' and 'inner' emergence, and a possible overall 'mechanism' justifying these types of emergence in human beings and possibly machines, are especially striking.

Key-words : Emergence, inner and outer emergence, Conversation Theory, descriptive and interpretative social/normative/relativist framework, machines, autonomy.

Emergence, diachronicité et machines. On défend ici la thèse que l'émergence [un terme à interpréter comme ayant à faire avec "les réponses autonomes ou opérantes dont un système est capable" (Pask, 75)] est, en dernière instance, un phénomène social. En tant que tel, en plus des "dimensions" temporelles et descriptives relatives au cadre mécaniste classique, l'émergence requiert aussi à titre de condition nécessaire une dimension d'attribution de sens. L'inclusion de cette dernière condition implique, à son tour, une réévaluation de la plupart des approches sur lesquelles l'émergence a été fondée jusqu'ici, ainsi que des approfondissements du concept lui-même. Parmi ces derniers, la distinction entre émergence extérieure et émergence interne ainsi que l'hypothèse d'un possible mécanisme global justifiant ces types d'émergence chez les êtres humains et peut-être les machines, sont tout particulièrement saisissants.

Mots-clés : émergence, émergence extérieure et émergence interne, théorie de la conversation, machines, autonomie.

Dedicated to the memory of Gordon Pask

* *Universidade Técnica de Lisboa*
Priv. Address *: R. Heliodoro Salgado 61, 2º -* 1170 Lisbon, Portugal
E-mail : p.medinamartins@mail.telepac.pt

1 - THE QUESTION OF EMERGENCE : AN OVERVIEW

1/1.1 - The question of emergence has recently become the subject of renewed interest, both from a formal logical or mathematical perspective, and from a point of view seeking its natural or artificial reproduction. Many factors have contributed to this situation, but amongst them the enhancement that computer technologies have provided to many of the contemporary disciplines has been especially important.

In this paper two such bodies of knowledge play a special role. Firstly, *biology* in general, the area *par excellence* from which the concept itself and its cognate problems came forth. Secondly, the importance which has progressively been ascribed to all the disciplines having *knowledge processes* as an *autonomous* (albeit quite general) host of research objectives, particularly those in which the 'knowing of knowing' has acquired a crucial relevance. In both cases, indeed — and beyond everything which straddles such different areas — there is an intimate underlying connection : they are both devoted to the study of the human being although looking upon him/her from different perspectives. And here, in this difference, lies one of the main problems that those to whom the question of emergence is the main focus of attention have to face. What *ultimately* underlies each of such viewpoints is a particular 'way of thinking', perpetuated from generation to generation since Descartes' epoch until quite recently. One of these viewpoints, that hereafter, for the sake of brevity, I shall name the *classical paradigm,* will be especially important in the further development of this work. This viewpoint has dominated, practically without opposition, the whole development of physical or 'hard' sciences ; by extension, it also aims at encompassing the evolution of the so-called 'soft' or humanistic disciplines.

I shall not dispute the fruitfulness of this classical paradigm in the realm of the physical sciences and cognitive disciplines (of which computational and neuro-sciences in general are illustrative examples) ; nor even in the understanding, simulation, and reproduction of *some* mental phenomena, particularly those pertaining to what Jackendorff (1987) calls 'the computational mind'. But, in my opinion, this supposedly unifying viewpoint tells us only a rather restricted part of the whole story. Human beings are far more complex and puzzling creatures than 'simple' serial or parallel logic machines. In particular, in addition to a brain and a "computational mind", each of us also has a body. And, since the paradigmatic model of the brain as an enormous

mass of neurons-like on/off electronic switches controlling the body has already been proved to be erroneous in the light of the most recent advances of neurobiology, the relationships between bodies and brains and, in turn, of both of them with minds, continue to be unsolved problems. As a further issue, the usual emphasis laid upon rational and cognitive processes *alone,* so characteristic of the classical approach, either excludes emotionality and "irrationality" almost by edict, or seeks their explanation solely in logical terms (as in some recent "cognitive" psychotherapies). This introduces a false and rather artificial division into the mind's functioning that is at variance with what is now known about 'normal' and pathological processes. Even the modelling of logical reasoning — one of the strongest bastions of AI and of mind-like simulations — undergoes severe criticisms when compared with the actual processes we use to get some apparently 'neat' conclusion from previously known (and also supposedly 'neat') premises. Indeed, the whole question of the coherence of our thoughts and utterances is beset with profoundly obscure problems. Thus, many of the terms, concepts and rules of inference that are used whenever we engage in conversation with other people (or even with ourselves) are far from being as neat and objective as the traditional logicians and their orthodox followers claim. On the contrary, they are essentially vague, they usually encompass several types of time-variable self-referentials ; consequently, they also entail a large variety of pragmatic, subjective and imprecise appraisals in which the individual's life experiences, his/her idiosyncrasies, etc., briefly, everything determining his/her *uniqueness,* play a role that cannot be neglected.

This vagueness, this subjectivity, one's individuality and uniqueness, bring sharply into focus the gap which still exists between the high standards of precision that prevail in mathematics, physics, engineering, biology and so forth on the one hand ; and on the other, the imprecision which pervades much of psychology, sociology, political sciences, history, philosophy, education, art and so on. Various reasons justify the sharp contrast between these two standards of precision, some of which are depicted in Table I. A prominent role is undoubtedly played by the deeply entrenched tradition that has framed the precise, exact and *quantitative reasoning* upon which the "hard" disciplines have been based for centuries. The possibility of quantification establishes a clear difference with respect to the "soft" disciplines centred on systems intrinsically associated with the complexity of the human *qua sentient* being, where *qualitative factors* cannot be overlooked or disavowed. These qualitative factors include values, norms, judgements, emotions ; unforgettable past influences which sometimes radically modify our

perceptions about the 'here and now', forcing us to react as though we were 'there and before', or even changing our possible futures.

Table I

'HARD' SCIENCES	'SOFT' SCIENCES
1) Neat separation 'Object'/'Environment'.	1) No sharp separation between 'Object' and 'Environment'.
2) All the 'objects' (human beings included) are regarded as "it" under scrutiny.	2) Human beings are mandatorily looked upon as human *qua sentient* beings.
3) 'Object' and 'Environment' are regarded as classical automata.	3) 'Object' and 'Environment' can, by no means, be looked upon as classical automata.
4) The interaction 'Object'/'Environment' is supposed to be framed by causal (efficient) rules. So, teleology is excluded almost by edict.	4) Although (efficient) causality cannot be overlooked, teleology and goal-directness are crucially important.
5) The interaction 'Object'/'Environment' is regarded as an exchange of symbols in some formal language.	5) Interactions amongst human beings are primarily based upon natural language.
6) This language is a syntactic entity only. So, semantics and pragmatics are excluded.	6) Semantics and pragmatics are crucially important.
7) The observer's attention is devoted to 'outer' event-processes. Furthermore : these 'observers' are, usually, supposed to be - neutral - objective - often omniscient	7) Attention is devoted to actions performed by human actors. These actors work, therefore, more in the sense of "participants in..." than in the classical sense of "observers of...".

8) Observers search for 'laws' ('natural laws') which, once discovered and experimentally confirmed (by means of successive repetitions) are
- integrally extended to the past ;
- assumed to be
  . objective,
  . based upon numerical magnitudes,
- referred to logically open classes ; so, 'universal'.
	8) These actions have, therefore, an 'external' and 'internal' parts, the latter dealing with the affective and cognitive thought-processes of those actors.
	So, their reconstitution is primarily concerned with
	- judgements ;
	- written/uttered in a natural language,
	in which emphasis is laid upon
	  . singularity,
	  . idiosyncrasy,
	  . singularity,
	  . uniqueness,
	- referred to logically closed classes ; so, applicable not to humans in general but to the man or woman who has performed this or that particular action

9) 'Natural laws' are usually described by differential/ difference/integro-differential equations or by automata theoretical approaches (deterministic or statistical). If integrable (or if statistically describable) the system that these equations model, allow therefore the setting up of prospections, forecasts and/or predictions.	9) Besides the impossibility of ascribing a precise numerical character to many or all of their describing variables, since neither the constancy of the 'object' under scrutiny nor the 'constancy of environmental conditions' are obeyed then, no individual predicting method (in the traditional sense) has been available so far.

1/1.2 - Generally speaking, owing to the

enormous influence that the mechanist or classical paradigm continues to exert on the contemporary scientific scene, many of its underlying technical procedures continue stubbornly to be applied to the analysis of human systems. In spite of this, it has also becoming increasingly clear that the complexity, the capability for autonomous and emergent behaviours, and the particular characteristics and specific problems that human systems exhibit, call for approaches that, both in spirit and in substance, have to be significantly different from the traditional methods.

In this paper — essentially devoted to the problem of emergence — I will lay special emphasis upon one of such non-conventional approaches: Pask's Conversation Theory (CT for short) albeit already modified according to my own research perspectives. As a matter of fact, Pask has created one of the few cogent alternatives to the poverty that most of AI's approaches to the emergence question exhibit, particularly when either the individual's previous life experiences, his/her idiosyncrasies, uniqueness, living experiences, etc., or his/her interactions with other individuals (and/or non-natural systems), are being questioned. From my perspective, however, the original version of CT suffers from severe limitations. It neither provides satisfactory answers to many of the aspects that our individual and social functioning encompasses throughout time (particularly those to which our emotional ties are referred) ; nor, *by itself alone,* is it able to ascribe that so desired *self-autonomy* and consequently *emergence* that I have long been endeavouring to assign to non-natural systems.

1/1.3 - The objectives of this paper become therefore twofold. On the one hand, to survey some of Pask's framing ideas as suitable clues for providing solutions to the problem of emergence, in particular when this emergence is taken in his sense of *psychic*, operant or autonomous behaviours (Pask, 1975). On the other hand, to provide an extremely condensed description of the way my neuro-fuzzy approaches to CT *may* overcome its limitations, particularly in the realm of emergent phenomena *in human beings*. Of course, by laying emphasis upon *psychic* and *human beings* it seems that I am excluding from analysis the areas of biological and neuronal mechanisms, which are the usual objective of research whenever emergent phenomena are under scrutiny. Nothing however could be further from my mind. Indeed, by paying attention to the preconditions that, conscious or unconscious, have begotten Archimedes' "Eureka!", or common utterances as "I had never thought of that!", I am, on the contrary stressing an aspect that somehow has been overlooked by those to whom emergence itself is

the main focus of attention : *that if no sort of meaning is assigned to what is regarded as an emergent phenomenon, then any discussion*

about it (as well as about any other conversational topic) becomes senseless. Furthermore : since such a meaning-assignment is, ultimately, a social phenomenon, then it follows that emergences (even those pertaining to the physical or quasi-physical domains) also become a matter of social interaction.

This is the thesis that pervades the whole of this work ; to support this contention will the major objective of the following considerations.

2 - EMERGENCE, MEANINGS AND CONVERSATION THEORY : A SURVEY

2/1 - 'Emergence' and some related problems

Amongst the questions that my preceding statement has certainly aroused in the reader's mind one in particular must, at this point of my exposition, be especially poignant. If the concept of emergence is ultimately assumed to be a social matter — in the sense that it depends on the perspectives of its observers/interpreters — then when can we speak of an actual, really observed 'emergence', about which nobody has doubt ?

The answer to this question depends on two prior conditions. Firstly, to the meaning according to which 'emergence' itself will be used in this paper. Secondly, to the importance ascribed to such a meaning-assignment in terms of a general theory of emergence, importance which precisely lays emphasis upon Pask's CT and its further extensions.

2/1.1 - Let us pay some attention to the first of these problems, i.e. to the meaning according to which emergence will be used henceforth. To this end, stated baldly, the first association which most of us makes whenever emergence is a matter of discussion is its identification with something *new,* that *comes out* or *appears from being hidden.* Especially poignant is the relationship set up between 'emergence' and 'new' or 'novel', a viewpoint shared by all whose attention is devoted to the subject. However, 'new' is an exceedingly dangerous term whenever uttered without adequate precautions. On the one hand, although the concept of emergence demands that something 'new' comes to light, the converse is far from being true. Indeed, 'new' data, facts, utterances and so forth occur everyday with each of us, but they do not *necessarily*

correspond to some kind of 'emergence'. The reader may consider, for instance, the stream of 'new' data we are always receiving at every moment of our lives from our habitual surrounding objects (both outer and inner); despite their actual and permanent 'novelty', no one would think of assigning any sort of 'emergent' attribute to such data. Sometimes, however, it may happen that such data *do* beget situations to which the term 'emergent' is undeniably applied. Where, therefore, does the difference lie ?

At the risk of being misunderstood and misinterpreted, I shall venture that three further attributes must be added to the aforementioned 'newness' in order that an actual emergence may be said to have arisen. Firstly, that we can positively assert that the data are not only "new" but also *unexpected*, *not foreseen*, *not (yet) predictable* or (in the realm of machines in general) *not previously programmed*. Secondly, that once they appear, they become thenceforward an integral part of the patrimony of the system to which they belong. In other words: that they acquire an *irreversible* character. Finally — and this is a mandatory condition — that a *novel meaning* not yet existing in the mental repertory of the participant observers has to be assigned to them.

2/1.2 - Many examples in which these attributes yield emergent situations can be pointed out. The growth of Science (a term which I am using here in the broad sense of any systematic, whole-embracing or *Weltanschauung* construct of the external and internal worlds) by means of new mental and/or technological 'discoveries' ; the process underlying all the artistic creations (painting, musical composing, sculpting, writing, etc.) ; the successive changes of perspective both of ourselves and of the 'external' world that all of those engaged in a psychoanalytic process undergo (yielding new insights, e.g., successive re-formulations of each one of our inner and outer contents that, in turn, give rise to new ways of looking upon those worlds together with the abandonment of older viewpoints) — can all be regarded as examples of emergences as defined here. In each case there is something new, not expected or foreseen ; something meaningful ; and something irreversible (in the previous sense that, once brought about or *realised* by someone, such a 'something' can *thenceforward* be looked upon as an *acquired change* either by the system, or by its observer/participant/interpreter, or by both).

2/1.3 - I have purposely laid emphasis on the terms 'meaningful', 'realised', 'acquired change' and 'observer/participant' since each one of them will allow us to excavate a little more some of the questions that the concept of emergence entails.

Let us begin with 'realised', 'acquired change' and 'observer/participant'. Suppose, at the outset, that I am following the classical mechanistic distinction between *system under scrutiny* and *environment*. Suppose now that this environment comprises *several* observers, not just a single one as the mechanistic framework traditionally considers. Then, the possibility these observers have of stating that their scrutinised object exhibits some kind of emergent behaviour depends on various conditions which are not usually taken into account in the analysis of emergence. To begin with, the system under scrutiny has to evince behaviours not yet foreseen or not yet predictable on the basis of those already known. Moreover, the observers have to realise that some alterations have arisen (not randomly but in a *definite* way). Finally, the observers have to realise that these alterations have somehow become an integral part of the present and future attributes of the system under scrutiny, and do not derive from any one of its observers isolation. The reason why I have been speaking of *several* observers, and not of a single one as in the traditional analysis, is to exclude individual illusions and to emphasize the irreversible character of the emergence.

Asserting this (which seems to be a bare triviality) I am however implicitly saying that each one of such observers must have been engaged in conversation with the others (hence their 'participant' feature), *sharing* and (eventually) *agreeing* that 'something' in the system has, indeed, actually changed during some observation. Obviously, this system can neither be undergoing rapid and unexpected random changes (otherwise it would very soon present no similarities at all with the original prototype, i.e. it would suddenly exhibit so distinct a character that ultimately the observer would be forced to assert that it had become *another* object) ; nor can those alterations take so long a period that, eventually, they may give rise to the idea that the system is time-invariant — a situation in which no detectable emergence may be said to exist. Similar arguments also hold for the *amplitude* of such changes which can neither be so small that they remain undetectable, nor so great that the limits either of the sense organs or of the scientific instruments such observers are using, may run the risk of being exceeded. In other words, there must exist a descriptive *space-time window* (the word 'space' being here used in a rather broad sense, albeit connoted with Lanczos' *configuration space*) within which

    i) the existence of the foregoing changes may be detected and described,

ii) apart from such modifications (allowing the observer to state that an emergence came forth) there must exist some kind of *similarity relation* in the object *before* and *after* the rising up of the features that have justified the observer's assertion that an emergence has been witnessed.

Let us then name hereafter *individual relativism* this particular relationship between each human observer and the class of systems he/she can be aware of. By the same token, let us call *public* or *social relativism* the (possible) relationship that a group of observer-participants *agree mutually to share* about some class of systems of which they are aware.

2/1.4 - The limits of *each one* of such 'windows' are obviously *private* since they depend not only on the range of the sensory receptors which each observer is endowed with, but also on his/her psychic constitution (here encompassing past experiences, the way they have been interpreted, etc.). Three immediate results can therefore be brought to light from this fact.

2/1.4.1 - Firstly, that even though some system is by itself providing data for its environment, if these data lie outside the limits framing the specific descriptive 'window' of this or that particular observer, then he/she will be unaware either of its cognate occurring processes or even of the object itself, everything working as though it were non-existent for the observer. However — and this an aspect the reader must carefully bear in mind — this individual ''non-existence'' is not solely restricted to physiological boundaries. *On the contrary, if for some reason the observer is incapable of assigning a meaning to the object he/she is attempting to scrutinise, then this object also remains opaque and impenetrable and it might as well not exist.*

If we look at the preceding considerations then we can see that two major subject matters are entailed in the concept of emergence.

The first one deals with the manifold descriptive 'dimensions' that every *individual* 'space-time window' may have. Such 'dimensions' enfold therefore the numerical variables that some classical observer has deemed relevant for the description of the system under scrutiny ; they also include elementary words which, after being somehow combined, concatenated and related, form a phrase or expression which may be used to make descriptions of the changes the system has undergone throughout time. The second (metaphorical) 'dimensions' deals with the eventual meanings the observer-
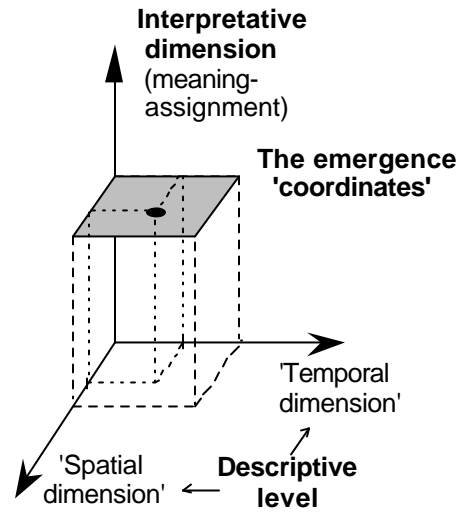
The three major 'dimensions' that the concept of emergence enfolds

Fig. 1

interpreters must have assigned to the foregoing changes. The intersection of some non-null descriptive level with the 'dimension' encompassing the possible meanings assigned to such a description defines the *interpretative level* ascribed to it. Together with the foregoing descriptive 'dimensions' ('space' and time included) these interpretations determine the 'reference-frame' according to which emergence will be interpreted thereafter. A rather symbolic representation of these 'frames' is depicted in Fig. 1.

2/1.4.2 - The second result is, to a great extent, an extension of these concepts. As a matter of fact, since physiological and psychological differences always exist amongst individuals (thenceforth called *M-individuals* according to Pask's terminology), it may happen that what for some observer, say A, is an actual datum can purely and simply be something non-existent for another observer, say B (for example, let the reader try to imagine the impossibility of explaining Monet's "Impressions du Soleil Levant" to a person born blind). Hence — and this is the third result I had in mind — each observer/ interpreter may or may not relate the object into their individual scheme of things in the same way. In other words : *that what we have called the 'same' material system, as well as its consequent emergent attributes and behaviours, can be incorporated into the affective and cognitive systems of the different observers/interpreters in different ways, with*

different relations and also with different meanings. Or, more briefly : the 'same' material object may not only be a different object for each one of its observers/ interpreters, but in addition its (possible) cognate emergences may or may not exist *simultaneously* for all of its observers/interpreters.

2/1.4.3 - The emphasis laid upon 'simultaneously' is here crucially important. For, if what one observer (the reader, or myself) regards as an 'emergence' is something already known for others, with an already socially accepted meaning-assignment, then how can we interpret, understand and accept our mutual descriptions ?

The answer to this question will precisely lead us to survey some of the standpoints of Pask's CT.

2/2 - Conversation Theory : a terse summary of its original (theoretical) version

Some of the tenets of CT have already been identified in the foregoing considerations. Another part, however, is condensed in Table II.

Two major columns are there inserted, of which the left-side includes some of the philosophical foundations upon which the great majority of the features pertaining to the 'hard sciences' of Table I rest. For the sake of simplicity, all of such foundations were inserted into the general label "Mechanistic Paradigm" although, from that philosophical perspective, they deal rather with the so-called "realism", a viewpoint that for centuries has framed (and still frames) the broad mainstream of science.

Table II

Because it is well known, only some of the characteristics of this point of view were there brought to light. On the other hand, the right side of Table II lays emphasis upon Pask's "Normative and Relativist Paradigm", a novel and fresh way of looking upon the world (at the time of its inception, of course), particularly useful as a means of overcoming the limitations of the classical paradigm in its endeavoured (but failing) extension to the psychological and social areas. The main features of such a paradigm are just those on which Conversation Theory has been based. In the sequel I shall survey some of them, in particular those which are more directly related to the question of emergence.

2/2.1 - These featuress of CT deal with the convergence of two major hosts of ideas which, at a first glance, seem to be bare trivialities. One is a philosophical approach — the "idealistic perspective" — which, because it is opposed to the aforementioned realistic attitude has been practically excluded from the entire body of contemporary science, 'hard' science in particular. The second idea — also unconventional when compared with the realistic features — encompasses two main facts : on the one hand, *nobody has a way of getting into the mind of anyone else* ; on the other, due to our specific genetic patrimony, to the life experiences we underwent, to the manner according to which we have interpreted them, etc., *each one of us is an undeniable idiosyncratic, personalised and unique entity who, altogether regarded, is different from all the remaining individuals of which humankind is composed*. This biological entity is just what was beforehand named an *M-individual*.

The convergence of these perspectives begets various ideas amongst which the following are especially important for an understanding of CT :

a) Firstly, observers are connected to what is naively called 'the world' by *their own observing acts*. Observers help to create and shape what is observed, so that observation is always *self* or *other-referenced*, never *it*-referenced ; this implies that we cannot apprehend the *meaning* of something through an external, 'out there', *absolute* and *pre-existing referent* but, on the contrary, by relating it to a context into which this something is inserted or within which it may be particularly regarded. This not only opens the way for the insertion of *subjectivity* into the observer's description, but also provides a radical alternative to the way of looking at knowledge acquisition and knowledge representation that

is prevalent in the "realist" approach, notably the expert systems of classical AI.

b) As a matter of fact, for the defenders of the realist approach, what is called the 'outer world' is, since the XVIIIth century (with Descartes, Galileo, Newton, etc.) an assemblage of forms reducible to more or less complex arrangements of non-living 'material particles', the existence of which — and this is the first aspect that the reader must bear in mind — is *independent of the observational process*. Knowledge of this world is, consequently, looked upon as some kind of commodity *logically distinct* from its knower. This means that its articulation may be expressed in *any convenient language*, and also, having being expressed, that it may be purveyed, transported, learned, taught and manipulated to some end. Therefore, knowledge consists in a collection of facts that are true of the world they represent. Thus, their *truth* or *falsehood* can be *objectively* determined and their place in a representation depends on whether or not they *correspond to* some object, state or relationship that exists in the world they represent. Furthermore, since the world is, by nature, a complex entity that is reducible into more and more simple units, and finally into fundamental elements, so knowledge of the world is equally representable as non-overlapping hierarchies of concepts where the place of each concept is fixed in relation to the others. Therefore, for those who follow the realist ideas, the problem of knowledge representation consists in finding a representational scheme that can capture the way the world *is*, in terms of facts that are true of it (together with the relationships between its facts) ; and in specifying a knowledge utilisation engine that can interpret those facts, so that conclusions that are *consistent with* and *true in* the world may be reached ;

c) One of the most remarkable problems that have to be solved by this formulation, occurs in the knowledge acquisition phase by an expert, where the opportunities for misrepresentations and misinterpretations are countless. In the realm of AI, where the realistic view has been prominent, these (possible) misinterpretations, misrepresentations, etc., are justified in terms either of the representational language used ("is a clumsy medium for expressing expertise") or because the knowledge engineer has difficulties in understanding the subject matter. But, whether or not representational problems of this type may exist, what is *never* disputed is the cogency of the epistemological attitude regarding how knowledge of the world may be acquired and represented ;

d) This criticism leads me to the idealistic current. For those who have defended this attitude (inheritors of the Greek Pythagorean school and developed since the XVIIIth century by Berkeley, Hume, Kant, Hegel and followers), the existence of a material world as the realist attitude claims is also undeniable except in an extreme Berkeleyan perspective. However, instead of directing their attention to the material 'stuff' of which that world is (eventually) made, the idealists lay, rather, emphasis upon the *meaning-assignment* in that world. The strongest argument supporting this thesis is that an object may be in the (material) world but, unless it has a meaning for its observer, it remains opaque and impenetrable to him/her, everything happening as though the object did not exist. Contrariwise, once some meaning has been assigned to the object (including its conceptualisation as an object of mystery), it becomes part of the world of meaning *for that observer;* obviously, this observer is not a classical one but, rather, a *subjective* observer of that object, this meaning that other interpreters may or may not relate the object into their 'scheme of things' in the same way ;

e) One of the most remarkable consequences of this attitude is that knowledge is regarded not as a reflection of the way the world 'really' is, say, a collection of independent bodies of facts, but, rather, as *connections that are created or constructed by knowers* to attach themselves to the world they live in. It follows, that a true/false Boolean logic becomes inappropriate for describing the idealist position. Instead, what is required is a *coherence theory of truth* (based on a logic of coherence, distinction and analogy), by means of which we can model our construction, not of *truth* but, rather, of *our coherent schemes about reality*. What is 'true' is what we can *agree on* at any particular time and place. Thus,'objectivity' in idealistic terms becomes a matter of *social agreement*. What is objectively so, is what we agree to be objectively so. If researchers see the world in the same way, it is because they have similar interests, values, purposes, motives, methods, and so forth. Agreement rests, not on the duplication of results, but on a commonality of perspectives which, in turn, produces similar results. As asserted by Gregory and Pask (1986) from whom some of the preceding considerations have been almost textually reproduced :

> In the macro world of objects (as well as in the micro world of quantum physics where the idealist position has been gaining more and more currency), idealism enables us to perceive and construe the objects around us in any of many different ways. But, doing so, we are also prohibiting or, at least, limiting (for the duration of our construction) other ways of construing. Furthermore, the constructions we do place upon the objects of our attention, undergo the 'tacit "agreement" of those objects if we are to trade any meaning with/through them that may be of value. I may, for example, interpret an object as being a desk. Having done this, my thoughts are

constrained by the notion I have applied to the object, so that I may now think of it as a surface for papers or as a set of containers for writing accessories; however, I am unlikely to think of it any more, as the host of other meanings for which I was previously free to assign to it. The success in continuing to treat the object as a desk and the reason why I assigned the desk-meaning to the object in the first place is because I expect the object itself to be capable of supporting the notion of "desk". It is not that the object *is* a desk in absolute terms but, rather, that the way it defines itself supports the notion of "desk" in a way that it does not support the notions of "flying machine" or anything else. Briefly : I do not create the existence of objects in the world but I am free to create meaning for them, this in such a way that their existence may or may not be able to support it. If they do support such a meaning (established by practical interaction and tested with them) then they contribute to the coherency of the world that we construct and *share* each other.

f) The emphasis here laid on 'share' is especially important for it brings to light another aspect of the idealistic attitude which, now, stresses the crucial role that language plays. Indeed, since knowledge and truth are not looked upon as *objective realities* but, rather, as *matters for agreement* within some universe of discourse (which does not exclude 'agreements to disagree'), then the importance of an appropriate language (not necessarily verbal) and the conversational aspect of the interaction amongst individuals become crucial aspects of the idealist position. As a matter of fact, knowledge is not merely articulated in that language; it is *allowed* by the language of its expression as well as being the *medium* through which two individuals interact by sending and receiving messages. These messages are not however regarded according to the traditional perspective (in which they are often viewed as telegrams that transmit or transfer information in some coding manner) but, rather, as *'productive/reproductive constructions'* that are externalised by each of us in our utterances when using them ;

g) This perspective is crucially important in terms of *knowledge-representation*. As a matter of fact, instead of 'objective truth', what these representations seek are consistencies. Thus, assuming that this consistency is compromised for some reason, then some representational process by means of which inconsistencies are detected and resolved, must also be taken into account. This resolution is based on a process of agreement that, in turn, depends on the perspectives of the knowers. Furthermore, as knowledge is not a reflection of the way the world is, its representation is not presumed to be capturable as a fixed hierarchy of nodes and relations. Rather, knowledge must be minimally representable as a heterarchy in which nodes may play more or less primitive roles with respect to other nodes, dependent on the particular perspectives, circumstances, and purposes pertaining when they are attended to. Therefore, the problem

of knowledge representation for an idealist consists in finding a means of representing the *process of understanding* and its outcome, i.e. what is *understood.* Representations can only proceed by a process in which the participants *share* the meanings that each participant attributes to the situation. This 'sharing through' (entailing conversational transactions in Pask's sense) results in a representation that is *true for both*, i.e. it is *coherent.* However, as a process, there is a difference between the situation they were both in at the start and at the end of the dialogue. This is because, when any conversation is taking place, it results in an exchange of concepts that are continually changed in the exchanging process. In this way, in addition to the representation of the shared meanings, any knowledge representation system created from the idealistic perspective must also take into account how the data structure changes by the entry of another perspective. Briefly : it must also represent *the process by means of which meanings become shared*. Or, more correctly, because of the idiosyncratic nature of each observer's point of view, *interpreted* meanings become shared. Indeed, as Glanville and Gregory (1986) assert :

> "We cannot share meanings since we have no way of really getting inside each other's head". In consequence "what each of us does is to build models of what the other means and re-iterate these models as being presentations of their understandings of the understandings of understandings.....in a (theoretical) eternally building regress". Practically, however, the (conversational) process terminates when the participants engaged in conversation (about some topic or domain within which they are exchanging the meanings that each concept has for them) "get to the point where each has an understanding of the domain not only in their own terms and from the original perspectives but also in the other's terms and from the other's perspectives. But, doing so, the information that was originated with each becomes forever changed. The participants *do not adapt to each other* ; they *change each other* into different individuals who, nevertheless, retain their original identities".
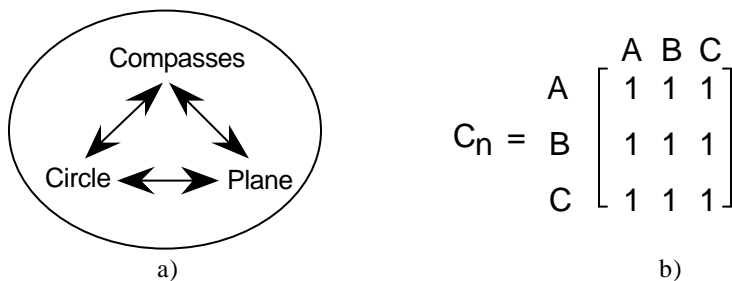
h) In order to understand the importance of this aspect let us take two classical examples extracted from a virtual machine named THOUGHTSTICKER that Pask and some of his followers constructed in the 1980s. To begin with, let us imagine that one person (Adam) is engaged in a dialogue with another person (Eve) about what his concept of 'Circle' is. Suppose, for instance, that the way according to which Adam represents his knowledge of 'Circle' takes the aspect shown in Fig. 2 a), meaning that he can *regenerate* his concept of 'Circle' by mentally combining his ideas of a pair of Compasses and a Plane. In these conditions there is no reason for Adam to deny the cogency of the two other cognate representations (correctly, permutations) depicted in Fig. 2 b) and 2 c).
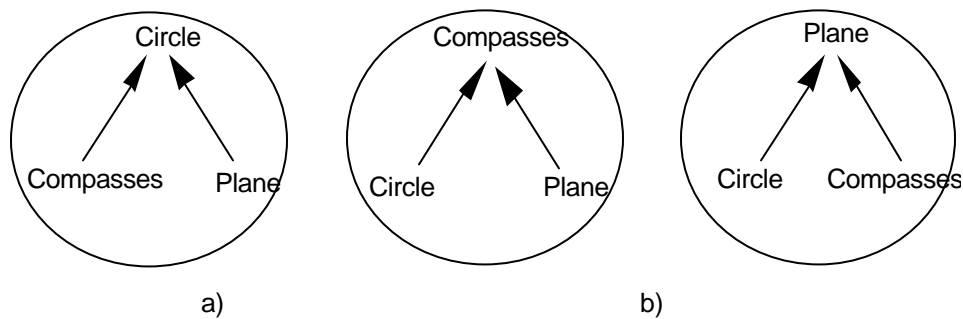
The setting up of cyclicity
Fig. 2

As a matter of fact Adam deems it possible for, if he knows what a Circle is and a Plane is, then he can get an idea of 'Compasses" as a device for inscribing a Circle on a Plane (Fig. 2b) ; a similar argument also holds for 'Plane' in Fig 2 c).

This property which is evinced by the bundle of concepts <Circle, Plane, Compass> (or any other bundle) is called *cyclicity* and entails the aforementioned notion of *coherency* (Fig. 3a).

$$C_n = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{ccc} A & B & C \\ \left[\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{array} \right] \end{array}$$

'Cyclicity' and its representation
Fig. 3

As a parenthetical commentary the reader must bear in mind that according to my own *psychological* interpretation of Pask's coherences (Martins, 1995a) such a cyclicity implies that the recollection of any of the concepts within a cluster yields the recollection of the remainder. Under this caveat, every coherence is not a set (according to the classical algebraic definition) but rather a *whole* or a *Gestalt* (in Wertheimer's sense), i.e. an overall relation in which, together with the *individual* concepts that the bundle itself entails, their mutual *relationships* must also be taken into account. This is tantamount to asserting that, if A, B and C are, for instance, individual concepts then the coherency itself is more adequately represented by a 3x3 matrix

$C_n= [a_{ij}]$, $(i,j=1,2,3)$ where the diagonal terms a_{ii} are 1, and the remainder elements a_{ij} (i p j) represent the aforementioned relationships (Fig. 3 b).

In order to continue this reasoning let us now imagine that Adam has already acquired several other ways of thinking about Circle, for instance, those shown in Fig. 4 where, for the sake of simplicity, the implicit derivational head arrows were not represented.

The interpretation which may be assigned to the picture is crucially important for an understanding of CT. As a matter of fact, what it means is that the concept 'Circle' is Adam's *repertoire* of the ways in which Circle can be understood. Or, in other words, that there is no 'absolute' definition of what Circle actually is but, rather, that this is the way Adam consistently and reliably thinks about Circles.

More precisely, that a concept is a stable repertoire of ways of knowing about that concept. Furthermore, for each of these ways (also called a perspective), the coherency requirement ensures that all the permutations obtain at the same time. In other words: that each topic/ concept in a coherent bundle is recalled or derived from a combination of the other terms. Thus, that concepts are remembered or derived from two, at least, other topics/ concepts; or consequently, that
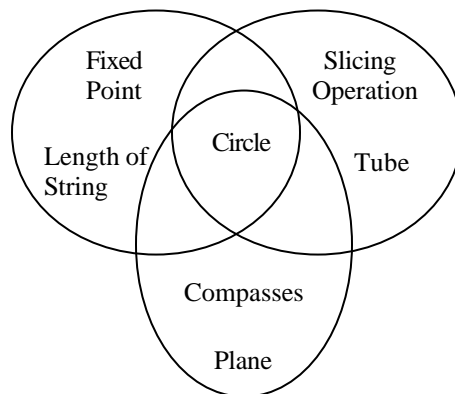
The various perspectives upon which 'circle' may be looked

Fig. 4

no piece of knowledge or concept is, in principle, isolated in our minds.

The second example brings to light the close relationships existing between CT and the (cognitive) relationships of machines/human beings from the point of view of that approach. Part of this mesh is represented in Fig. 5. Suppose, indeed, that we begin with the assemblage of coherences that a human being or machine has in his/her/it mental repertoire, named an *entailment-mesh* in Pask's terminology. Suppose also that A is engaged in conversation with another M-individual, say B, about, for example, the concept of 'bench'.
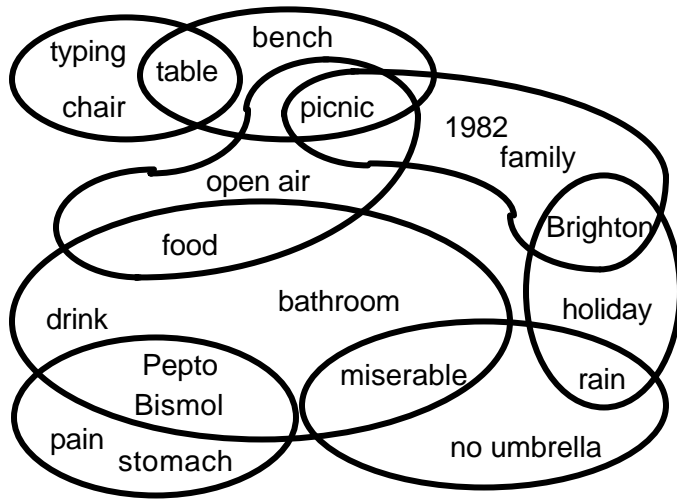
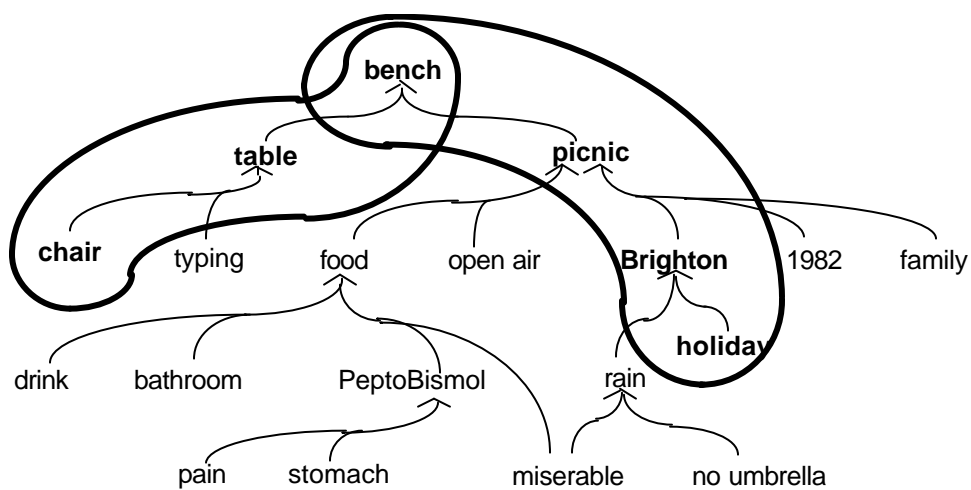Part of the entailment mesh of an M-individual A

Fig. 5

This is tantamount to asserting that 'bench' is the main conversational topic for A and B or (in A's case) that the selection of this topic concept is computationally translated by the use of a modal operator (called *Prune* in Pask's proto-language Lp). The result of this operation in A's mesh is represented in Fig. 6.
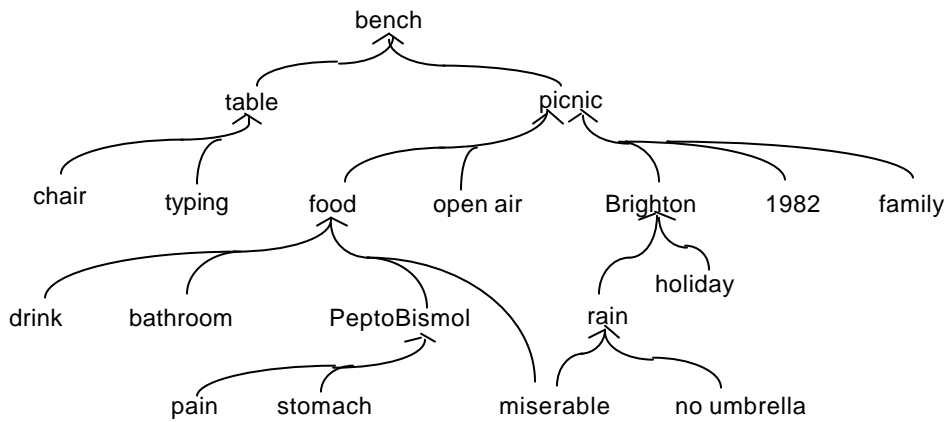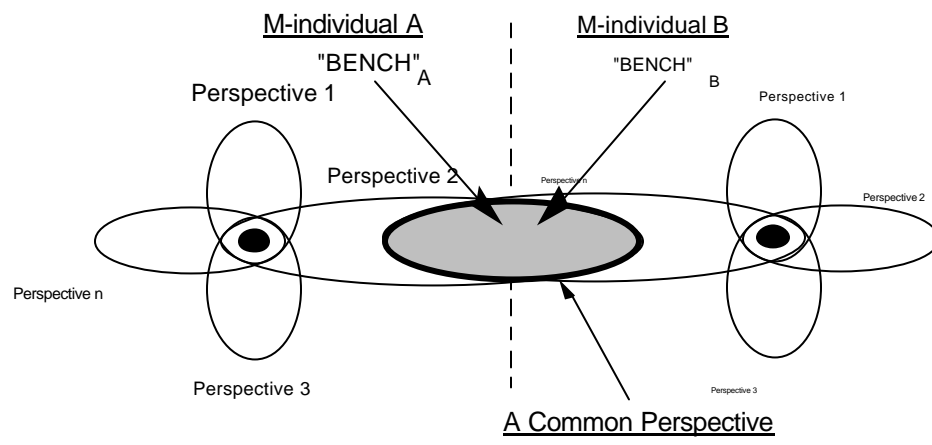

Pruning under 'bench'

Fig. 6

Owed to the different life experiences that A and B have undergone throughout their lives it is almost impossible that *altogether regarded*

B's pruning of 'bench' is identical to the meanings that A may have assigned to his/her 'bench'. For instance, even in A's mind, 'bench' can be related either to <table, chair> or to <picnic, Brighton, holiday> — selective prunes (also named *SelPrunes*, another of Lp's operators) that validate (each one of them) a *possible* meaning of 'bench' — as the simplified representation of Fig. 7 depicts. If, however, B's repertoire entails a meaning/perspective of 'bench' which *matches with* that lying in A's mesh, then A and B are in conditions of *understanding* what each other are speaking about. This overlapping/intersection is, just, the commonalty of perspectives previously referred to as a mandatory condition for mutual understanding in general (Fig. 8) Briefly, when engaged in conversation A and B share *conceptual models* of each other. This is an aspect of CT which lays a crucial emphasis not only upon the prominent role which the individual experience plays in the perspective which each of us has constructed of what is commonly named 'the world' (either socially or individually looked upon) but also on what we, human beings, have of idiosyncratic, personalised and distinctive.

Two (possible) meanings of 'bench' belonging to A's conceptual repertoire

Fig. 7

M-individual A   M-individual B
"BENCH"ₐ   "BENCH"_B
Perspective 1   Perspective 1
Perspective 2   Perspective n   Perspective 2
Perspective n
Perspective 3   Perspective 3

A Common Perspective

A and B *understand* what 'bench' means for both; or A and B *agree with* the (interpreted) meanings they both *share* about 'bench'

Fig. 8

2/3 - Conversation Theory and some of its (original) limitations

2/3.1 - In the preceding considerations it was assumed that A *and* B had in their conceptual repertoires some meaning(s) already assigned to 'bench'. But it may happen either that one of these M-individuals did not have this (or other) concept in their mind or that, instead of a topic concept, the conversation is focused on a phrase uttered by one of the interlocutors - usually, indeed, we do not communicate using simple terms but *ordered concatenation* of words. If the M-individuals engaged in dialogue are both adults, belong to the same cultural level, have common interests and so forth, it is probable that no communication problems arise (eventually they may lead either to psychological troubles such as those the so-called *double-bind* yields or, on the contrary, to psychic healing). But forgetting these psychological consequences of conversation, the important aspect to be stressed is that if one or several of those (usually unconscious) 'contractual' pre-conditions fail(s) then the original version of CT becomes unable to cope with such situations.

The reader can obviously argue that if someone does not know some word used by her interlocutor then what she has to do is to ask for its meaning, a strategy which *mutatis mutantis* may also be used about the real or apparent unknown meaning, possible to be assigned to a phrase or expression. However, from the viewpoint of CT, the answer to this argument is far from being easy, *especially if one of the interlocutors is not a human being but a machine*. As a matter of fact, Pask's CT is not an all-embracing and already completed theory. On the contrary, and notwithstanding all of his successful contributions, *from my perspective* it suffers from some severe limitations, particularly when the following aspects (amongst others) have to be taken into account :

i) The first deals with the notion of *coherence* itself. Indeed, coherences that are exchanged in *human* utterances entail far more than concepts. For instance, and *beyond verbs*, we also use a countless host of terms such as 'great', 'small', 'high', 'low', 'tall', 'short', 'young', 'old', 'middle age', 'close to', 'near', 'far', 'likely', 'approximately', 'more or less', 'almost', 'very' and so forth, which are used to quantify, qualify, compose or modify the meaning of our propositions (and consequently of our perspectives). All of these terms (context dependent, since they also require universes of discourse whose range is not only personal and idiosyncratic but also variable according to the *local* conversational topic into which they are inserted) have a *fuzzy character*. However, such a fuzzy perspective is not included in CT.

This is not surprising for those to whom Pask's work is familiar. As a matter of fact, from its very beginnings, CT was always tailored in terms of *modal*, not fuzzy, *logic* in which, in addition, *algebraic constructs* have been excluded. By the same token, although subjectivity in general may be looked upon as an integral part of the whole approach, the treatment it undergoes is primarily made in logical terms. So, all the *affective content* which such words, utterances and so forth may eventually enfold, has also been neglected in CT's original version.

ii) Secondly, owing to the idealist philosophical foundations upon which the theory rests, the distinction that any classifying system establishes between 'more specific' and 'more general' concepts (for example, between 'roses' and 'flowers') requiring the inclusion of those generality-levels that in linguistics and psycho-linguistics are named hyperonymy/hyponymy (and consequently yielding some kind of hierarchy) - have also been excluded from CT. Likewise, also the *order* which the elements of some utterance or written expression must obey (having in mind to assign some possible meaning to it) has also not been taken into account in CT's coherences. This fact begets an additional problem, *now related to the origin, genesis and further temporal development of our thought processes* either based upon free associations or on 'approximate' (fuzzy) premises, not encompassed in the approach.

iii) This temporal exclusion yields a third fundamental 'limitation' dealing now with the cognate suppression of everything related to the diachronic evolution that the interlocutors *and/or* the system whom they were engaged with in conversation undergo through time. But (amongst other consequences) to be engaged in a dialogue or *to explain* something to a child is not the same as talking with an adult. This means, briefly, that 'understanding' itself cannot be looked upon as an *absolute* concept as CT claims. Indeed, there may be some situations in which two individuals, say A and B, engaged in conversation have not full commonalty of perspectives (in the limiting case they can even be mutually exclusive and conflicting as well). Under these caveats, if 'full understanding' is numerically expressed by the number 1 and 'null understanding' by the number 0 then, *all possible intermediate situations will fall within the interval [0, 1] and 'understanding' itself will acquire a fuzzy character*. Furthermore : if A works, momentarily, as a 'teacher' and B as a 'student' then A's 'explanation' must thenceforth be looked upon as a *goal-directed task* (in the sense that such an 'explanation' seeks the maximisation of B's understanding). But this purposiveness must, in addition, satisfy an *adequacy*

requirement — a condition that the theoretical/computational tenets of CT also do not take into account.

iv) Fourthly, there is a question which brings to light the ancient mind-body problem, especially when regarded from the aforementioned diachronic perspective. As a matter of fact, Pask´s CT (as well as its computing applications, namely THOUGHTSTICKER or some of its further developments) *are, both, typical examples of 'mind-like' systems alone*. Thus, the existence of *a 'body-like' interface* so that, for instance, a 'table' becomes not a bare word or written concept but a specific 'outer object' with which some corporal movements and sensations may be associated, become (amidst other computing/robotic goals) mandatory requirements. This is nothing however but the host of objectives that my own research programme aims at achieving.

v) Finally, although related to the preceding 'limitations', there is a question which brings to light the temporal 'dynamics' of the mesh, in the sense that throughout time and throughout the various interactions with different interlocutors, each of us (or any of Pask's virtual machines) is always changing the (fuzzy) values attached to the concepts or to their inter-relationships lying either within any of the coherences of the mesh or even modifying the mesh itself through new recollections, new interpretations, etc. These, in turn, may beget the selection of new preferential associative/inferential 'paths' which, by the same token, may change the way we interact with other interlocutors (ourselves included) providing new insights, etc.

All of these factors are discussed in (Martins, 1998). For the moment let us however restrict our attention to those that, more intimately, are related to the emergence question.

3 - FURTHER DEVELOPMENTS OF CT AND EMERGENCE : 'OUTER' AND 'INNER' EMERGENCES, CURIOSITY AND UNDERSTANDING
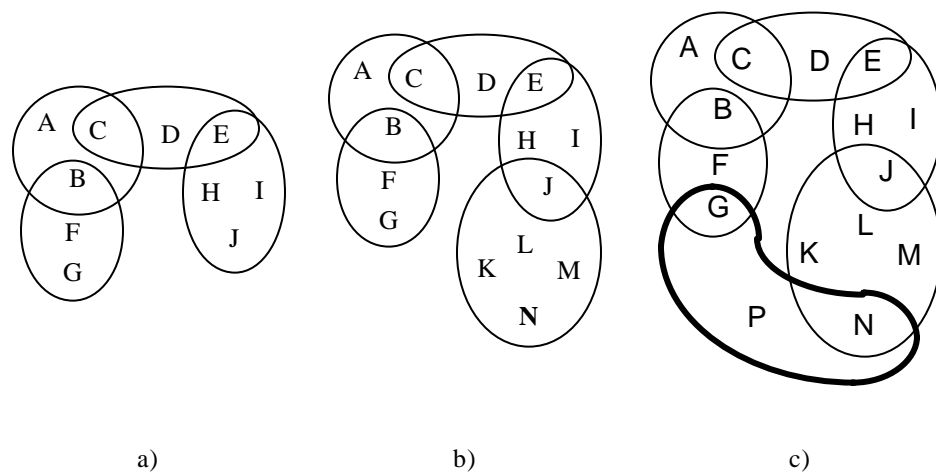
One of the main objectives of my research programme is to overcome some of these 'limitations' of Pask's CT (Martins, 1995 a). This essay formulates the goal of endowing non-natural systems with a physical, intellectual and affective autonomy; consequently, it also raises the question of emergence that this paper seeks to enlighten. Let us then return to what I said in section 1 about utterances such as "I had never thought of that!" or "This is the solution I have been looking for, for centuries!" often found in teaching, in psychotherapy, in psychoanalysis, in research, in processes entailing some kind of artistic

creation, etc., all of which are nothing but the externalisation of some kind of emergent phenomenon. Taking into account what was said above about the dynamics of entailment-meshes what can indeed be the 'mechanism(s)' underlying such emergences ?

Generally speaking I believe that the answer is altogether framed by two categorical types of emergence — which I have named *outer* and *inner* emergences — both of which entail, in turn, other factors (such as, for instance, needs, wishes, relational dynamics, etc.) to be cursorily mooted later. For the moment let us direct our attention to the foregoing 'outer' emergence.

3/1 - 'Outer' emergences and some related questions : meanings and the completion of wholes

3/1.1 - An 'outer' emergence is one which comes forth in consequence of a dialogue or conversation with someone who, *purposively* or not, is able to induce in the mind(s) of their interlocutor(s) a new and hitherto non-existent relational meaning between certain concepts. Figs. 9 a), b) and c) depict this kind of emergence.

a)   b)   c)

An outer emergence due to an external suggestion
Fig. 9

Indeed, by means of the new perspective <G, N, P> uttered by A, the interlocutor B becomes able of linking "G" to "N" through "P" (from here the foregoing "I had never thought of that!"). This is not all, however. In the work referred to above (Martins, 1995a), aimed at overcoming one of the limitations of CT (namely, the meaning-

assignment to a *phrase* or *expression* uttered by someone), I proposed a 'principle' governing one aspect of our mental functioning — namely *the principle of the completion of wholes.* This principle, cursorily stated, claims that

i) There is in the human mind a tendency to group under the same label (Gestalten) 'objects' A_i (i=1,2,...,n) amongst which a fuzzy similarity relation exists. If such 'objects' (a term to be interpreted in its widest sense) are *linguistically* described by attributes A_{ik} mutually shared by the A_i then these A_{ik} will too be interpreted as forming a *whole* to which a *unique* symbol is ascribed ;

ii) whenever two or more *distinct* 'objects' A_1, A_2, etc., are so strongly related to another, say M, that the recollecting of one of the A_i (i=1,2,...) yields the recollecting of M (and vice-versa), then there is a (psychological) tendency to set up a (real or imaginary) relationship amongst the A_i, so that these latter, together with M, form a whole/coherent whole (in Pask's sense).

Algebraically interpreted, this principle provides a possible explanation both for the so-called 'semantic closure' and for the meaning-assignment process. This is just the result I had in mind to stress about Fig. 9. As a matter of fact, when the introduction of "P" gives rise to a semantic closure in the open bundle of coherences represented in Fig. 9 b), this corresponds a potential but hitherto inexistent whole-embracing meaning. Briefly : it is as though the initial coherences were regarded as 'parts' of a partially or totally incomprehensible phrase which, after the introduction of <G, P, N>, may become meaningful and, consequently, *understood*.

3/1.2 - A concrete example of this situation (especially poignant for it provides a possible explanation for the self-understanding of simple phrases both for human and non-natural beings) is given in (Martins, 1995a). I shall not repeat the reasoning there exposed but, for the further comprehension of one of the possible manifold emergence mechanisms, there are some general ideas that the reader has to bear in mind. Suppose in this sense that A is an adult and B a four-year old child who asks A what a computer is. Suppose also that in his/her mental 'data base', A has the following two possible perspectives to answer B's question :

P_1 : A Computer is an <u>Electronic Device</u> containing several <u>Units</u>. These <u>Units</u> are made of <u>Chips</u> which contain <u>Thousands</u> of <u>Integrated Circuits</u>. Some of these circuits are <u>Electronic Oscillators</u>

(<u>Monostable</u>, <u>Bistable</u> and <u>Astable</u>) which, once suitably triggered, beget a series of <u>Internal Instructions</u>, commonly called <u>Program</u>. Based on it, the Computer receives <u>Complex Input Data</u> and treats it, so that an adequate <u>Output</u> <u>Information</u> is provided.

P_2 : A <u>Computer</u> is that <u>Kind of Television</u> that <u>Uncle John</u> has in his <u>Office</u>.

It is obvious that, engaged in conversation with a *real* child, P_1 is clearly out of court. I only introduced it, having in mind to lay some emphasis upon the aforementioned *adequacy*, i.e. that when two individuals are engaged in a *fair* conversation, they tend to 'adapt' their 'inner reference frames' (their minds, language, etc.) to each other, so that their understanding from each other's perspective becomes *maximal*. Having this in mind, let us then assume that A utters P_2. Under this proviso let us more attentively examine the possible processes (and cognate problems) that such an utterance yields in the minds both of B and also of A. In this sense, when A says that "A <u>Computer</u> is that <u>Kind of Television</u> that <u>Uncle John</u> has in his <u>Office</u>", he is providing a explanation for "computer" which, albeit meaningful for him, presupposes a host of previous assumptions. Beyond the obvious knowledge that *this* child must already have of "Uncle John" and "Television", A is also assuming in fact that (a) B has already been in the office of Uncle John, (b) there, the child has been looking at the computer (even though he did not know what this term signifies), (c) B is already able to understand what the fuzzy relation "a kind of (something)" means. In other words: that the child can already set up a similarity or analogy *relation* between 'things', in this case "television" and which A calls "computer".

For the sake of concreteness let us divide these presuppositions into three distinct, although cognate, categories : one dealing with the recollecting/interpretation (from B's side) of the *individual* concepts or perspectives entailed in A's utterance ; the second with their *condensation* (including the construction of relations between them) ; and thirdly with the *setting up of such analogies or similarity relations* yielding the formation of an overall Gestalt image or 'integrated whole-meaning' of "computer". The evolutionary stages of this process-formation of meaning-assignment are depicted in Fig. 10 a), b), c). To this end let us assume, at the outset, that "Kind of Television" recollects in B's mind "Television" to which "a <u>Big Black Box</u> I have at <u>Home</u>, in the <u>Dining Room</u>, where I can see <u>The Ninja Turtles</u>" is a *possible* associated perspective.

Also, let an initial perspective of "Uncle John" be given by "a Nice Man, who gives me Pleasant Gifts at Christmas and on my Birthdays". Finally, let us assume that "Office" is "a Room where a Desk and Chairs exist".
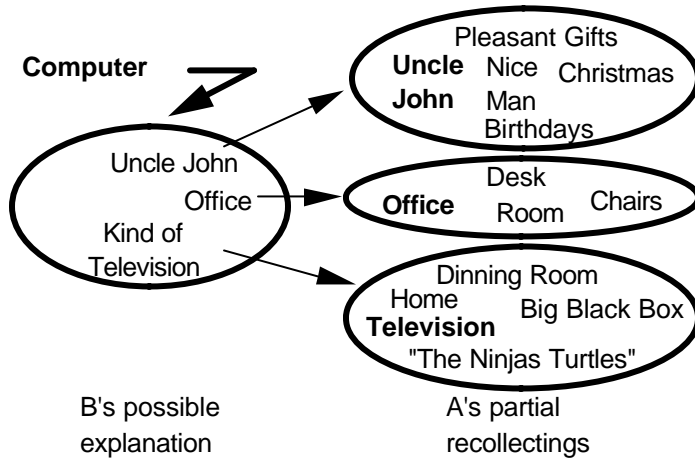

Fig. 10a)

This situation is symbolically depicted in Fig. 10 a), where the concepts written in bold type represent Pruning operations and the remaining terms symbolise their Selprunes.

Looking at the picture, the reader can see that, *if nothing is added to it*, then *no integrated meaning is ascribed* to A's overall utterance. In other words, although *partial* meanings may be assigned to the concepts uttered by A, B is unable to bind them together. Consequently, at this stage of B's process, A's explanation (as a coherent and meaningful *whole*) works as though it were *globally incomprehensible* for the child. To continue our examination of the process, two further possibilities must be taken into account :

i) B is able *by himself* to recollect some other perspectives relating his interpretations of "Television" to "Uncle John" and to "Office" (which requires that B must already have had *some* previous knowledge of such relations) ;
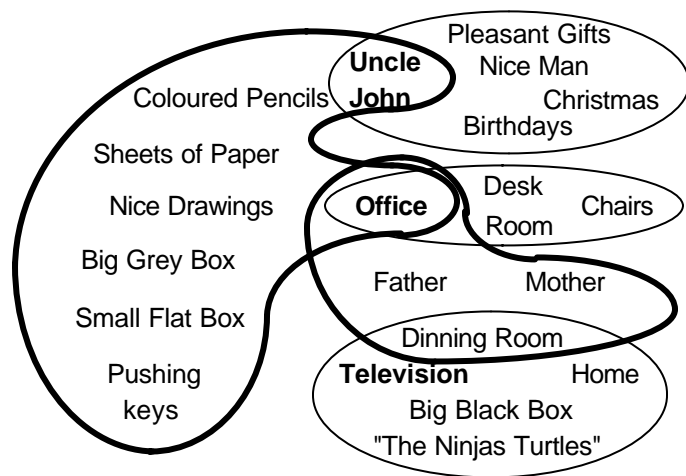
ii) or, on the contrary, B does not remember (or does not know at all) this binding relation, situation in which he has to ask for what such a relation may be.

Let us begin with i) for the sake of simplicity. Suppose in this sense that, in B's mind, the "Office" of Uncle John recollects "Father", "Mother" and "Dining Room". Suppose also that he is *really* able to recall something like "Whenever I am at the Office of Uncle John, he gives me Coloured Pencils and Sheets of Paper with which I can make Nice Drawings. I also remember that, being there, I saw a Big Grey Box and a Small Flat Box, with which I play Pushing Keys" (Fig. 10 b). This
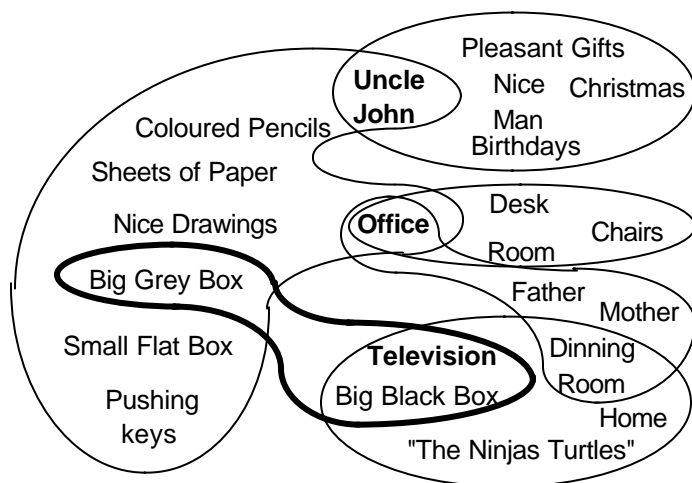
means that "Uncle John" becomes *either* "a nice man who gives pleasant gifts by Christmas and birthdays" *or* someone to whom some pleasurable memories related to "coloured pencils, sheets of paper, nice drawings" and so forth are, also, connected.

However, relating "Uncle John" to "Office" in this way, they become bound in a fashion that clearly facilitates the setting up of a similarity relationship between the "Big Grey Box" and the "Big Black Box which the Television of the Dining Room" is supposed to be (Fig. 10 c).

If B is able to set up this last relation *by himself*, then it can be said that A's explanation has been *fully understood* by the child. Otherwise, in his interaction with A, B will plausibly ask for the possible connection between A's utterance and that explanation. But this is just the situation previously referred to in ii) which, from the point of view of emergence, is clearly inserted into the designation 'outer emergence' previously stressed. As a matter of fact, when A said that "A Computer is a Kind of Television that Uncle John has in his office",

b)

c)

An evolutionary description of the meaning-assignment process
Fig. 10

A was *inducing* in B's mind a similarity relation between the Big Black Box that B has in the dinning room and the Big Grey Box existing in Uncle John's office. This implicit suggestion yields in B's mind the formation of a new coherence (that one just depicted in Fig. 10 c) which, binding together "Uncle John", "Office" and "Television" provides an *unified meaning* to A's explanation. In other words : B will

be in conditions either of understanding A's explanatory utterance or of assigning an overall, albeit, *unique* meaning to *this* 'explanation'. Thus, this mechanism not only overcomes one of the aforementioned limitations of CT, but also of provides a possible explanation of the so-called 'outer' emergences previously referred to.

3/1.3 It is worth noting that in terms of the overall objectives of this work, the foregoing 'mechanism' encompasses understanding, interactions, meaning-assignment and emergences in general. Although so far I have only been dealing with social 'outer' emergences, this mechanism also suggests possible clues for inscribing the question of emergence in the reference frame stressed in 2/1.4, and for approaching an explanation of 'inner' emergences.

Let us assume that what for some observer had always been looked upon as a stable system begins to exhibit those new, unexpected, irreversible and not yet predictable changes that so far have been related to emergences. I shall not scrutinise the countless *particular* reasons (either physical or psychological) which may beget their effective appearance. But whatever they may be, I believe that three prototypical situations exist that all the emergence scrutinies have to take into account, particularly when they belong to the realm of hard sciences :

a) In order to provide an explanation of the new observed features that is coherent with his past knowledge, the observer is forced to add new 'spatial' dimensions and new descriptive attributes/relations to those hitherto considered. However, insofar as these additional dimensions and meanings do not correspond to a novel overall meaning, the *general* meaning assigned to the system remains the same.

b) The number of 'spatial' dimensions and the corresponding attributes/relations by means of which the system is described are *potentially* the same but, for some reason, the observer assigns a new and different meaning to the system.

c) Both the 'spatial' dimensions and the meanings assigned to them undergo significant changes over some period of time.

These extreme situations are depicted in Fig. 11 a), b) and c) where, for the sake of simplicity, all the possible temporal cuts corresponding to successive observations, are condensed in a single symbolic 'plane'.

Three prototypical emergent situations
Fig. 11

The reader must, however, bear in mind that this temporal 'plane' actually represents a more or less extended period of observation period. Hence, it obeys not only Collingwood's 'principle of minimal time' (Collingwood, 1970) but also its cognate conclusion: particularly, that the image which each of us constructs of the world depends (amongst the other features already mooted so far) on the time we spend in its observation.

The figures are practically self-explanatory. It is however interesting to relate them to some of the more conventional approaches that have aimed at a definition of emergence. Generally speaking — and without any intention of discrediting such proposals — it seems that many such attempts have been deeply influenced by some of the mechanist standpoints. Especially poignant are, in this sense :

i) The idea that emergence is *almost exclusively* associated with the 'spatial' dimensions of the general framework condensed in Fig. 11, together with a rigid and often *predetermined hierarchy* related to them ;

ii) The features are ascribed to the *observer* of some emergent phenomenon, who either is *not explicitly taken into account* or, when considered, is essentially the *objective observer* of that paradigm ;
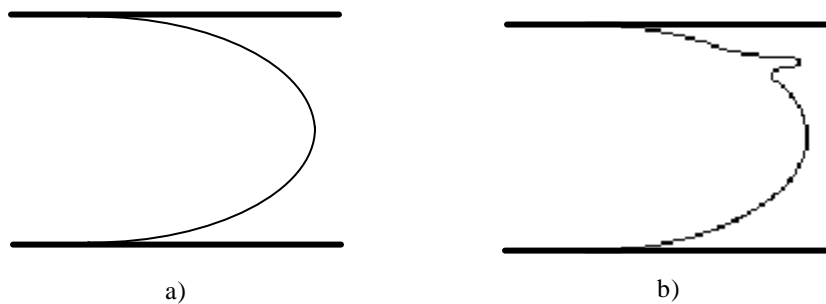
iii) The consequent *absolute character* assigned both to this (adult and unchangeable) observer *and/or* to the observed emergent *events*. This view neglects both the interpretative, relative and social dimension

such events do always have ; and their temporal dimension and also their relative durations.

There are many definitions of emergence in the literature, but I shall mention only two. One, owed to Heylighen (1990) who recognizes that no satisfactory theory exists to explain emergent properties or to specify the conditions for their appearance. He characterises emergence as "a classical concept of systems theory, where it denotes the principle that the global properties defining higher order systems or wholes (e.g. boundaries, organisation, control, ...) can, in general not be reduced to the properties of the lower order subsystems or parts". The second example is presented by Peter Cariani (1998). For him, "emergence is the process by which new, complex order arises from what is simpler. Novel structures and functions can either arise by creating new combinations of pre-existing primitives, or alternately by creating new primitives de novo. Examples are rearrangements of strings of letters vs. creation of new types of letters (novel combination within a fixed notational system vs. extension of the notation itself via new categories)". My criticism is that these proposals need to be augmented by explicitly adding to them both a temporal dimension and a meaning dimension. On the one hand, "higher order systems or wholes" or "a new rearrangement of letters, sensations, ideas and so forth" do not suffice *by themselves alone* to define something meaningful. As a further issue, this 'something meaningful' is only meaningful *for someone* — either an M-individual or a group of M-individuals sharing the same perspectives (what Pask called *P-individual*, where 'P' stands for 'psychologically characterised' individual). This is a fact which always presupposes individual interactions, conversation, shared agreements, etc., as well as a minimal time and a minimal 'space', this in order that such a 'something' is knowable, understandable and accepted (or, eventually, rejected *now* and accepted *later* or vice versa). But these conditions require an overall framework similar to the one I have proposed.

Assuming the cogency of this criticism, then Cariani's idea of the second type of emergence ("creation of new primitives de novo", together with some of my already added meaning) corresponds to the situations that Figs. 11 a) and c) depict. Fig. 11 a) portrays quite well what (for me) is the easiest emergent situation. Some phenomenon exists to which an observer has already assigned some general meaning, say A. However, throughout further observations and albeit holding such general meaning, the observer is forced to add new descriptive attributes in order that the new observed phenomena be taken into

account. A physical example of this situation — taken from Prigogine (1971) — is depicted in Fig. 12 a), b). *Both* pictures portrait the velocity profile of a *laminar* flow in a pipe. However, whereas Fig. 12 a) provides a typical representation of this type of flow, in Fig. 12 b) such a profile presents a protuberance owed to some small fluctuation dE_c in the kinetic energy of the system. If, for , t Æ•, the fluctuation that dE_c undergoes is such that the system's response tends to restore the initial equilibrium state, then the flow remains stable and dE_c tends to become null. On the contrary, if dE_c tends to grow throughout time, then a new type of flow will also tend to arise. This phenomenon becomes especially poignant whenever Reynold's number surpasses the value that corresponds to the emergence of a *turbulent* flow. But this distinction between 'laminar' and 'turbulent' flow is crucially important for the emergence question, particularly to the situations that Fig. 11 embraces.

a)   b)

Fluctuations in the velocity profile of a laminar flow
Fig. 12

Indeed, from a physical point of view, what the preceding assertions ultimately express is that the emergence of a new structure (corresponding to what physicists have named 'turbulent flow') is a consequence of an instability, everything working as though, instead of a response tending to restore the disturbed equilibrium state, the aforementioned fluctuation is, on the contrary, *amplified and gives rise to the emergence of a new and totally different streamline pattern*. Hence, from the perspective of Fig. 11, what this means is that — as long as the fluctuations are small enough to hold the values of the variables within the initial stable boundaries — the observer
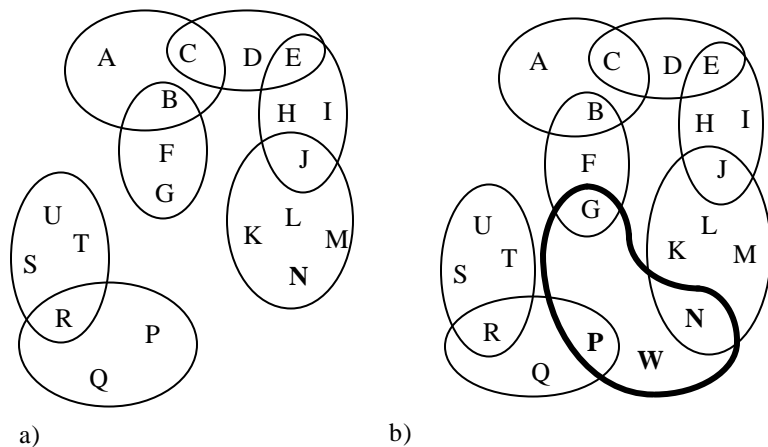
i) through an almost continuous observation of the phenomenon, and

ii) using similarity relationships between two successive observations,

can interpret all the observed changes as modifications of *the same structural type of flow*, to which he assigns a meaning corresponding to the term 'laminar'. This is just which Fig. 11 a) depicts, *if and only if the duration of the phenomenon is important for the observer*. For, if — instead of the almost continuous observation previously assumed — the interval of time elapsed between two consecutive observations is long enough for the laminar flow to have changed to the other 'turbulent' structure, then he can be faced with a *situation of discontinuity*, as though there were no possible relation between its behaviour 'before' and its 'after' (a difference that brings to light the definition of emergence afforded by Rosen and quoted by Cariani, as "the deviation of the behaviour of a system from the behaviour predicted by a model of that system"). Furthermore: if the observer has no knowledge of hydrodynamics, then he will have to find a new term and a new meaning to denote the new state of affairs as well as to search for the diachronous reasons justifying its emergence — which will surely require the introduction of new descriptive attributes, new individual time scales that allow him *to fill in the 'blank spaces' between observations*, etc. ; briefly, what Figs. 11 c) portrays. Once again, however, this is just another illustration of the statement that the idea that each of us constructs of the world depends on ourselves, e.g., on our individual perspectives and, consequently, on our social interactions as well as on the relation between the duration of our observations and the duration of the observed/interpreted processes each of us is scrutinising. But beyond its relationships with the meaning-assignment question, this result also lays emphasis upon the other type of emergence previously referred to : 'inner' emergences.

3/2 - Inner emergences, the completion of wholes, curiosity and mesh-dynamics

3/2.1 - This type of emergences does not require *in itself* any outer dimension, although the influence of others is, of course, always present, particularly in terms of memories. Generally speaking, the reader may think of it as something tantamount to the well-known 'stream of consciousness', to Thomson's autistic thinking (Thomson, 1975) or — within the scope of CT — to another of the *practical* limitations of Pask's theory, now related to one of his Lp operators called *Saturation*. Cursorily exposed, this operator deals with the possibility that the mesh's user has of setting up *self-associations* between the topic concepts it contains, so that no further coherences can be added to it. From Pask's perspective the major idea lying beneath the implementation of this operator is that the more interconnected a
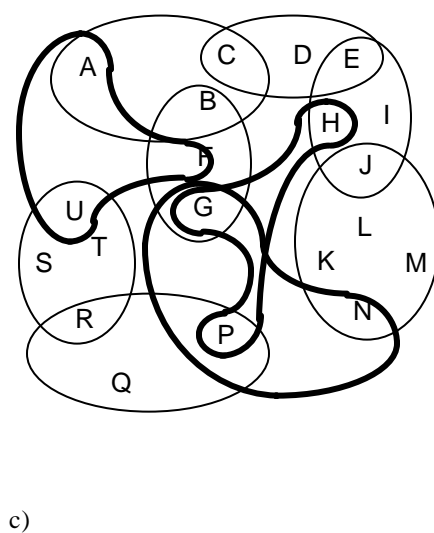
mesh (or a mind) is, the more efficient its functioning will be. Psychologically speaking this is an opinion which *in principle* I agree with (a proviso that is also shared by all psychoanalysts), since it would mean that from some initial topic we should be capable of reaching any other topic, everything working as though no 'forbidden' areas could exist in our minds.

a)   b)

Also, from the point of view of the emergence question, such a possibility of endowing a mesh with a 'mechanism' yielding self-associations is crucially poignant since it begets new and *unexpected* coherences beyond those actually observed or purposively suggested. The reader may recall the Archimedean "Eureka!", Kékule's dream that gave rise to his discovery of the cyclical structure of some organic carbon compounds, or the expression "I had been looking for that, for

c)

An example of inner emergences
Fig. 13

centuries!"— in each case, without explicit 'outer' influences). A sketch of such possible associations are shown in Figs. 13 a), b), c). One (Fig. 13 b), deals with the perspective <G, N, P, W> which, similarly to what happened in the example of Fig. 9, was used to provide a connection between the bundles (<P, Q, R>, <R, S, T, U>) and the remaining

clusters, *but now supposed to arise without any outer interference*. The other (Fig. 13 c) is part of the mathematical combinations that 'Saturation' entails, only stressed as a means of bringing the problem to light. As a matter of fact, it can be demonstrated that any m-saturated mesh on n nodes (say, a mesh where all possible m coherences have been utilised without contravening the Genoa rule) is isomorphic to an algebraic entity called a *Steiner system* — in particular $S(m, m+1, n)$ — where the Steiner system $S(t, k, v)$ is a block design in which each subject of t elements from v appears once (only) in a design of k-element blocks selected from the v elements. The blocks of the design may be identified with the locally cyclic coherences and the elements of the design with the nodes of the mesh. Thus, assuming that we are concerned with coherences of 3 nodes, the first number of nodes that can exhibit full saturation is 3 (every node is derivable from the other two). Steiner showed that the next number of nodes exhibiting full saturation for a 2-maximally saturated mesh is 7, 9, 13, 15, 19, 21, 25, etc. As a further result, any maximally 2-saturated mesh of n nodes has $n(n-1)/6$ locally cyclic blocks in it. But, although there are existence-theorems for $S(2, 3, n)$ and $S(3, 4, n)$, few known systems exist for $k > 4$. And even for $m = 2$ and $k = 3$ there are 80 Steiner systems on 15 elements. This shows not only that the computation of Steiner systems (even when their existence is known) is far from being a simple task but also that, *from a psychological perspective, no one can a priori guarantee that a particular cogent meaning may be assigned to some (algebraic) combination of elements.*

3/2.2 - To overcome some of these problems — which, in essence, are quite similar to the questions that the so-called 'complexity trends' are seeking to solve, albeit using mechanist approaches — I have been pursuing an approach in which a great deal of attention is paid to what, generally speaking, I may call a *relational dynamics*. I shall not go into details concerning the host of problems this 'dynamics' entails (although it is already capable of providing answers to the possible ultimate 'mechanisms' on which inner emergences rest), since they have been described in (Martins, 1998). Three of its major framing ideas deserve however special attention :

i) The importance ascribed to the biological fact that human beings are altricial, incessorial and nidiculous animals whose physiological survival during the long period that childhood lasts, demands a total dependence on the parents' support ;

ii) The great deal of emphasis which has been laid upon the *diachronous* or *temporal* 'dimension' attached to all human beings

from their earlier stages of life to adulthood as well as to the role that *affections* play in the setting up of the first inter-relationships both with the earliest significant figures and the 'outer' and 'inner' worlds of babies and children ;

iii) This obviously presupposes the existence of *social interactions* as well as the importance ascribed to the aforementioned 'outer' emergences.

Translated into a different language, what these conditions yield is the construction of a general framework consisting of two complementary 'mechanisms' which — if no traumatic or innate disturbances exist - tend progressively to converge towards the setting up of a unified bodily mind. One of these 'mechanisms' is primarily (although not exclusively) biological and comes from i) above. It deals with the psychological consequences that the periodical *homeostatic disruptions* commonly named 'needs' (of which hunger is crucially relevant in the earliest stages of the baby's life) have for the progressive organisation of his/her mind. Such disruptions give rise to 'experiences of satisfaction' (Freud, 1895) — the appeasement of the need, relief, satiety and pleasure. Once repeated several times, such experiences beget in turn the triggering of *wishes* as well as the setting up of *emotional ties* between the child and the figure who becomes physiologically and psychically significant for such an appeasement (usually the 'mother'). Obviously, the way according to which such needs and wishes are or are not fulfilled also entails the emergence of other feelings such as frustrations, anxiety, fear, etc., or on the other hand, of attachment links, of love, affection, trusting and so forth (although never excluding the unavoidable existence of frustrations and conflicts). But whatever the intensity of their *affective balance* may be, the creation of the ties between child and parents always yields the progressive formation of an overall structure named by Freud the *superego* — whose role in the setting up of inner and outer emergences, becomes crucially important.

The progressive erection of this structure cannot be presented here in detail. Generally speaking, it deals with the conditions which the child *imagines* as being necessary to obtain the parents' love (love and solicitude which, indeed, protects the child from the dangers that threatens him/her not only from the outer world but also from 'within'). The price the child has to pay for this security is a *fear of loss of love*. And, in order to avoid it, the child loves himself, judges himself, condemns or punishes himself in the same way that his/her parents have loved, judged, condemned or punished him. In these conditions

everything works as though love or condemnation were dependent on the agreement or disagreement of a critical or moral *inner* image (an *internalised standard*) which, ultimately, is nothing but a copy of the parentals' demands. This 'standard' (Freud's 'superego') plays two crucial roles in the future development of the child. On the one hand (psychologically speaking), it works as some kind of 'inner *motive force'* which, by means of *identifications* and *projective identifications* (as psychic 'mechanisms') impels the child to a psychic growth, tending to make him do the same things that parents do. Briefly : to acquire a physical and psychic *autonomy* — to which interpretation, understanding, wishes, desires, curiosity and emergences are intimately related. But as far as this autonomy tends to increase — and here the physical independence that the acquisition of any type of *self-displacement* represents becomes crucially important — it also begets, on the other hand, the earlier prohibitions, warnings, punishments, etc., in brief the *physical* and *emotional constraints and rewards*, which, thereafter, will also become inseparably related to the superego.
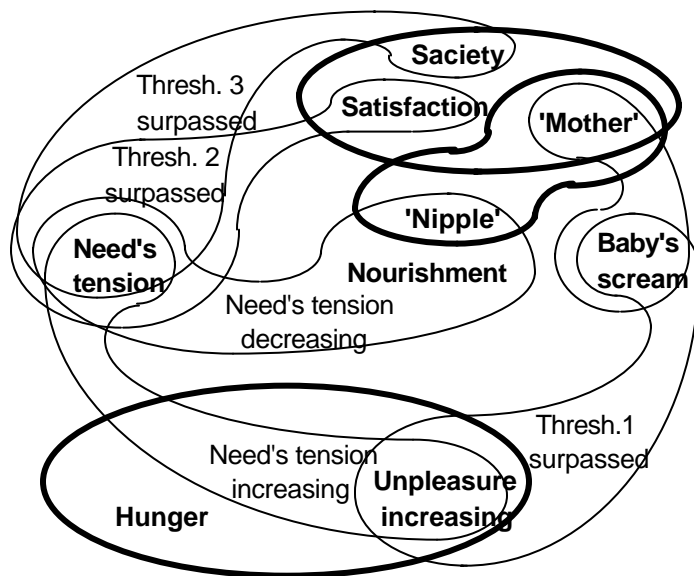
3/2.3 - These affective and cognitive 'mechanisms' entail an extreme complexity which cannot be adequately rendered here (Martins, 1990, 1992, 1993, 1994, 1995, 1996, 1998). Several conclusions directly related to the immediate goals of this work can, however, be drawn from the foregoing cursory considerations. Firstly, if the affective balance between "constraints" and "autonomy" tips towards the latter, then the 'outer' world becomes the source of intense and *pleasurable curiosity and expectations*. This leads to an increasing development of the child's 'inner' world where the role played by imagination, fantasy, setting up of analogy relations, etc., together with the projective identifications regarding the significant figure(s), tend to contribute to the setting up of new associations, new relationships, etc. ; briefly, emergences in the sense that, since the beginnings of the work I have been stressing.

The same arguments apply to the second of the overall conclusions that I wish to draw. For, if we look backwards at the problems that beset Pask's 'Saturation' operator, then it can be seen that with this relational dynamical approach, the number of possible combinations that Steiner analysed becomes enormously restricted. Indeed, it is one thing to seek a mathematical description of *all the possible* combinations existing amongst 2000 terms (the normal vocabulary of a 5 year-old child) ; it is another to consider the progressive building up of a more and more elaborate mesh in which — thanks to the associated affective values, to the correspondence (or not) to some aspects of the 'outer' world, to the

number of times a group of coherences is presented and the order in which its entailed concepts are recollected, etc. — there is an 'automatic' selection which prevents some combinations being made and, simultaneously, allows others to come forth.

For instance, if we look at Fig. 14 where only a small part of the preceding survey is depicted — a 'flat' representation of a much more complex 'n-dimensional' picture into which the preceding 'mechanisms' may be inserted — then it is easy to see how *meaningless* it becomes to set up a combination entailing, for example, <displeasure increasing, baby's scream, satisfaction >.

Obviously, the existence both of these constraints and of the aforementioned 'motive forces' impelling the child for an increasing exploration of his/her world does not stop here. Indeed, throughout one's life, the superego receives further contributions from later successors and substitutes of the parents — teachers, public figures, admired social ideals and so forth — and, through them, of the family, of racial and national traditions, etc.

Part of the physiological and affective mechanism of hungriness
Fig. 14

Thus, they work as some kind of internalised image, mirrored from the outer milieu into which some individual is inserted, an image which

prolongs (usually by means of unconscious mnemic residues) a person's own parental relationships. This (usually unconscious) influence of the past upon the ego's actual and contemporary reactions is precisely what led Freud to assert in his "Second Introductory Lectures" that

> "...the superego is the vehicle of tradition and of all the time-resisting judgements of value which have propagated from generation to generation. The past, the tradition of the race and of people live in the ideologies of the superego and yields only slowly to the influence of the present and to new changes".

However, to state the question in these terms is tantamount to laying emphasis upon the temporally changeable inter-relationships that may exist between an *'outer P-individual'* (representative of some group, school of thought, religious believers, etc.) and an '*inner P-individual*' (corresponding to some symbolic part of an M-individual). In the course of time, and by means of conversation, these inter-relationships can mutually promote or prevent the generation of 'outer' and 'inner' emergent phenomena. In contrast with the traditional mechanist/reductionist perspective, this is the unconventional approach to emergence — normative, relativist and social — that I have sought to present.

References

Collingwood R. (1970).*The Idea of Nature*. Clarendon Press Oxford.

Cariani P. (1996). Emergence of Signalling Primitives in Adaptative Timing Networks. Draft Paper. Journée Scientifique de l'ARC, ENST, Paris.

Freud S. (1964). *On Metapsychology: the Theory of Psychoanalysis*. The Pelican Freud Library, reprinted edition.

Grumbach A. (1996). Emergence et niveaux d'organisation : un problème d'observateur ? Draft Paper. Journée Scientifique de l'ARC, ENST, Paris.

Gregory D. (1986). Philosophy and Practice in Knowledge Representation. *Human Productivity Enhancement*, Vol I, Ed. J. Zeidner., Praeger,

Gregory D. and Pask G. (1986). Conversational Systems. *Human Productivity Enhancement*, Vol I, Ed. J. Zeidner., Praeger,

Gregory D. (1987). The Representation of Knowledge : Pask's Lp. Systems Research

Heylighen F. (1991). Self-Organization, Emergence and the Architecture of Complexity. *Communication and Cognition*.

Izard C. and Kagan J. and Zajonc R. (1990). *Emotions, Cognition and Behavior*. Cambridge University Press.

Jackendorff R. (1987). *Consciousness and the Computational Mind*. MIT Press.

Medina Martins, P. R. (1990). On the Simulation of Depth Psychological Processes. Xth European Meeting on Cybernetics and Systems Research, Wien.

Medina Martins P. R. and Rocha L. (1992). The In and the Out. XIth European Meeting on Cybernetics and Systems Research, Wien.

Medina Martins P. R. et al. (1993). CYBORGS : A Fuzzy Conversational System. Final Report, NATO Internat. Programme on Learning Systems. March.

Medina Martins P. R. (1995a). Metalogues : An Abridge of a Genetic Psychology of Non Natural Systems. CCAI, on *'Self Reference in Cognitive and Biological Systems'*, Vol. 12.

Medina Martins P. R. et al. (1995b). Objectal Relations : a Cybernetic Approach. SPA, Lisbon.

Medina Martins P. R (1998). The Dynamics of Entailment-Meshes. 2nd International Conference on Computer Anticipatory Systems, Liège 98.

Memni D. (1996), Emergence et niveaux d'explication. Draft Paper. Journée Scientifique de l'ARC, ENST, Paris.

Pask G. (1975). *The Cybernetics of Human Learning and Performance : A Guide to Theory and Research.* Hutchinson Educational.

Pask G. (1976). *Conversation, Cognition and Learning*. Elsevier.

Pask G. (1976). *Conversation Theory: Applications in Education and Epistemology*. Elsevier.

Pask G. (1980). Developments in Conversation Theory. *International Journal for Human Machines Studies*, Vol. 13, pp 357-411.

Pask G. (1980). Concepts, Coherence and Language. Vth European Congress on Cybernetics and Systems Research, Wien.

Pask G. and Gregory D. (1986). Cognitive Process Research. *Human Productivity Enhancement*, Vol I, Ed. J. Zeidner., Praeger.

Pask G. (1994). The Lies We Tell To Make Life True. XIIth European Congress on Cybernetics and Systems Research, Wien.

Thomson, P. (1975). *The Psychology of Thinking*. Pelican Books