ADAPT

IST–2001-37173

Artificial Development Approach to Presence Technologies

Deliverable Item 2.1

A theory of intentionality

Delivery Date: September 28, 2003

Classification: Public
Responsible Person: Riccardo Manzotti – LIRA-Lab University of Genoa
Partners Contributed: Jacqueline Nadel, Giorgio Metta

Short Description: Presence corresponds to the feeling of being there that, in turn, is conceivable only in a conscious being. A conscious being is a system that experiences (feels) something. This capability of feeling something depends on what is called the aboutness of phenomenal states, a property which is related to the intentionality of mental states. Thus we claim here that in order to understand the feeling of being there or presence we need to understand the nature of aboutness and intentionality in a conscious being. Any project that aims at understanding human perception thus needs to take issue with the notion of first-person experience or feeling. As it is argued in the following, this task can be profitably approached if we leave behind the dualist framework of traditional Cartesian substance metaphysics and rather adopt a process-metaphysical stance. We begin by sketching the outline of a process-ontological scheme whose basic entities are called "onphenes". From within this scheme a series of constraints on a possible architecture capable of intentionality and aboutness is formulated. An architecture abiding these constraints is capable of epigenesis driven by onphenes. Since an onphene is a process in which the occurrence of an event creates the conditions for the occurrence of another event of the same kind, an onphene-based architecture allows external events to cause the repetition of other events of the same kind. In an artificial system, this propensity to repeat events can be considered as a functional reconstruction of motivation.

Project funded by the European Community under the "Information Society Technologies" Programme (1998-2002)

Summary

1. Introduction

Presence refers explicitly to the notion of being there which, in turn, refers to the understanding of phenomenal experience. Phenomenal experience is practically coextensive with a traditionally taboo word in many fields: consciousness. However, such ostracism is no more necessary and we can advance scientific hypotheses about the nature of subjectivity and the physical condition required for its emergence (Chalmers 1996; Damasio 1999; Edelman and Tononi 2000; Crick and Koch 2003; Gallese and Metzinger 2003). This is unavoidable if we want to deal scientifically and practically with subjective phenomenal experience (presence).

A series of terms used in psychological language does not have a direct correspondence into robotics and physical sciences (for instance, representation, intentionality, and meaning). An attempt at dealing with the sense of being-there must challenge these concepts by providing a unitary framework in which these terms can be used. A tentative "theory of intentionality and of the sense of being there" will be formulated as a working background at the beginning of the projects (a tentative theory of meaning and intentionality).

This theory will make a series of hypotheses in order to bridge the gap between cognitive systems and physical systems. These hypotheses must put forward a candidate structure as the structure responsible for the occurrence of meaning and representation. They must provide necessary and sufficient criteria to locate the occurrence of meaning and representation.

This theory must propose a structural difference that will be used to test if an artificial system is processing information and if it develops intentional motivations. A possible way to this objective is by defining in physical terms what we mean by sensory-motor loop and by defining the role of epigenetic and phylogenetic development in a system. We believe that a crucial factor in developing true meaningful representation is the capability of a system of producing internal epigenetic criteria for further development. Therefore the success of the candidate theory will be evaluated by checking its capability of expressing cognitive and mental terminology in terms of objective structural conditions (like sensory-motor loops, causal relation between experiential events and subsequent system development).

On the other hand, a scientific theory of intentionality must be translatable into a series of experiments and suggestions about how to replicate the physical conditions for the emergence of phenomenal experience (presence). According to the present theory, two concepts are fundamental: i) the developmental nature of the subject and ii) the physical continuity between the subject and the external environment.

Within Adapt, there will be two lines of enquiry: the first line goes along the design and implementation of an artificial system whose architecture should implement an epigenetic intentional structure. Epigenetic means that the development of the structure is not completely dependent on the initial design (such as the process of imprinting or conditional learning in biological systems). Intentionality means that the possible content of the structure is not defined a priori but it is a result of the actual occurrence of an event. The second line of enquiry will investigate human postnatal development with the aim of at least in part validate the theory of intentionality presented here.

A caveat has to be made here on the usage of two words: ontogenesis and epigenesis. Often, their use is somehow confused. Especially in biological literature, ontogenesis is often related to the individual growth according to some internally contained genetic plan. On the other hand, the word epigenesis emphasizes the role of the active interaction with the environment during development. For this reason, in the following, we will use in prevalence the term epigenesis (and the related adjective epigenetic). However, according to our approach the term ontogenetic could have been acceptable as well since we use it mainly as the opposite of phylogenetic. In other words, the use of the pair ontogenetic/phylogenetic mirrors the dichotomy between nature and nurture, between what is hard-wired or innate and what is derived from experience and development.

Another caveat refers to the two words ‘aboutness’ and ‘intentionality’ that could be misleading. Historically the term ‘intentionality’ has been introduced in modern literature by Franz Brentano as the essential property of mental states and hereafter it has been extensively used (Brentano 1874/1973; Searle 1983). Traditionally there have been two different definitions for intentionality: one is related to the capability of agents of fostering intentions towards future actions and goals (Dennett 1987). On the other hand the term ‘aboutness’ is more recent and it is somehow narrower in its scope since it does not require any reference to agency or

purposefulness but it is used to refer to the fact that mental states point to something in the world (a phenomenal quality, a fact, an event, a physical property, an event, a relation between events). Aboutness and intentionality have always been considered scientifically awkward to the point that many authors despaired to find any good "candidate" for them (Fodor 1998). To avoid misunderstanding we will use here the two terms with the following meaning. Aboutness refers to the fact that mental events refer to something else (they are about something: a pain, a taste, a color). Intentionality refers to the fact that mental events refer to something else under some purpose or expectancy. The concept of intentionality is thus related to the concept of agency and responsibility.

In Table 1, we represent the relation between these concepts. From an empirical point of view, aboutness is related to embodiment and situatedness and thus it requires at least a capacity for stimulus dependency. Similarly, phenomenal experience has been insofar identified only in biological beings capable of stimulus dependency. Pure mechanical reactive systems are not supposed to be endowed with phenomenal experiences. A higher level of aboutness – that we call here intentionality – is produced in systems capable of stimulus expectancy and agency. The difference is that – in the case of agency – phenomenal experiences are perceived as part of the agent's history and are thus enriched in their meaning (the feeling of lemonade becomes the feeling of 'my' lemonade). Are there systems with intentionality without aboutness? We believe there are no real cases because they should be disembodied purposeful system: a contradiction in terms. However, an imaginary example could be the character HAL 9000 in "2001 A Space Odyssey".

In the following paragraph, we will try to show the identity between phenomenal experience and representation; hence the need of understanding the nature of representation in order to understand the nature of presence. In section 3, a working hypothesis about the causal nature of representation is described. The type of causal relation is a particular case of reciprocal causation, here termed *onphene* for practical and theoretical reasons. In section 4, the relation between representation, intentionality and meaning is discussed. In section 5, we will try to show that this kind of causal structure shows many similarities with the existence of goals and motivations in agents and particularly that the development of new motivations is needed in order to have onphenes.

	Phenomenal experience	**Aboutness**	**Agency**	**Intentionality**	Examples
Pure mechanical reactivity	NO	No	No	No	Simple machines
Stimulus dependency	**Yes**	**Yes**	No	No	Animals without the feeling of agency, disorganized schizophrenic patient, "implicit consciousness"
Stimulus expectancy and purposefulness	**Yes**	**Yes**	**Yes**	**Yes**	Agents, human beings, systems with epigenetic and phylogenetic motivations
No stimulus dependency but purposefulness	No	No	**Yes**	**Yes**	No real cases, as an ideal example HAL 9000 in "2001 Space Odyssey"

Table 1: The relation between the couple phenomenal experience/aboutness and the couple Agency/Intentionality.

2. Presence and phenomenal experience

Most subjects (both artificial and natural) can be classified on the basis of their behaviors. Different sensory, cognitive and motor capabilities determine how an agent is capable of interacting with the environment (Arkin 1999). However there is an aspect of agents which cannot be trivially reduced to their behaviors. Biological agents, especially humans, seem to have a direct experience of what happens to them; further they experience a feeling of ownership and responsibility in their phenomenal experience. Experience in human agents is something that is produced as an effect of their own history and existence. It is reasonably safe to assert that human beings, primates and most mammals have a capability for experiencing what happens to them. This aspect of their functioning is startling since there is no apparent explanation that justifies it. This capability of having an experience of bodily states and external events (and even abstract thoughts and concepts) has been called consciousness (in this context could be also called "presence")). Chalmers argued that there are two orders of problem in explaining a complex agent like a human being: one is the cognitive-behavioral-functional problem labeled the 'easy problem', the other is the phenomenal problem labeled the 'hard problem' (Chalmers 1996). That is not to say that the problem of cognition is in itself easy, but to say that the problem of the phenomenal side of the brain is somehow of a different order (Crick 1994; Edelman and Tononi 2000; Gell-Mann 2001; Hameroff 2001; Zeman 2001; Zlatev 2001; Churchland and Churchland 2002; Perruchet and Vinter 2002; Rees, Kreiman et al. 2002; Thau 2002).

The problem of consciousness has been defined the 'hard problem' for several reasons: (1) since consciousness is not related to behavior[1], it is difficult to see how it can be empirically tested; (2) there is no objective definition for the phenomenal experiences; (3) it is difficult to perform experiment *in vivo* on human beings.

On the other hand, it is difficult to underestimate the importance and the potential impact of any discovery related to consciousness. Human beings live literally immersed inside their own

[1] As it is not clear how phenomenal experience would influence behavior if it does at all.

consciousness. In a great number of fields, from medicine to entertainment industries, the capability of interfering with the subjective side of cognition would be of immense value (Kulli and Koch 1991; Pepperell and Punt 2000; Koch and Crick 2001; Zeki 2001; Manzotti, Sandini et al. 2003).

The capability of feeling is strongly related, if not the same, with what has been defined 'presence', or the feeling of being there. It can be argued that the concept of presence is conceivable only in a conscious agent. Presence can be defined as the consciousness (or the having of the experience) of being in the world: to be in a certain place, with a certain body, smelling, tasting, touching, hearing and seeing certain things. Experience, in the sense required by the concept of presence, is always synonymous of conscious experience. We cannot conceive a totally unconscious agent having experiences in any meaningful sense related to presence. For this reason the understanding of consciousness is mandatory for any research concerning presence.

An engineering approach to the problem of consciousness consists in trying to design and to build an artificial conscious being. Insofar this approach has been pursued only theoretically but there are no *a priori* reasons why it should not be successful (Steels 1995; Aleksander 1996; O'Brien and Opie 1997; Manzotti 1998; Schlagel 1999; Aleksander 2000; Martinoli, Holland et al. 2000; Togawa and Otsuka 2000; Aleksander 2001; Buttazzo 2001; Manzotti and Tagliasco 2001; Manzotti and Tagliasco 2002; Perruchet and Vinter 2002). Yet a formidable obstacle to an engineering approach to consciousness is represented by the fact that a conscious being is an entity that *feels*, not an entity that *does* something. And naturally before building a conscious being we need to have a theory for what a "conscious being" is.

3. Phenomenal experience and representation

In order to model conscious experience, we need to concentrate on three characteristics that, in combination, provide a working hypothesis about the nature of conscious experience. As conscious beings we know that experience (i) is always an experience of something i.e. every experience has a given content and (ii) experiences can be phenomenally distinguished on the basis of their content. For instance, between the experience of drinking lemonade or looking at a painting by Mondrian there is a difference in their content. Further, setting aside extreme forms of skepticism, it is an introspective part of our natural epistemic stance that (iii) the content experienced represents external events (Dennett 1969; Block 1988; Dennett 1988; Fodor 1990; Chalmers 1996; Bickhard 2001), at least in ordinary perception. In a sufficiently wide sense of the term, introspection commits us to the 'representationalist' standpoint. The claim that conscious experience occurs, that it has a phenomenal content, and that phenomenal content is tied to what happens, can be summed up as the thesis that to be conscious of something is to have a representation of that something.

On the basis of these considerations, we can put forth two hypotheses: (1) there are occurrences of events that correspond to contentful phenomenal experiences; (2) these occurrences represent something. According to this point of view, an event has a phenomenal content only if it has a representational role (Fodor 1981; Millikan 1984; Dretske 1993; Dretske 1995; Clark and Tornton 1997; Bickhard 1999). Hence conscious experience is the occurrence of events with phenomenal content and, by implication, with a representational role.

The notion of a representation is commonly taken to be more than an abbreviation for the claims (i) through (iii) above. Often, it is used as an explanatory notion with a meaning of its own – a representation is something that presents (or re-presents) something else. If we adopt an object-ontology and assume that the world is composed only of objects, then a representation is an object which presents another object. Yet how could we make sense of the kind of 'presentation' involved here, how could we think of an object 'referring' to another? One way to approach this question is by taking representation to be somehow effected by similarity. This approach – called the 'copy theory of representation' by Nelson Goodman (Goodman 1974) – however does not

get us anywhere since the identity principle establishes that an object is just itself, and that no object can be another object at the same time.

From within an object-ontological setting the only other way to understand representation is to view it as a relation. The representing entity and the represented one are linked by some kind of abstract relation (semantics, aboutness, or intentionality). However, this is in conflict with a commitment to physicalism which admits one relation only, namely, causality. Between events or objects there are only causal relations, there are no intentional, teleological, formal or semantic relations. In recent years there has been several innovative analysis of the probabilistic, manipulative, counterfactual and structural approaches to causation (Pearl 2000). It is not a coincidence that most of the attempts to naturalize semantics, perception and representation are based on some kind of causation (Grice 1961; Armstrong and Malcom 1984; Haybron 2000). However, if we accept that representation is any type of causation then representation becomes a ubiquitous phenomenon.

If we want to hold on to the insight that phenomenal experiences correspond to the occurrence of representations, while at once holding on to a commitment to physicalism, a physical interpretation of representation must be found. Obviously the classical dualist Cartesian model of representation as a relation between physical and mental items is not helpful here. However the dualistic model of representation (the representing item is something different from what it represents) is unavoidable given an underlying ontology that divides the world in separate objects.

Two entities are separate if their existences, at any given instant *t*, are mutually independent. (If you destroy your computer, nothing happens to mine at the same instant). All objects which do not stand in parts/whole relations are separate. If an object-ontology is accepted the dualist model is the only way to make sense coherently of the introspective data of conscious experience. As long as we describe ourselves as entities that are separate from what they are conscious of, we have taken on board the supposition that experience is some kind of duplication of the external world inside the internal domain of the subject.

In fact, the dualism of the external domain and the internal domain is independent of the additional qualification of the external as physical and the internal as mental. It is a contingent detail that in the XVII century the internal domain could only be assessed as mental while

nowadays we also have a neurophysiological description of this domain. The object-ontology representation always implies the existence of a dualist counterpart (either a copy or a *relatum*) of the world, no matter how that counterpart is characterized. In this respect current neuroscientists assign to the brain the same role Descartes attributed to the *res cogitans*. However, since it is quite clear that the brain taken as a physical object cannot contain copies or isomorphic relational counterparts of the external world (i.e. of something with completely different physical properties), it remains a mystery how the brain can represent the external world.

For an object-based account of representation the only viable strategy is to leave physicalism behind. One might suppose, then, that the brain contains *qualia* or pure phenomenal qualities, the modern version of the secondary qualities, which unfortunately can only be identified in terms of what they represent and thus they are a label for a problem rather than a solution. Alternatively, one might embrace a functionalist stance according to which brain states are representations of something in the world due the fact that they possess a functional role for the organism. The functional domain is an abstract domain of input/output relations built on top of, and always extraneous to, physical events.

In short, traditional object- or substance-ontologies are committed to explaining representations (and hence the mind) in a dualist fashion. Conversely a Cartesian dualist standpoint is naturally compatible with a substance-ontology (even though it does not entail the latter). Consider the following three claims:

a) the world consists of separate substances or objects
b) the mind represents the world (or the mind is equivalent to a set of representations of the world)
c) representations are different from what they represent (dualism)

As worked out consistently in XVII century metaphysics, a commitment to a) and b) entails a commitment to c). If the mind represents the world, and the world and the mind are made of separate entities, the mind must be a separate entity from the world; thus representations must be separate (and hence different) from what they represent.

However, is claim (a) indeed an *a priori* truth? We can reject the assumption that the world made of separate substances for at least the following three reasons. First, there is enough evidence

from particle physics to militate against classical objects or substances as a type of fundamental entities (Cramer 1988; Zohar 1990; Stapp 1998; Auletta 2000). Second, the claim is no logical necessity—as we shall see presently, a different ontology can be formulated. Third, besides the problem of representation there are a number of fundamental *ontological* difficulties arising for object ontologies (Seibt 1990).

If (a) is rejected, it is important to settle for the right type of alternative ontology. An event-based ontology, if not implemented carefully gets us into yet another type of difficulty. From a scientific point of view, the idea of a single event is a nomological absurdity. In contemporary science we cannot speak of anything which is not the object of an experiment, the result of a measurement, the target of an observation, or a postulated interaction whose results we observe. If something is not directly or indirectly observed, it is not part of what is empirically known. However, in order to be observed an event must be in relation with other events. It must somehow 'present' itself to other events. But then we must admit that in science there are no isolated events—events are derived entities, namely, interactions of processes. Isolated events or autonomous static objects are abstractions: like the Euclidean point or line, which are not part of the real world. Unfortunately these abstractions have been misunderstood as the real world, in the sense of Whitehead's 'fallacy of misplaced concreteness' (Whitehead 1925).

In the following we will argue that once we replace the traditional object-ontological framework with a suitable process-based alternative we can deal with the relation between the mind and the world without falling in the dualistic trap. We advocate the following alternative set of assumptions:

a) the world is an assembly of processes which are not necessarily separate
b) the mind represents the world (or the mind is equivalent to a set of representations of the world)
c) representations are not different from what they represent (monism)

Assumption (b) has remained unchanged, while (a) has been changed and, as a result, (c) also. In the following we will try to show how a process-ontology permits us to: i) account for representation without dualism, ii) treat the mind as a set of representations; iii) ensure true knowledge of the world without solipsism or dualism.

Experiments and measurements are ways of dealing with relations. If this were true – and we claim that there are no counterexamples falsifying it – each event is a relation in itself, and each relation is an event. This is completely different from the Cartesian approach. On the same basis we can observe that each event must have in itself what its content has to be, yet what it is can well be its content. To look for what could be the content for representing the event A, where to look if not in the event A itself? In this way it is possible to avoid distinguishing between the occurrence of an event as pure existence, the occurrence of the representation of that event and the relation constituted by that event. The reason is that they always occur conjointly. Hence they are the same thing seen from three different points of view. From a practical point of view it is reasonable to use one word to refer to this unique principle that can serve a triple role (as existence, as representation and as relation).

If phenomenal experiences correspond to the occurrence of representations, a physical candidate for representation has to be proposed and discussed. The classical Cartesian model of representation is not helpful here because it cannot be physically translated. Instead of thinking to the problem of representation in term of relation between separate entities or separate objects, representation can be addressed using a more event oriented ontology (Whitehead 1927/1978; Davidson 1980; Steward 1996; Manzotti and Tagliasco 2001; Manzotti 2003). By using an event- or a process-ontology a different picture stems out. A tentative causal relation between events, something which could be considered a representational process, can be looked for.

4. A candidate for phenomenal experience and aboutness: the onphene

If presence is the result of the occurrence of a large set of representations and if such representations are a kind of causal entanglement, a candidate for this kind of causal process has to be proposed. It is useful to distinguish two kinds of representation here: i) pure phenomenal representations which occur independently of the existence of an agent that recognizes them as its own representations; ii) representations that occur as the result of expectancy or a decision of an agent that uses them as expression of its existence. The first case is, in a way, simpler but nearer to Chalmer's hard problem of phenomenal experience. Further, the first case has to be logically explained in order to be able to explain the phenomenal experience related to agents. We will deal with the first case in this section and with the latter in section 5.

It is an empirical fact that not all causal relations seem to show a representational efficacy. Using our own phenomenal experience as an empirical background, it is easy to see that many events which have some kind of causal influence on our brain states are nonetheless representationally neutral. For instance all those chemical and physical events which are located between what we see and our neural activity are complete transparent to us; they do not contribute in any way to the phenomenal content of our experience, thus they are not represented. The same is true for those events that are causally responsible for the occurrence of the objects we see but that do not take part in our phenomenal experience.

As a paradigmatic example, a familiar case can be used: a TV screen. Although we are phenomenally conscious of what happens on the surface of the screen, we are totally oblivious of what happens earlier to produce the images we see. In other words, there is a long causal chain of causes of which we represent only one in a particular position. We will propose here that the particular event that is phenomenally represented in our experience is the event which is entangled in a unique and distinguished causal relation with our brain events and with our development.

There are physical phenomena that can be used to give a rough idea about what could be this kind of causal relation. They are physical phenomena in which the physical continuity between

the represented object/event and the representing object/event is much less evident than the usual. The rainbow is one of these examples and it will be used here as a paradigmatic example.
When the sun is sufficiently low on the horizon and sheds its rays at the right angle with respect to a cloud (or a sufficiently large volume of drops of water in suspension in the air) an observer (either a human being or a camera) in a position between the sun and the drops of water can see a beautiful arch with all colors of the spectrum. Each drop of water reflects the sunlight in the same manner, yet only those drops, which are in a particular geometrical relation between the observer position and the direction of the sunrays, are seen as part of the rainbow. The position of the rainbow depends on the position of the observer as well.
The rainbow is an object that cannot be defined in any meaningful sense independently of the point of view from which it is seen. In this sense the rainbow, although it is constituted by a set of physical drops of water reflecting the light in a certain way, cannot be defined without knowing where and how it will be seen. For instance, it is not possible to fly under a rainbow or to sidestep it. The rainbow will move accordingly to the position of the observer. Furthermore it is possible to argue that the rainbow is a private physical phenomenon since two different observers always see two physically (even slightly) different rainbows. Since two separate observers occupy two different positions in space, they select different rays of light and thus different drops. Every observer sees a different rainbow. This can be verified by moving rapidly (for instance driving a car): the rainbow seems to be moving together with the observer. In reality by changing its position, the observer modifies the properties of his/her point of view and therefore continuously selects new sets of drops. A fundamental question arises at this point. Does the rainbow exist independently of the act of observation? We claim the answer is no. In fact, if there were no eyes looking at the rainbow (or a camera), that phenomenon that we call rainbow would not produce any effect. Furthermore the selection of a particular set of drops as constituting the rainbow cannot be done independently from the point of view. As we have seen, changing the position of the observer changes also the drops of water that are part of the rainbow. Even if it is possible to argue that an expert physicist knowing the position of each drop, of the sun and the point of view could calculate the projection of the rainbow on the observer's retina, such a calculation requires as an essential element the knowledge about the point of view, which cannot be omitted. The rainbow is not there, as a bridge of stone. The

rainbow occurs only when it is seen. It is a case in which there is no distinction between the cause and the effect. The cause (the arch of drops) is not there as a distinctive whole until it produces an effect (for instance a projection in the observer's retina). The rainbow is different from other everyday physical phenomena. For instance a building or a crystal can be defined independently of an observer by taking into account all those particles, which have a particular property. In the case of the rainbow, among the essential properties that a particle has to possess in order to be considered part of the phenomenon, there is the relation with the observer. The point is that we cannot separate the cause from the event, as well as we cannot separate the events from their relation. The effect is responsible for the existence of the cause.

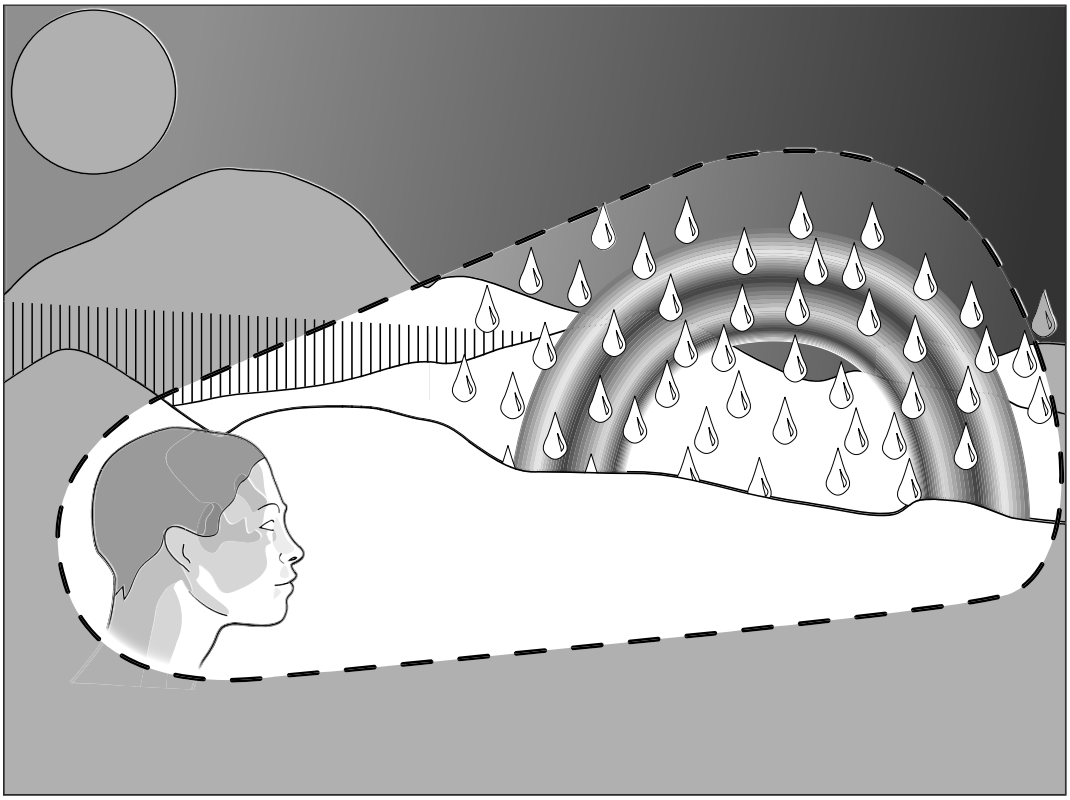

Figure 1: A rainbow.

In a similar fashion, in conscious representations a similar structure is recognizable. For instance, if we have six dots on a wall they are not producing any physical effect as a result of their joint existence. However, if they become a perceptual object for an observer they produce an effect that is responsible for their existence as a whole. The effect (the perception in the observer) is responsible for the existence of the cause (a particular pattern) as a whole. If there were no cognitive system it will be extremely implausible that they would produce any effect as a whole. Each dot would exist by its own. However if a human being were in front of them, she could recognize them as a whole. If she had not seen anything like that before, she would not have had the corresponding phenomenal content until she saw it for the first time. Why was she able to see it for the first time? Among all possible reasons, there is at least one that was necessary: the six dots were there to be seen. It was a startling moment. The six dots produce, for the first time, an effect because of their joint existence. For the first time they produced an effect as a whole. If existing is the same as "producing effects" (something which does not produce effects is something which does not exist), the six dots as a whole come into existence when they produced an effect, and thus when she *saw* the six-dots-whole for the first time. There is no way for physically defining the existence of the six-dots-whole without making use of the cognitive system that reacts to them. The six-dots-whole literally comes into existence when the six dots produce an effect. Before they were just an arbitrary possibility of universe with no more existence than an arbitrary cluster composed, for instance, of a pen on my desk, a drop of water in the Caspian sea, a helium atom inside Alpha Centauri. In the same way, arbitrary clusters of entities are just "*an abstraction/remaining a perpetual possibility/only in a world of speculation*". On the contrary when a cluster of entities produces an effect *as a whole* it comes into existence. The unity of the object is a consequence of its capability of producing an effect as a whole. For many clusters (or objects) they can produce a joint effect only by means of the interaction with a cognitive agent.

A similar structure is recognizable in most perceptual events. The rationale that emerges is that whenever we have a mental representation there is no real distinction between the represented event and the representing one. Both occur conjointly and as a unity: a natural unity. This unity

can be seen as a natural candidate for representation since it contains what it has to represent, there is no need to suppose a duplication of reality as in the Cartesian framework. On the other hand all the problems deriving from theories advocating identity between brain processes and mind processes are solved since the physical substratum of the mind is no more just a neural activity located in the brain and separated from what it should refer to; the physical substratum of the representations corresponds to a process which starts in the external world and ends in the brain. Finally, the problems of the functionalistic approach are solved. The process, which is responsible for representation, is no more an abstract functional level in search of its own interpreter; the responsible event is a concrete event that is identical with that part of the external world, which should be represented.

In this way a mind no longer correspond with an emergent property of a system duplicating the external reality in some internal code. The mind is *enlarged* to that part of reality that the mind represents. In fact representation is no more a *re-presentation* but it is just a *presentation* (a synonymous of occurrence to us) of reality inside a system of events.

This process could be called in several different ways. In literature there have been similar candidates as Whitehead's prehension (Whitehead 1933), reciprocal causation (Newman 1988; Hausman 1998), intentional relation (Manzotti 2000; Manzotti 2001). Instead of the term "intentional relation", we will use here a new term in order to avoid any unnecessary and misleading connotations with other standpoints: **onphene**. Onphene is a new word composed by *ontos* (existence), *phenomenon* (representation), and *episteme* (being in relation with). The choice of these three terms comes from the fact that the onphene is a physical process (that is, something that exists), it corresponds to a phenomenal content (thus a representation), and finally it puts in relation external reality with the conscious mind (thus being in relation with).

Thanks to the onphenes, instead of having to deal painfully with the unsolvable obscurities of the problem of representation, a unified theory of the mind, the body and the environment can be developed. According to this theory the distinction between a representing brain which has to carry the burden of the Cartesian soul and a represented body (plus environment) is arbitrary and unnecessary. The neural activity in the brain does not have to access mysteriously the content of the external world, developing representations. Each representation corresponds to a process that is physically continuous from the external event (the environment) to the brain. In this fashion

the mind does not correspond to a property of the brain, or to a new emergent property of the brain: the mind corresponds to that part of reality, which becomes entangled in the activity of the subject. The distinction between the mind and the world is revealed as a misleading useless metaphor. On the other hand, in a strict sense, by definition nobody has ever experienced anything outside of its own mind and insofar no pure disembodied mind has ever been described. In this way representation and relation are both resolved, at the same time, in the less troublesome category of existence. To represent something means to be that something. There is no need for the activity of the brain to possess a dual state, an invisible property corresponding to a phenomenal experience even more mysteriously linked to the external events. The identity principle is safe: if A has to be represented it can be done only by A itself which must become part of the subject. In the same way, if the mind has to access a series of events that mind must be that series of events. Hence the mind is metaphorically enlarged to a larger physical domain than the brain alone at a certain time *t*. All those events, which constitute the content of the conscious mind, are physically part of the mind. There is no more dualism: there is just one reality. There are no more representations as something different from the represented events: there are just events that constitute processes. When we feel something is because the process we are is extended to comprehend those new events we experience. If something is the content of our experience, it is never just a phenomenal experience opposed to a real world. Events that previously had no part in our developmental history become entangled in our internal process (or equivalently we enlarge to that something). The traditional boundary between an internal domain and an external domain is resolved in favor of a broader idea of occurrence of a process. There is no more an external occurrence of something that must be somehow reproduced in the inside of the soul/mind/brain: there is just an occurrence of a process that is both the outside and the inside.

The onphene is thus a particular kind of physical process. As a working hypothesis we propose that it could be the carrier of those phenomenal contents that constitute conscious experiences. According to this hypothesis a moment of presence, a moment of conscious experience, would then be constituted by a series of processes (onphenes) going from the consciously perceived events (color, shapes, smells, tastes, and so on) to the observer brain. Of course the brain's architecture must be such that onphenes can terminate in the brain itself. In the following

paragraph we will argue that to host onphenes the architecture must be capable of developing new motivations.

In order to understand what constitutes a sense of 'being there', or specifically presence, it is important to understand the physical and causal basis of aboutness and thus of phenomenal experience. It has been argued in some of our previous works (Manzotti 2000; Manzotti 2001; Manzotti and Tagliasco 2001; Manzotti 2002) that the sense of being-there depends on the sense of agency which is strongly related to intentionality and development, where representations depend on the capability of developing epigenetically new motivations. Developing motivations is important because it is the basis of the development of an agent. Motivations are the means by which the history of the system becomes the cause of the future actions and development of the agent itself, thereby constituting the foundation for intentionality. We will concentrate here on how (and if) intentionality can be embodied into artificial agents: that is robots.

The problem of how to create a coherent representation that implements intentionality and that can be embodied within an artificial agent is the core of the literature discussed here. Although, the issue of representation remains problematic, we can accept the standpoint according to which to represent something means to possess an 'arrow' pointing at that something. In this sense, the term 'representation' is synonymous with aboutness: the capability to refer to. For example, perception itself is a form of representation. When subjects perceive the world, they represent it.

Certain critics of Artificial Intelligence believe that aboutness is an inherently biological phenomenon, in that it is inherently an embodied phenomenon, and a product of natural selection (Dreyfus and Hall 1982; Dulany 1999; Byrne 2001; Crick and Koch 2003). That is, there is currently no artificial system that has its own aboutness, given that no artificial system has been able to duplicate the conditions of embodiment that determine intentionality in biological organisms. To date, there have been few research efforts (mostly philosophical) that have addressed the issue of aboutness as it relates to embodiment (Searle 1980; Dennett 1987; Harnad 1990; Dreyfus and NetLibrary Inc. 1992; Bickhard 1998; Metzinger 2003). In fact, Artificial Intelligence has addressed mostly the problem of how to modify behavior in order to achieve a fixed set of goals such as sensori-motor coordination. Machine learning techniques such as

reinforcement learning or supervised learning algorithms[2] were employed (Clark and Toribio 1998; Sutton and Barto 1998; Pfeifer 1999/2001).

Several authors (Agazzi 1981; Dennett 1991; Dennett 1995; Dennett 1996; Dennett 1998; Zlatev 2001) stated that there is no particular reason to assume that aboutness is an exclusively biological property, which is an argument adopted by John Searle (Searle 1980; Searle 1983; Searle 1984; Harnad 1987; Harnad 1990; Searle 1992; Harnad 1995). That is, if aboutness were to be embodied in an artificial entity that was capable of human-like interaction and development within its environment, then such an entity would constitute a system capable of meaning. This conjecture is supported through a thought experiment inspired from Wittgenstein's work (Wittgenstein 1953/1995). The main counter argument (Searle 1980) is that aboutness in human beings is a product of the causal features of the brain, and that certain brain processes are needed for intentionality. This argument extends from Searle's celebrated "Chinese Room" argument (Searle 1983) in which the symbol grounding problem is referred to as the problem of intrinsic meaning. The argument challenges the core assumption of symbolic Artificial Intelligence that in order for a symbolic system to be able to generate behavior indistinguishable from that of a person the system must have a mind.

In that debate (only briefly reported here), the proposed theory aims at solving the problem of aboutness with a kind of causal process which should be responsible for the development of beings (natural and artificial). In this respect, aboutness would not be an exclusive property of biological material neither would be the result of a third-person interpretation. Aboutness would be a way of linking together events belonging to the system's history and development.

[2] Clearly the literature on machine learning is extremely large and this is an extreme simplification.

5. Motivations and intentionality

From a causal point of view, an onphene is a peculiar process. It links the occurrence of the cause with the occurrence of the effect. It corresponds to a situation in which the occurrence of an event produces not only an effect but also the condition for the occurrence of the causal relation between that kind of cause and that kind of effect.

There are six dots on a wall distributed along a circle at equal distances (they could be interpreted as a hexagon). If nobody recognizes them as a whole, they are just six dots. They do not constitute a whole in any meaningful sense: they do not produce any effect as a whole. If an observer recognizes them as a whole, for the first time, they produce an effect in the observer's brain as a result of their joint existence (in fact the observer could have shouted "I see a hexagon!"). Is the effect inside the observer's brain and the "six dots as a whole" outside? No, because before the act of observation there were "six dots as a whole" neither inside nor outside the observer's brain. The whole, which is the content of the observer's perception, is at the same time inside and outside. It occurs because the act of observation and what is observed are a unity. There is no inner event and outer event, just one event: the physical process that is responsible both for the hexagon and its recognition. The six dots as a whole arise when they produces a joint effect. Intuitively, it (the six dots as a whole) is the result of its effect (the recognition of a hexagon in the observer's brain).

We propose to consider an onphene as a process in which the cause and the effect are mutually necessary (as in the rainbow's example). It is something which makes a difference in the reality, something which occurs. There is no existence without producing effects; there is no representation which is not something that occurs; there is no occurrence which is not the content of itself (the representation of itself); there is no relation which does not makes a difference; there is no representation that does not produces effects; there is no relation which is not the bearer of a representation.

How are these concepts related to a theory of mind? By making use of onphenes and events, instead of representations, it is possible to build a theory of the mind that does not require representations and that, at the same time, is compatible with the existence of phenomenal

experiences. The reason for this is that representations were mainly needed to re-present in the mind what is happening elsewhere. By making use of the onphenes, the mind is just a piece of reality which has its own coherence and unity. The coherence and unity are just a result of the structure of the onphene. The mind, according to the onphene-based standpoint, does not need to emerge, does not need to be different from reality. The mind is the reality, a piece of it, at least. How a part of reality becomes a unity is explained by epigenetic processes that link together different causal chains into a unified causal process which is usually referred to as the "will" of an individual. Separate phylogenetically induced motivations and goals become more and more merged into a unified epigenetic giant causal process which is the subject.

According to this point of view, in some sense, the conscious mind is the process of development itself, albeit based on a particular kind of causal relation. The conscious mind is a kind of occurrence, a way of taking place of events: cognitive and sensori-motor development is necessary to make it happen. The conscious mind is a process and the physical structure of the agent is necessary to make such a process take place.

The agent is more than just a physical structure. It is a history capable of propagating itself towards the future by means of motivations and purposefulness. In a physical agent we must distinguish two different classes of motivations: phylogenetic motivations and epigenetic motivations. The former ones are hard-wired in the structure (in a biological being they are innate in the genetic code; in an artificial system they are explicitly designed). The latter are derived from experience and thus are caused by past experience. Among the most important phylogenetic motivations we must mention curiosity which is a general tendency of the system to extend itself to new classes of stimuli. Other phylogenetic motivations can be more goal oriented (like looking at faces, supporting social interactions, paying attention to human voices, to brightly colored objects, moving objects, etc.). On the other hand, epigenetic motivations are the result of the interaction with the environment. They must derive from something that entered in the past experience of the system. In this chapter we will concentrate on this second kind of motivations, since they are what constitute the individual history of an agent.

If a system had only phylogenetic motivations its behavioral structure would be completely innate and fixed. It could not adapt to the environment and it would not have an individual developmental history. If a system had only ontogenetic motivations it would only repeat what

happens to it (it would lack curiosity for instance). It is the interplay between these two types of motivations that permits to the individual agent to emerge as an individual flow of events linked by the development of epigenetic motivations.

An epigenetic motivation is here defined as a process whose probability to occur is increased because of the existence of a physical structure and because of its own occurrence. Epigenetic motivations and phenomenal experiences are related, according to our hypotheses, because they share the same causal structure. Both are processes and both are the same kind of processes. Further they are both constitutive of what the subject is. A motivation is here defined exactly in the same way. A motivation is a process which the system repeats again and again. Goals are the result of habit and produces habits. The biggest difference between the two is that a phenomenal experience does not to produce necessarily a stable modification in the subject's structure, while the motivation must propagate its effects through time thanks to a modification in the subjects' structure. A phenomenal experience is an onphene that can occur only once in history, like a particular rainbow. A motivation is a process that has to link different moments in the subject's life thereby constituting an agent.

To have an epigenetic motivation is to have an intention. To have an intention of getting something means to be in relation with an event. We can make an example of an agent capable of developing new motivation as a result of its experience. An agent is exposed to the presence of a person named Susan, and as a result the agent has got Susan in his field of view. If this process becomes a motivation, the agent will behave with the goal of repeating as much as possible the process of seeing Susan. The process has already taken place. Its taking place modified the structure of the agent in such a way that there is 'having Susan in the field of view' among its goals. The occurrence of the process 'seeing Susan' has increased the possibility of its own repetition. This increase of the probability of the occurrence of a process is due only to the occurrence of a certain event. If Susan would have not shown up, the system would have not modeled its criteria around her visual appearance. A caveat; what increases its possibility to happen again is not the appearance of Susan (which depends only on Susan), but the process of Susan being seen and recognized by the system. In the case the system incorporates 'seeing Susan' among its goals, will act in the future accordingly in order to keep Susan at the centre of

its visual field, thereby increasing the probability for that kind of process to occur again and again.

Many examples can be made. Playing a sport increases the probability of doing it again. There is nothing like playing tennis in order to become fond of playing tennis. Seeing someone is a necessary condition in order to be willing to spend more time with that person (spending more time will naturally translate into repeating those physical processes in which that person or her/his action are perceived by us). Of course all this processes, in order to become part of the motivations of the agent have to be compatible with its previous structure. For instance, eating rotten food does not become a motivation, although there are proofs that in extreme cases abnormal situations might become part of the subject's structure (imprisonment and such). An intentional structure must therefore be able to comply with the largest possible number of external events. As a rule of thumb it must be constituted by three separate modules: a module for building categories and being able to distinguish between different kinds of events, a phylogenetic module for bootstrapping the system at the beginning and orienting it towards certain classes of events, an epigenetic module who will contain and produce the new criteria.

Basically, an epigenetic motivation has the same causal structure of an onphene: a reciprocal causation structure. In fact, it is an onphene. However it is an onphene that together with other onphenes propagates itself in the future (it increases the probability of its own occurrence). If some action had proved itself successful, it will be repeated.

Between aboutness and intentionality there is a strong relationship. They share the same causal structure. They are both onphenes. Yet a motivation takes place inside a more complex structure of causal relation that we call 'agent'. By using this approach we can differentiate between phenomenal experiences which require simple aboutness (reciprocal causation *per se* as in the case of the rainbow) and full representations of the world which require intentionality (reciprocal causation as a modification of the agent history, namely motivation).

Here is a table that recapitulates the main points of this theory.

Event	something which takes place
Process	when an event is caused by another event
Reciprocal causation	when the occurrence of the effect is responsible for the occurrence of the cause.
Onphene	a physical process in which cause and effect are mutually necessary. A process with reciprocal causation.
Represented event or object	an onphene from the point of view of the cause
Representing event or representation	an onphene from the point of view of the effect
Phenomenal experience	an onphene
Motivation	an onphene that is part of a history of onphenes in such a way as to repeat itself or rather a process whose probability to occur is increased because of the existence of a physical structure and because of its own occurrence
Action	an event provoked by a subject as a result of a motivation

Table 2: A list of terms used and their definitions

We can summarize our rationale in the following:

- The conscious mind is a unified collection of phenomenal experiences.
- A phenomenal experience is something which has content; it can be considered a representation of something.
- Unity and representation can be helpful to understand consciousness.
- There are two kinds of unity: the unity of the subject that acts and the unity of the object that is perceived.
- A representation is something which has unity and that is capable of accessing something else.
- This general capability of accessing something else is what has been called traditionally intentionality or aboutness.
- The classic theories of representation struggle against a dualistic conception of reality: on one side the thing to be represented, on the other side the representing thing/event/state.
- An alternative approach can be proposed. A representation is not a thing different from the represented thing.
- The represented thing and the representation are two different ways of looking at the same process: a continuous physical process. They do not exist separately.
- This kind of continuous process occurs whenever a reciprocal causal relation is instantiated.
- The reciprocal causal relation is called onphene because of its role in shaping the experienced reality and endorsing intentionality
- Onphenes endorses the unity of the object and the subject.
- Phenomenal experience exhibits aboutness and corresponds to onphene.
- The onphene is typical during cognitive epigenetic development where experiences give rise to capabilities and motivations.
- There are two kinds of motivations: phylogenetic and epigenetic.
- The development of new epigenetic motivations is crucial because it is through the development of new motivations that the system unifies itself.
- A motivation is something which collects a large group of other intentional relations. Furthermore a motivation guarantees the temporal continuity of behavior. Therefore goals permits to glue together different intentional relations (or experiences).

- A motivation is an onphene that, being part of a flow of other onphene (the agent), is capable of repeating itself by increasing the probability of its own occurrence.
- Motivations endorse the unity of the subject and intentionality.
- A representation is an onphene in an agent.
- A conscious mind, an agent, is a group of intentional relations or onphenes which occur as a unity since they are the cause of at least a motivation.

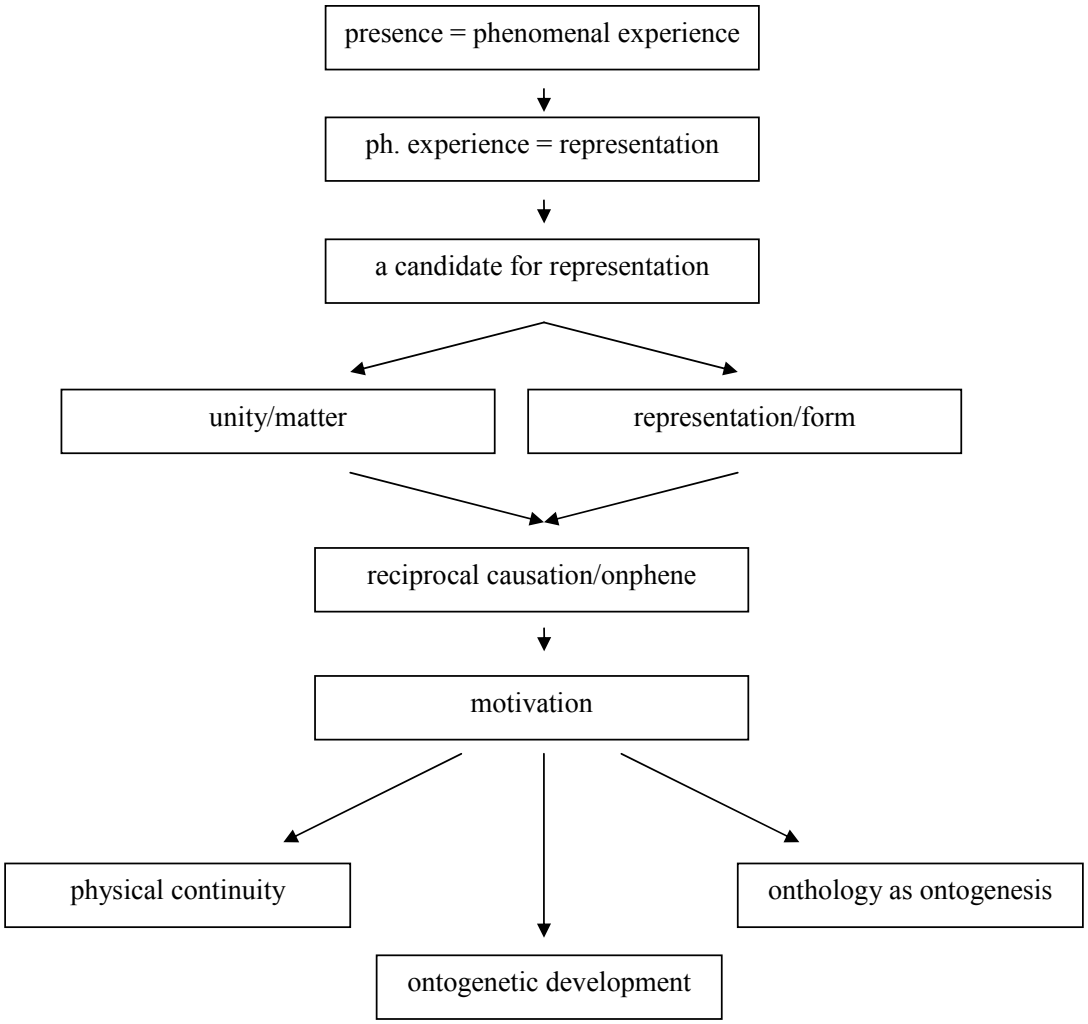

Figure 2: Summary of conceptual relations of the proposed theory.

6. The end precedes the beginning

The resulting theory of consciousness can be considered a theory of the phenomenal consciousness based on the nature of processes involved in development. We say phenomenal consciousness since some author distinguished between an access consciousness, which is not necessarily phenomenal, and a phenomenal consciousness (Block 1997; Chalmers 1997). We refer to the latter. However it is a scientific and empirical theory that addresses explicitly the problem of phenomenal experience and makes prediction on the necessary physical structure of events. Thus it can be verified or falsified. It is a theory that addresses the nature of the feeling of being there, or presence. It is a theory that presents the subject as the result of the self organization of the environment in which the developing agent acts. If the world were devoid of subjects, it will be a much different place in causal terms. Everything that happens will happen only once. If shapes or drawing were to be carved in stone by the capricious whirlpool of air, their appearance will not increase the probability of their repetition. On the other hand, in a world populated by subjects, every process that becomes part of their structure increases its probability of repeating itself. Events propagate themselves in a new and interesting way. As a result shapes, colors, behaviors, actions are repeated countless times all over the Earth both as phenomenal experiences and as the result of actions that have to be phenomenally perceived. According to our hypotheses, the subject can be seen as a set of processes that becomes progressively integrated during development.

We can recapitulate our proposal for what regards presence as follows: the content of the conscious mind is not ‘inside the head’; it is how the world is organized because of the existence of the agent. The agent is nothing more than the collections of processes or onphenes that take part into its development. Phenomenal experiences are not mysterious and ineffable subjective qualia, they correspond to physical processes.

There is a series of compatible empirical facts that can be quoted here:

1) Subjects that lack the direct contact with a category of physical events (born blind people for instance) also lack the corresponding phenomenal conscious content (born blind

people miss the phenomenal content of color and visual phenomenal conscious experience).

   a. Exception: the only adverse evidence is the not so clear description of a patient with congenitally absent limbs that reported a proprioceptive phenomenal experience of them (Brugger, Kollias et al. 2000). Nevertheless, in that case: i) there was not a totally missing sensory modalities as in the case of born blind people; ii) there are other possible psychological explanations; iii) data can be questioned.
   b. Rationale: phenomenal content comes from the outside. It is a consequence of the direct physical continuity with the external environment during development: it is not internally produced.

2) Localized brain damage is often content specific. Whenever a certain brain area is locally damaged the most frequent result is a specific functional impairment. Functional impairment is almost always accompanied with the disappearance of the corresponding phenomenal content (Moutoussis and Zeki 1997; Zeki 2001).
   a. Rationale: each processing unity is somehow connected with a corresponding content.
3) Although it can be claimed that there are functional processing that do not correspond to any phenomenal conscious content, all perceptual phenomenal content must be linked to corresponding functional processing (Zeki 2003).
   a. Rationale: the functional processing of the brain is necessary, albeit perhaps not sufficient, to have phenomenal conscious experience.
4) Every cortical area seems to be capable of developing the function of every other cortical area as long as it has enough time and it's actually activated and it can do it during development. For instance, different cortical areas can develop language capabilities. Infants with extended left brain damage develop perfectly normal language skill in the right hemisphere (Finlay, Darlington et al. 2001; Nelson and Luciana 2001).
   a. Rationale: there is no hardwired genetically-determined algorithm in the brain. In the brain there are the tools to develop algorithm, not the algorithms in them.

5) The capability of the cortex seem to be accomplished through some kind of architectural universal principle applied many times to different signals, sources of information, internally derived signals. There is no evidence of different neural structure for sensation, perception, beliefs production, and thinking. They seem to be processed by a similar structure which could be held responsible for the processing of different kind of content.
    a. Rationale: same structural principle, applied to different incoming signal, gives raise to different phenomenal content and different functional processing.
6) The fine neural structure of the brain cannot be defined by a genetic blueprint for a series of reasons: i) there is not enough space in the genetic code to store all possible neural connections; ii) the genetic incompleteness of the brain is one of the key to understand its efficacy, its variability, and its adaptability to different environmental conditions. Furthermore, there is evidence that the particular cortical content depends on which sort of data it is exposed to (Zeki 2001; Rees, Kreiman et al. 2002).
    a. Rationale: since the brain cannot totally be a result of genetic based development, there must be some basic principle which can define the brain structure on the basis of the external events.
7) Lack of physical candidate for unity in the brain: physical brain activity does not have any kind of strong physical unity. As it has been showed by Cajal, brain activity corresponds to a series of spatially and temporally separated processes. There is no physical criterion to glue them all together. All explanation for binding have insofar proved to be unsuccessful.
    a. Rationale: the unity of the mind (which can be split in two: the unity of the subject and the unity of the perceived object) does not have any physical candidate inside the brain.
8) The phenomenal conscious content cannot be either outside or inside the brain. It cannot be outside since it will be physically (temporally and spatially) separated and different from the neural events going on in the brain. It cannot be inside since the brain processes don’t have the necessary properties to be considered good candidates for the phenomenal conscious content (they do not have unity, they do not have color, shape, taste, etc.).

These elements, albeit not a conclusive proof, seems to be coherent with our hypothesis: the conscious mind is not a property of the brain itself (or whatever information processing structure): the conscious mind is a way of referring to an historic collections of onphenes involved in the agent development.

At the basis of all of this there is the intentional relation of onphene, which is – as we described before – a physical process that can play the role both of the represented event as well as of its representation.

7. References

Agazzi, E. (1981). "Intentionality and Artificial Intelligence." *Epistemologia* IV: 195-228.
Aleksander, I. (1996). *Impossible Minds: My Neurons, My Consciousness*. London, Imperial College Press.
Aleksander, I. (2000). *How to Build a Mind*. London, Weidenfeld & Nicolson.
Aleksander, I. (2001). "The Self 'out there'." *Nature* 413: 23.
Arkin, R. C. (1999). *Behavior-Based Robotics*. Cambridge (Mass), MIT Press.
Armstrong, D. M. and N. Malcom (1984). *Consciousness and Causality: A Debate on the Nature of Mind*. Oxford, Blackwell.
Bickhard, M. (1998). "Levels of Representationality." *Journal of Experimental and Theoretical Artificial Intelligence* 10(2): 179-215.
Bickhard, M. (1999). Representation In Natural and Artificial Agents. *Semiosis. Evolution. Energy: Towards a Reconceptualization of the Sign*. E. Taborsky. Aachen, Shaker Verlag.
Bickhard, M. (2001). The Emergence of Contentful Experience. *What Should be Computed to Understand and Model Brain Function?* T. Kitamura. Singapore, World Scientific.
Block, N. (1988). What Narrow Content is Not. *Meaning in Mind: Fodor and his Critics*. B. Loewer and G. Rey. Oxford, Blackwell.
Block, N. (1997). On a Confusion about a Function of Consciousness. *The Nature of Consciousness*. O. Flanagan and G. Guzeldere. Cambridge (Mass), MIT Press.
Brentano, F. (1874/1973). *Psychology From an Empirical Standpoint*. London, Routledge & Kegan Paul.
Brugger, P., S. S. Kollias, et al. (2000). "Beyond re-membering: Phantom sensations of congenitally absent limbs." *Proceedings National Academy of Sciences* 97(11): 6167-6172.
Buttazzo, G. (2001). "Artificial Consciousness: Utopia or Real Possibility." *Spectrum IEEE Computer* 18: 24-30.
Byrne, A. (2001). "Intentionalism defended." *Philosophical Review* 110: 199-240.
Chalmers, D. (1996). The Components of Content. *Philosophy of Mind: Classical and Contemporary Readings*. D. Chalmers. Oxford, Oxford University Press**:** 608-633.
Chalmers, D. (1997). Availability: The Cognitive Basis of Experience. *The Nature of Consciousness*. N. Block, O. Flanagan and G. Guzeldere. Cambridge (Mass), MIT Press**:** 421-423.
Chalmers, D. J. (1996). *The Conscious Mind: in Search of a Fundamental Theory*. New York, Oxford University Press.
Churchland, P. S. and P. M. Churchland (2002). "Neural Worlds and Real Worlds." *Nature Reviews Neuroscience* 3: 903-907.
Clark, A. and J. Toribio (1998). *Machine intelligence: perspectives on the computational model*. New York, Garland.

Clark, A. and C. Tornton (1997). "Trading spaces: Computation, representation and the limits of uninformed learning." *Behavioral and Brain Sciences* 20: 57-90.
Crick, F. (1994). *The Astonishing Hypothesis: the Scientific Search for the Soul*. New York, Touchstone.
Crick, F. and C. Koch (2003). "A framework for consciousness." *Nature Neuroscience* 6(2): 119-126.
Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, Harcourt Brace.
Davidson, D. (1980). *Essays on Actions and Events*. Oxford, Clarendon Press.
Dennett, D. C. (1969). *Content and consciousness*. London, Routledge & Kegan Paul.
Dennett, D. C. (1987). *The intentional stance*. Cambridge (Mass), MIT Press.
Dennett, D. C. (1988). Quining Qualia. *Consciousness in Contemporary Science*. A. Marcel and E. Bisiach. Oxford, Oxford University Press.
Dennett, D. C. (1991). *Consciousness explained*. Boston, Little Brown and Co.
Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York, Simon & Schuster.
Dennett, D. C. (1996). *Kinds of minds: toward an understanding of consciousness*. New York, Basic Books.
Dennett, D. C. (1998). *Brainchildren: essays on designing minds*. Cambridge (Mass), MIT Press.
Dretske, F. (1993). "Conscious Experience." *Mind* 102(406): 263-283.
Dretske, F. (1995). *Naturalizing the Mind*. Cambridge (Mass), MIT Press.
Dreyfus, H. L. and H. Hall (1982). *Husserl, intentionality, and cognitive science*. Cambridge (Mass), MIT Press.
Dreyfus, H. L. and NetLibrary Inc. (1992). What computers still can't do a critique of artificial reason. Cambridge (Mass), MIT Press.
Dulany, D. E. (1999). "Consciousness, connectionism, and intentionality." *Behavioral and Brain Sciences* 22: 154-155.
Edelman, G. M. and G. Tononi (2000). *A Universe of Consciousness. How Matter Becomes Imagination*. London, Allen Lane.
Finlay, B., R. Darlington, et al. (2001). "Developmental structure in brain evolution." *Behavioral and Brain Sciences* 24: 264-308.
Fodor, J. (1998). *Concepts - Where Cognitive Science went Wrong*. Oxford, Oxford University Press.
Fodor, J. A. (1981). *Representations: philosophical essays on the foundations of cognitive science*. Cambridge (Mass), MIT Press.
Fodor, J. A. (1990). *A theory of content and other essays*. Cambridge (Mass), MIT Press.
Gallese, V. and T. Metzinger (2003). "Motor ontology: The representational reality of goals, actions and selves." *Philosophical Psychology*.
Gell-Mann, M. (2001). "Consciousness, reduction, and emergence. Some remarks." *Annals of the New York Academy of Sciences* 929: 41-49.
Goodman, N. (1974). *Language of Art*.
Grice, P. (1961). *The Causal Theory of Perception*.
Hameroff, S. R. (2001). "Consciousness, the brain, and spacetime geometry." *Annals of the New York Academy of Sciences* 929: 74-104.
Harnad, S., Ed. (1987). *Categorical perception: the groundwork of cognition*.

Harnad, S. (1990). "The Symbol Grounding Problem." *Physica* D(42): 335-346.
Harnad, S. (1995). Grounding symbolic capacity in robotic capacity. *"Artificial Route" to "Artificial Intelligence": Building Situated Embodied Agents*. L. Steels and R. A. Brooks. New York, Erlbaum.
Hausman, D. M. (1998). *Causal Asymmetries*. Cambridge, Cambridge University Press.
Haybron, D. M. (2000). "The Causal and Explanatory Role of Information Stored in Connectionist Networks." *Minds and Machines* 10: 361-380.
Koch, C. and F. Crick (2001). "The zombie within." *Nature*(411): 893-893.
Kulli, J. and C. Koch (1991). "Does anesthesia cause loss of consciousness?" *Trends In Neurosciences* 14(1): 6-10.
Manzotti, R. (1998). *Emotions and learning in a developing robot*. Emotions, Qualia and Consciousness, Casamicciola, Napoli (Italy), World Scientific.
Manzotti, R. (2000). *Intentionalizing nature*. Tucson 2000, Tucson, Imprint Academic.
Manzotti, R. (2001). Intentional robots. The design of a goal seeking, environment driven, agent. *DIST*. Genova, University of Genoa.
Manzotti, R. (2002). "Why Physicalism and Constructivism Will Never be Able to Understand the Mind." *Psycoloquy* 13(6).
Manzotti, R. (2003). "A process based architecture for an artificial conscious being." *Axiomathes*: in Press.
Manzotti, R., G. Sandini, et al. (2003). "From behavior-based to motivation-based robots." *Robotics and Autonomous Systems*.
Manzotti, R. and V. Tagliasco (2001). *Coscienza e Realtà. Una teoria della coscienza per costruttori e studiosi di menti e cervelli*. Bologna, Il Mulino.
Manzotti, R. and V. Tagliasco (2002). "Si può parlare di coscienza artificiale?" *Sistemi Intelligenti* XIV(1): 89-108.
Martinoli, A., O. Holland, et al. (2000). Internal representations and Artificial Conscious Architectures, California Institute of Technology.
Metzinger, T. (2003). *Being no one: the self-model theory of subjectivity*. Cambridge, Mass., MIT Press.
Millikan, R. G. (1984). *Language, Thought, and other Biological Categories: New Foundations for Realism*. Cambridge (Mass), MIT Press.
Moutoussis, K. and S. Zeki (1997). "Functional segregation and temporal hierarchy on the visual perceptive systems." *Proceedings of the Royal Society of London* 264: 1407-1414.
Nelson, C. A. and M. Luciana, Eds. (2001). *Handbook of Developmental Cognitive Science*. Cambridge (Mass), MIT Press.
Newman, A. (1988). "The Causal Relation and its Terms." *Mind* xcvii(388): 529-550.
O'Brien, G. and J. Opie (1997). "Cognitive science and phenomenal consciousness." *Philosophical Psychology* 10: 269-86.
Pearl, J. (2000). *Causality. Models, Reasoning, and Inference*. Cambridge, Cambridge University Press.
Pepperell, R. and M. Punt (2000). *The Postdigital Membrane. Imagination, Technology and Desire*. Bristol, Intellect(TM).
Perruchet, P. and A. Vinter (2002). "The Self-Organizing Consciousness." *Behavioral and Brain Sciences*.
Pfeifer, R. (1999/2001). *Understanding Intelligence*. Cambridge (Mass), MIT Press.

Rees, G., G. Kreiman, et al. (2002). "Neural Correlates of Consciousness in Humans." *Nature Reviews* 3: 261-270.
Schlagel, R. H. (1999). "Why not Artificial Consciousness or Thought?" *Minds and Machines* 9: 3-28.
Searle, J. R. (1980). "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 1: 417-424.
Searle, J. R. (1983). *Intentionality, an essay in the philosophy of mind*. Cambridge (Mass), Cambridge University Press.
Searle, J. R. (1984). *Minds, brains, and science*. Cambridge (Mass), Harvard University Press.
Searle, J. R. (1992). *The rediscovery of the mind*. Cambridge (Mass), MIT Press.
Steels, L. (1995). Is artificial consciousness possible? *Consciousness: Distinction and Reflection*. G. Trautteur. Napoli, Bibliopolis.
Steward, H. (1996). *The Ontology of Mind. Events, processes, and states.* Oxford, Clarendon Press.
Sutton, R. S. and A. G. Barto (1998). *Reinforcement Learning*. Cambridge (Mass), MIT Press.
Thau, M. (2002). *Consciouness and Cognition*. Oxford, Oxford University Press.
Togawa, T. and K. Otsuka (2000). "A model for Cortical Neural Network Structure." *Biocybernetics and Biomedical Engineering* 20(3): 5-20.
Whitehead, A. N. (1925). *Science and the modern world*. New York, Free Press.
Whitehead, A. N. (1927/1978). *Process and Reality*. London, Free Press.
Whitehead, A. N. (1933). *Adventures of ideas*. New York, Free Press.
Wittgenstein, L. (1953/1995). *Ricerche filosofiche*. Torino, Einaudi.
Zeki, S. (2001). "Artistic Creativity and the Brain." *Science* 293: 51-52.
Zeki, S. (2001). "Localization and Globalization in Conscious Vision." *Annual Review of Neuroscience* 24: 57-86.
Zeki, S. (2003). "The Disunity of Consciousness." *Trends in Cognitive Sciences* 7(5): 214-218.
Zeman, A. (2001). "Consciousness." *Brain and Mind* 124(7): 1263-89.
Zlatev, J. (2001). "The Epigenesis of Meaning in Human Beings, and Possibly in Robots." *Minds and Machines* 11: 155-195.