## To My Lifetime Teachers

*A. Stephan*
*M. Amini*
*A. Shafiee*

# The Conscious Mind Revisited

## An Informational Approach to the Hard Problem of Consciousness

Seyedreza Maleeh

University of Osnabrück

A thesis submitted for the degree of Doctor of Philosophy
at the University of Osnabrück

**September 2008**

# Contents

# Chapter 1
# Introduction

## 1.1 The Sketch of the Present Work

According to several compelling arguments, such as the logical possibility of zombies, the knowledge argument, and the like, which a so-called naturalistic dualist may pose, materialism is false.[1] The doctrine of materialism (or physicalism) holds that the only thing that can be truly proven to exist is matter. Its claim is that fundamentally, all things are composed of material and all phenomena (including consciousness) are the result of material interactions; therefore, matter is the only substance. So, according to materialism, physical facts exhaust all the facts about the universe. By naturalistic dualism I mean a branch of property dualism, which declares that phenomenal consciousness is ontologically independent of physical properties, but arises from a physical substrate in virtue of certain contingent laws of nature. In other words, the thesis claims that consciousness is a fundamentally new feature of the world as it does not logically supervene on physical properties.[2] Such an idea, which considers consciousness as a feature ontologically over and above physical features of the world, requires a construction of new fundamental psychophysical laws specifying the dependence of phenomenal to physical properties. In the most speculative part of his book, *The Conscious Mind* (1996), David Chalmers proposes his Double-Aspect Theory of Information as such a law, which brings out a crucial link between the physical and the phenomenal. According to double-aspect principle, "whenever we find an information space realized phenomenally, we find the same information space realized physically. And when an experience realizes an information state, the same information state is realized in experience's physical substrate."(Chalmers, 1996, p. 284)

As far as the notions of information are concerned, Chalmers' view has a lot in common with what is discussed by Shannon (1949)[3]. He formalizes Shannon's idea on information by appealing to the concept of an Information Space:

---

[1] The arguments against materialism will be briefly posed in the next section.

[2] For more detail on the notion of supervenience and its varieties, see Chalmers (1996)

[3] Shannon's theory mainly focuses on communications, control systems, and computers defining mathematical measure of the amount of information contained in a given message. It also introduces a quantitative method for expressing and analyzing the degradation of information during transmission, processing and storage. The theory is linked to statistical processes and works with quantitative expressions of the uncertainty about the outcome of possible alternatives and the amount of the information expected to be obtained, once one such alternative has actually occurred.

> "An information space is an abstract space consisting of a number of states, which I will call information states, and a basic structure of difference relations between those states." (Chalmers, 1996, p. 278)

As we will see in chapter 6, such a view suffers from the lack of a strong explanatory and predictive power and leads to counterintuitive notions such as panpsychism, the view that all parts of matter involve mind.[4] Among the probable ways, which may make the basic idea of double-aspect theory explanatorily stronger, perhaps each one from a different respect, this essay focuses on the notion of the information that Chalmers employs, namely, Shannon syntactic information. The proposal of this work is that exploiting such a concept makes the theory speculative rather than realistic, while confronting with many questions left open. So, if the substitution of a new notion of information can explanatorily give Chalmers' thesis more power of connecting the relevant observables in the relevant experiments and if such a substitution can contribute more predictive power to the theory, then we have enriched the idea in the right way.[5] There may also be other suggestions to enrich the thesis from different aspects dealing with which is out of the scope of this essay.

There are some alternative notions of information, which may fit double-aspect principle much better. In chapter 2, in addition to discussing Shannon's theory as not being a genuine theory of information, two other alternatives, namely, Dretske's (1981) interpretation of semantic notion of information and Roederer's (2003 & 2005) interpretation of pragmatic information are examined.[6] Semantic concepts are concerned with both signs and the objects to which they apply. The core idea in chapter 2 is that neither syntactic nor semantic concepts of information can give a universal and

[4] There is also a more holistic view according to which the whole universe is an organism that possesses a mind.

[5] As far as we are talking about observables and confirmation of predictions of a theory by experiments, we are indeed talking about the physical part of the theory or, more precisely, the physical correlates of consciousness. As phenomenal consciousness is not reducible to its physical correlates, it is not possible to measure it or deal with it experimentally.

[6] The way concepts of information are categorized in chapter 2 is due to the well-known triad syntactic, semantic, and pragmatic concepts of information. This way of categorization of the notion of information encompasses Shannon's theory as a syntactic concept of information in which the interrelation of the signs is studied, regardless of what they refer to, that is, meaning. I should also hint the reader that although Shannon's mathematical approach to the notion of information was quite new and innovative, neither Dretske nor Roederer was the founder of semantic and pragmatic notions of information, respectively. Each notion has a history. A selective, brief history on the latter notions will be posed in section 1.3.

comprehensive definition of information, potentially applicable to all sciences. It is claimed that in this respect, pragmatic concept of information would be a good candidate. Generally, pragmatic notion of information comprises relationships between signs and their impacts on those who use them. The proposal of this essay then would be that, in general, syntactic and semantic notions of information are included in the notion of pragmatic information and the two former notions are special cases of the latter.

In Roederer's reading of pragmatic information, the notion of "interaction" is projected as a basic, primordial concept; two fundamentally different classes of interactions are introduced, namely, physical interactions and information-driven ones. In physical interactions, there is direct energy transfer between two interacting bodies; complete knowledge of the initial conditions in such a system leads to the knowledge of its final state at any later time. Such systems are bidirectional in principle.

In information-driven interactions, however, there is no direct energy transfer between the two interacting bodies. In other words, such bodies are decoupled energywise. Complete knowledge of the initial conditions will not help in knowing the final state of the two bodies interacting informationally. In such interactions, the concept of "purpose" comes into play as a key discriminator between information-driven interactions and physical interactions. The claim then would be that *original* information-driven interactions only happen in the realm of *natural biological* systems. There can also be *derived* information-driven interactions, which happen in artifacts. The idea is that artificial systems (or parts of an artifact) reflect the information *we* put in them by determining the initial conditions or providing the setup of measurement. Derived information-driven interactions only happen artificially. They do not happen in nature or just happen by chance.

Information-driven interactions happen between systems, or parts of a system, the complexity of which exceed a certain degree. In chapter 3, the notion of complexity together with some related notions such as "self-organization" and "emergence" are discussed. The claim is that complexity and self-organization alone do not represent information. Information is only represented in self-organizing systems, which enjoy information-driven interactions between their constituent parts. The latter only happens in the realm of biology. Regarding the notion of "emergence," the essay adopts the position according to which every phenomenon is weakly emergent but consciousness. As the notion of "initial condition" also plays a crucial role in discerning information-driven interactions from physical ones, it is also dealt with in detail in chapter 3.

In the next section, we will see that consciousness and intentionality go hand in hand: any successful theory of consciousness should account for intentionality successfully. In chapter 4, the notion of "intentionality" has been given a new natural account in virtue of pragmatic information. According to this account, every system with information-driven interactions between its parts is an intentional system. Whether the system is a mental system or not, whether the system is a biological system or an artifact, such intentionality is intrinsic to the system. The combination of the new information-based account of intentionality proposed in this essay with double-aspect principle, dealt with in chapter 6, explains how psychological aspects of consciousness link up the phenomenal ones.[7]

Chapter 5 deals with David Bohm's (1990) notion of "active information" and the way such a notion relates mind to matter. It is shown that his account of active information in his hidden variable theory is very similar to the account of pragmatic information. However, in the interaction between wave and particle at quantum level, according to Bohm, we witness nonlocality. Nonlocality is direct influence of one object on another independently of geometrical distance. Wave and particle also interact non-mechanically, meaning that their interaction is not reducible to any of the four fundamental interactions known in physics. Contrary to the mechanical interactions in which there is energy and matter transfer between two interacting physical objects, in the non-mechanical wave-particle interaction, according to Bohm, there is no matter and energy transfer between these two interacting components comprising the whole system. These two properties of the wave-particle interactions, that is, being nonlocal and non-mechanical, are prohibited in information-driven interactions occurring in the classical world.[8]

---

[7] As we will see in the next section, there are two aspects of consciousness, which exhaust the conscious mind: psychological aspect and phenomenal one. The former can be analyzed in a way wherein we have access to some information to use it in the control of behavior, while in the latter the conscious mind is characterized by the way it *feels*.

[8] The reader should bear in mind that the way the notion of "mechanical" or "mechanics" is used within Bohmian paradigm in this essay is different from what is meant by the terms in Newtonian paradigm (discussed in chapter 3), which deals with the macroscopic, non-relativistic, and non-quantum domain and which is a branch of physics. "Mechanical" and "mechanics" in Bohmian paradigm, as interpreted in this

If Bohm's description of the wave-particle interaction corresponds to reality (as we can never be sure of this; individual quantum domain is inaccessible), then such a nonlocal, non-mechanical interaction can be good evidence for the existence of a sort of interaction in which information is processed phenomenally. For such a nonlocal, non-mechanical interaction is not reducible to mechanical one. The latter may give rise to conscious experience.

Not only is pragmatic information the most comprehensive notion of information, compared to the syntactic and semantic notions, but also the fittest to Chalmers' double-aspect principle according to which when information is processed physically, at least in some cases, it is also processed phenomenally. The notion of pragmatic information can successfully explain many mental processes and states while Chalmers' notion of information is unable to do that. This is the claim proposed in chapter 6.

As stated earlier, according to Chalmers, the appropriate notion of information to adopt for the double-aspect principle is Shannon syntactic information proposed in virtue of Bateson's slogan "the difference that makes a difference:" Information is physically realized in the system consisting of a light wired to a switch, for example, when the change we make in the position of the switch *causes* the light to go on or off. In chapter 6, I criticize such a notion, stating that "the difference that makes a difference" principle adds nothing to our knowledge of the nature of information. The principle does not go further than the causal explanation of any phenomena. If information is nothing but causality without any more properties, why do we need the notion of information at all?

So, the suggestion would be to replace "the difference that makes a difference" principle with pragmatic information. This will provide the double-aspect principle with much more explanatory and predictive power while skipping the counterintuitive issues such as panpsychism. On the other hand, it modifies Roederer's interpretation of pragmatic

essay, have been used in a broader sense of the terms; they are used equivalently to the use of the terms "physical" and "physics" respectively in their conventional interpretation. In such an interpretation, physics is the science of matter and motion, as well as space and time. Physics uses the concepts of energy, force, mass, and charge in explaining phenomena in nature. Every interaction in physics is reducible to one of the four "weak," "strong," "gravitational," and "electromagnetic" interactions. I will deal with such interactions in more detail in the next chapter.

information in the following way: despite the explanatory power of Roederer's account of pragmatic information, his account of consciousness fails to capture the subjective character of it. The failure stems from his physicalist point of view. He considers consciousness as arisen from the coherent interplay between cortical and limbic functions. The claim of the essay is that such an account of consciousness at best can be regarded as a candidate for being a neural correlate of consciousness (NCC).[9]

So, the proposal in this essay is that accommodating pragmatic information in double-aspect principle modifies both theses: on the one hand, it contributes to Roederer's reading of pragmatic information, as an originally physicalist thesis, another aspect, *viz.*, a phenomenal one. On the other hand, it substitutes the explanatorily powerful notion of pragmatic information for the syntactic information, originally used in the double-aspect principle by Chalmers.

There are also some crucial questions left open in my proposal, for example: what are the mechanisms or criteria according to which phenomenal aspect of information is realized? What are the criteria according to which phenomenally realized information corresponds to *my* experiences? Regarding the idea that experience arises from its physical base, how is the correspondence between experience and its physical base established?

The above questions are decisive ones showing that the thesis I put forward is still in its infancy. Furthermore, consciousness is an inaccessible domain, like individual quantum domain. In consciousness studies, there are no strict criteria to distinguish a thesis as the final one. We can only choose theses with more explanatory and predictive power as better ones. What is proposed in this essay is an attempt to locate the thesis among the latter theses.

[9] The definition of a NCC given in the program of the ASSC (Association for the Scientific Study of Consciousness) concerns it as what follows: a neural correlate of consciousness is a specific system in the brain whose activity correlates directly with states of conscious experience.

## 1.2 Consciousness: an Introduction

When we, as conscious beings, perceive and think, there is a chain of information-driven interactions leading to the physical activation of the relevant parts of the brain and body. But there is also an internal aspect associated with such interactions, which leads to our direct, intimate, and subjective knowledge of our perception and thought. In other words, we have immediate knowledge of how it feels when such processes happen; there is something it feels like to be a cognitive agent. To borrow Nagel's (1974) famous phrase, we can say that a creature or a mental state is conscious if there is something it is like to be that creature or in that mental state. "Equivalently, we can say that a mental state is conscious if it has a *qualitative feel* – an associated quality of experience. These qualitative feels are also known as phenomenal qualities, or *qualia* for short. The problem of explaining these phenomenal qualities is just the problem of explaining consciousness. This is the really hard part of the mind–body problem." (Chalmers, 1996, p. 4)

Furthermore, although we have direct knowledge of our own experiences, trying to know facts about others' consciousness will simply fail. Such an asymmetry, epistemic asymmetry, brings an important distinction between two systematic approaches about, namely, first-person perspective versus third-person perspective: "From the inside, consciousness seems all-pervasive, self-evident, and undeniable […] from the outside, firsthand exploration of the consciousness of others just seems to be out of the reach of ordinary scientific methods, others' experiences being neither directly nor non-inferentially verifiable." (Güzeldere, 1997, p. 25) The root of such an asymmetry lies at the subjective quality of consciousness.

The problem stated above can be seen from a different angle, i.e., the asymmetry of our knowledge of objective and subjective facts: "If all we knew about were the facts of physics, and even the facts about dynamics and information processing in complex systems, there would be no compelling reason to postulate the existence of conscious experience." (Chalmers, 1996, p. 5)  Such an asymmetry has been labeled as the "hard problem" of consciousness. The hard problem of consciousness is thus the problem of experience. Chalmers also gives an account for the easy problems: "The easy problems of consciousness are those that seem directly susceptible to the standard methods of cognitive science, whereby a phenomenon is explained in terms of computational or neural mechanisms." (Chalmers, 1995, p. 201- 202) To have given clearer accounts on the easy problems versus the hard problem of consciousness, first we need to explore the concepts of mind together with the concepts of consciousness, examining how these concepts relate to one another and how they play role in introducing problems associated with consciousness.

### 1.2.1 Two Concepts of Mind

There are two quite distinct concepts of mind: a phenomenal concept (or conscious experience) on which mind is characterized by the way it *feels* and a psychological[10] concept on which mind is characterized by what it *does*. These concepts are "aspects" of the mind and there should be no question of competition between these two notions of mind:

> "Neither of them is the correct analysis of mind. They cover different phenomena, both of which are quite real […]. A specific mental concept can usually be analyzed as a phenomenal concept, a psychological concept, or as a combination of the two." (Chalmers, 1996, p. 11-12)

Indeed, these two concepts exhaust the mind.

Having the two concepts in mind, it seems noteworthy to examine some mental terms to see whether they are phenomenal, psychological or the hybrid of the two concepts. Let us start with "learning." There seems no qualitative feel accompanying learning.

> "For something to learn, at a first approximation, is for it to adapt its behavioral capacities appropriately in response to certain kinds of environmental stimulation." (Chalmers, 1996, p. 11)

Thus, learning must lie within the psychological concept of mind.

[10] By the term "psychological" I mean the same view that cognitive science adopts, that is, those aspects of mind responsible for the causation of behavior, regardless of being conscious or not.

What about strong emotional reactions? Certainly, there is an immediate felt aspect and an internal phenomenal character accompanying psychological aspects of emotions. Doubtless, there is a strong intrinsic and qualitative feel associated with, for example, thinking of my willingness to hug my little daughter after being a year far from home. So, I undergo some hybrid states including both phenomenal and psychological aspects of mind.[11]

In such discussions, beliefs, as well as desires, hopes, thoughts, and the like appear to be more interesting and complicated because of their *intentional* character. Intentionality has to do with directedness, about-ness or of-ness of mental states. Our thoughts, hopes, beliefs, desires, and the like are intentional in a sense that they are always directed on, or at something. I have discussed the notion of intentionality in length in chapter 4.

As far as the relationship between consciousness and intentionality is concerned, three philosophical groups are discernable:[12]

Group 1: Philosophers belonging to this group hold that these two mental aspects of the mind can be studied separately. Among the advocates of such an idea, Wilfrid Sellars (1956), Gilbert Ryle (1949), Wittgenstein (1953), Putnam (1975), Fodor (1991), and Donald Davidson (1983, 1986) can be mentioned. As evidence, let us have a look at Fodor's justification on the distinction between phenomenal consciousness and intentionality:

> "It used to be universally taken for granted that the problem about consciousness and the problem about intentionality are intrinsically linked: that thought is ipso facto conscious, and that consciousness is ipso facto consciousness of some or other intentional object […]. Freud changed all that. He made it seem plausible that explaining behavior

[11] Chalmers (1996) suggests a sort of test to see whether a mental notion M is primarily psychological or not. To see whether a mental notion M is psychological or not we should ask: "could something be an instance of M without any particular associated phenomenal quality? If so, then M is likely psychological. If not, then M is phenomenal, or at least a combined notion that centrally involves phenomenology." (Chalmers, 1996, p. 18)

[12] Others may categorize such a relation in different ways. For example, Charles Siewert (2003) distinguishes four philosophical positions that are noteworthy to be mentioned here: "a: consciousness is explanatorily derived from intentionality. b: Consciousness is underived and separable from intentionality. c: consciousness is derived but inseparable from intentionality. d: Consciousness is underived from, inseparable from, and essential to intentionality." (Siewert, 2003, p. 22)

> might require the postulation of intentional but unconscious states. Over the last century, and most especially in Chomskian linguistics and in cognitive psychology, Freud's idea appears to have been amply vindicated […] dividing and conquering − concentrating on intentionality and ignoring consciousness − has proved a remarkably successful research strategy so far." (Fodor, 1991, p. 12)

Group 2: This group contains those who try to analyze consciousness totally in terms of intentionality. Philosophers among this group are divided into two distinct subgroups: the first subgroup analyzes conscious states in terms of higher-order states that represent them. The main proponents of such a theory are David Rosenthal and Peter Carruthers. Philosophers belonging to the second subgroup, such as Fred Dretske and Michael Tye, among others, have adopted another strategy. They focus on the first-order intentional content of conscious states, and advocate representationalism. I will come back to the ideas of both subgroups while dealing with the notion of introspection.

Group 3: This group grounds intentionality in consciousness. Here, John Searle can particularly be mentioned, among others, as the main advocate of such a thesis. Stating his "connection principle," Searle holds that:

> "Only a being that could have conscious intentional states could have intentional states at all, and every unconscious intentional state is at least potentially conscious […]. There is a conceptual connection between consciousness and intentionality that had the consequence that a complete theory of intentionality requires an account of consciousness." (Searle, 1992, p. 132)

As will be apparent in chapter 4 and section 6.3, the idea of this essay has a lot in common with Searle's thesis on intentionality though in many respects I criticize Searle's account of intentionality. The idea I put forward in this dissertation on the relation between intentionality and consciousness also benefits from the claims of group 2 generally. In short, the thesis of this proposal is that there is a sort of bidirectional connection between consciousness and intentionality: natural biological systems are originally intentional; such systems are potentially conscious. On the other hand, consciousness *only* occurs in the originally intentional systems. Note that intentionality in the present work has been given a new natural account that will be depicted in chapter 4.

Let us turn our attention again to the notion of belief, as the notion has a central role in philosophy of mind. Consider group 2 and consider, for instance, David Rosenthal as the representative of the first subgroup and Michael Tye as the representative of the second one. According to Rosenthal, "we are conscious of something when we have a thought about it. So a mental state will be conscious if it is accompanied by a thought about that state. The occurrence of such a higher-order thought (HOT) makes us conscious of the mental state." (Rosenthal, 1997, p. 741) Note that in Rosenthal's view, "'thought' has a generic sense, which applies to any propositional mental state, regardless of its mental attitude" (Rosenthal, 1997, p. 742) such as believing, hoping, expecting, and the like. In such a view, beliefs can be construed as higher-order thoughts that are functionally and causally connected to the lower-order mental states.

> "But it does square nicely with materialist views. For the account holds that what makes conscious mental states conscious is their causing higher-order thoughts that one is in those mental states. And the materialist can reasonably maintain that this causal pattern is due to suitable neural connections […]. The present explanation, moreover, has precise empirical consequences that one could reasonably hope to test. For it implies not only that conscious mental states are accomplished by distinct higher-order thoughts, but also that some causal mechanism exists that connects mental states to corresponding higher-order thoughts." (Rosenthal, 1991, p. 467)

With the above statements in mind, we can successfully place Rosenthal's view on beliefs within the psychological concept of mind.

Tye also does not consider any phenomenal aspect for beliefs. It seems to him "not implausible to deal with these cases [,among them beliefs,] that insofar as there is any phenomenal or immediately experienced felt quality to the above states, this is due to their being accompanied by sensations or images or feelings that are the real bearers of the phenomenal character. Take away the feelings and experiences that happen to be associated with the above states in particular cases, and there is no phenomenal consciousness left." (Tye, 1995, p. 14)

The above claim shows that Tye excludes beliefs from phenomenal aspect of mind, but according to what follows, includes them in the psychological properties of mind:

> "Talk of belief compartment is not intended to presuppose that there is a special, spatially discrete region of the brain dedicated to beliefs, any more than talk of the memory space of a computer demands that there be some single, contiguous physical region within the machine in which information is stored. Rather, the idea is that there is a *functionally* discrete region. Sentences that are tokened in this functional belief space are alike in that they play a common functional role with respect to other mental sentences, sensory inputs, and behavioral outputs. Beliefs, on this view, are states that function in a distinctive way with respect to perceptual information, desires, and behavior. Any sentence token that occupies a role of the sort that is distinctive to beliefs belongs in belief compartment." (Tye, 1995, p. 98)

Concerning the above statements, there remains no doubt that in his view, beliefs are psychological properties in a sense mentioned above, that is, in the sense that has to do with functional and causal roles in the control of behavior.

However, the way philosophers who belong to the group 3 deal with the case is quite different. According to John Searle, as the prominent proponent of the idea of grounding intentionality in consciousness, we should distinguish between what he calls "as-if" intentionality or "interpreter-relative" intentionality on the one hand, and "intrinsic" intentionality on the other. In short, he holds that without consciousness, all that is present is "as-if" intentionality. Intrinsic intentional systems are at least potentially conscious.[13] Thus, experience is essential for anything to count as belief.

As mentioned earlier, Searle's thesis on intentionality will be discussed and criticized in chapter 4. For the time being and for the purpose of this section, suffice it to say that as

[13] In Searle's vocabulary, "intrinsic intentionality" and "original intentionality" are synonyms. In the terminology adopted in this work, however, I will distinguish between the two expressions. See chapter 4 for more detail.

far as the question of whether beliefs should be considered the psychological aspect of mind or phenomenal one, the above discussion can be summarized through the following quotation:

> "We can note that there is at least a *deflationary* concept of belief that is purely psychological, not involving conscious experience; if a being is in the right psychological state, then it is in a state that resembles belief in many important ways, except with respect to any phenomenal aspects. And there is an *inflationary* concept of belief, on which conscious experience is required for truly believing, and perhaps even on which a specific sort of conscious experience is required for truly believing that P."(Chalmers, 1996, p. 18)

What counts here is that there seems nothing in the interpretation of belief that stands over and above the psychological and phenomenal concepts of mind. Other intentional states such as thoughts, hopes, desires etc. could also be treated in the same manner.

### 1.2.2 Two Concepts of Consciousness

Concerning the above concepts (or aspects) of mind, it seems quite natural to consider consciousness also having two quite distinct − but often related − concepts, that is, phenomenal and psychological concepts of consciousness. In what follows, I will discuss these two notions to highlight the distinction.[14] To prevent confusions and to have a clearer look at some philosophers' ideas, I distinguish between "creature" and "state" consciousness in each of the two notions of consciousness.

- *Phenomenal Consciousness*

Perhaps the most mysterious and perplexing concept of mind is phenomenal consciousness. In November 1992, magazine Discover devoted an issue to the ten great unanswered questions of science among which one could see the questions "What is

[14] There are also some other terms and expressions that show such a duality in philosophy of mind; some of them are as what follows: "access" versus "phenomenal" consciousness (Block, 1995), "O-story" versus "I-story" (MacKay, 1980), the "easy" versus "hard" problems (Chalmers, 1996), and metaphysical versus epistemic explanations (Levine, 1983).

Consciousness?" together with "Does Chaos rule the Cosmos?" and "How big is the Universe?." Note that by the word consciousness in the above questions, it is aimed at phenomenal consciousness, not any other aspects of the term. This concept of consciousness subsumes phenomenal aspects of mind. Being conscious phenomenally is just an instantiation of an intrinsic, subjective, and qualitative feel. It is the existence of phenomenal consciousness that leads to the hard problem of consciousness mentioned above, and this is this concept of consciousness at which this essay aims in order to find an explanatorily powerful psychophysical law, a law that can successfully bind the phenomenal aspect of consciousness to the other one, that is, the psychological aspect.

The idea in this essay is that not only is there no conflict between the phenomenal and psychological concepts of consciousness, but they also exhaust the conscious mind. This is what Güzeldere calls integrationist intuition:

> "In contrast with the segregationist intuition is what I call the *integrationist intuition*: what consciousness does, *qua consciousness*, cannot be characterized in the absence of how consciousness seems, but more importantly, that how consciousness seems cannot be conceptualized in the absence of what consciousness does." (Güzeldere, 1997, p. 11)

As stated earlier, I will divide phenomenal (and also psychological) aspects of consciousness into the creature and state consciousness.

Creature Consciousness. According to Thomas Nagel's (1974) famous "what it is like" criterion, a being is conscious just if there is "something that it is like" to be that creature. In other words, to say that a being is conscious (in the phenomenal sense) is to say that the world seems or appears from that creature's mental or experiential point of view − in a subjective way.  Thus, bats are conscious because world seems to them through their echo-locatory sense in a subjective way that we, as humans, are not able to understand − from our point of view − that what such a sense of consciousness is like from the bats' point of view.

State Consciousness. A state is phenomenally conscious if it has experiential properties. The totality of the experiential properties of a state are “what it is like” to have it. So, a mental state is conscious if there is something it is like to be in that mental state. Another way of putting this is to say that a mental state is conscious if it has a qualitative feel. Qualitative feels are also known as phenomenal qualities or qualia for short.[15] Switching from synonyms to examples, we have phenomenal conscious states when we have visual, auditory, tactile, olfactory, and taste experiences; when we have the sense of hot and cold, pains, or bodily sensations. Moving to some experiences associated with objects generated internally, we can instance mental imagery, conscious thought, emotions, desires, and the like. There can be some other rich experiences that are derived from the combination of two or more of these experiences. Such experiences are very useful for the experience of the unity of consciousness, that is, the way some experiences are tied together in a way that seem as a single experience to the experiencer.

- *Psychological Consciousness*

This concept of consciousness can be analyzed in a way wherein we have access to some information to use in the control of behavior. Chalmers (1996) calls the psychological aspect of consciousness as “awareness” in the broader sense of the term. According to him, phenomenal consciousness is always associated with psychological consciousness, but the vice versa does not necessarily hold, unless we take a narrower sense of the notion of awareness explained in chapter 6. There are several varieties of the psychologically conscious mind, briefly explained in the following.

Creature Consciousness

Sentience. In this sense, a creature is conscious if capable of sensing and responding to its world (Armstrong, 1981). A sentiently conscious creature is a stimulus-response entity with a certain level of competence and access to or capacity to retain some certain

[15] Many philosophers do not consider these two terms, i.e., phenomenal and qualitative states, to be equal; see, for example, Van Gulick (1993, 2004). Also, existence and nature of qualia has been very controversial; see Block (1980), Jackson (1982), Levine (1983), Dennett (1990), Churchland (1985), Shoemaker (1990), Chalmers (1996), and Clark (1993). I am not to deal with different uses of “qualia” in this entry. In this essay, I use the term in what I think is the standard way to refer to those properties of mental states that type those states by what it is like to have them.

amount of information about its world. It is also capable of placing itself in situations wherein information is constantly being acquired. Such a definition seems very ambiguous because it is not clearly defined to what extent sensory capacities are sufficient for being conscious in this sense. Can we consider a butterfly conscious in this sense?

Awakeness. One might say that a creature is conscious just if it were awake and normally alert. This notion falls in psychological consciousness since it makes sense to think about a creature having experience while sleeping. Thus, what here counts is the analysis of awakeness in functional terms as an ability to process information about the world and treat it in a rational fashion.

Self-Consciousness. A creature can be considered self-conscious if that creature has the capacity for second-order representation of the creature's own self. Self-consciousness can also be analyzed in terms of the access to a self-model. As Allen (2005) points out, because of its second-order character ("thought about thought") the capacity for self-consciousness is closely related to questions about "theory of mind" in nonhuman animals and [even in young children] — whether they are capable of attributing mental states to others. Questions about self-consciousness and theory of mind in animals are a matter of active scientific controversy.

Transitive Consciousness. As Rosenthal (1997) states, we can use the word "conscious" in two distinct ways: transitive and intransitive consciousness. The former is when we speak of our being conscious of something. Here, there is a direct object that we are conscious of, and because of this direct object, Rosenthal calls it transitive consciousness. Thus, when a creature is transitively conscious of an object, that being is a conscious creature. Being transitively conscious of something is a relation that a person or other creature bears to that thing. Note that the direct object can be physical or mental things alike. We may be conscious of a tree or a nice song, or of a mental state. It must be specified that our transitive consciousness of our mental state relies on neither inference nor observation, because if we are to be conscious of a state with benefit of ordinary inference, that inference must be a conscious inference and it makes the issue circular.

State Consciousness

Intransitive Consciousness. As stated above, as a conscious creature, we can be conscious of a mental state transitively. This makes that state a conscious mental state. Thus, according to Rosenthal's reading (1997), a conscious mental state is compound of two things: a mental state, which is not conscious *per se* and one's being transitively conscious of it. Rosenthal emphasizes that being conscious of a mental state does not cause that state to have the property of being intransitively conscious; rather, a mental state's being intransitively conscious simply consists in one's being transitively conscious of it. In short, the core of Rosenthal's idea is that we are conscious of something when we have a thought about it. So a mental state is conscious if it is accompanied by a thought about that state. Recall that the thought must be arrived at non-inferentially and non-observationally.

Introspection. Mainly, there are two distinct kinds of proposals conceiving introspection: introspection either as a form of inner sense (the higher-order perceptual model − HOP − or simply, the perceptual model), or as a form of higher-order thought (HOT, or the conceptual model). The pioneer of both ideas in the philosophical debate is David Armstrong (1968, 1997). In what follows, I will discuss these two concepts in relatively more detail.

HOP. Following Locke, who put forward the theory of consciousness as "internal sense" or "reflection," and Kant that made it "inner sense," Armstrong  states the his doctrine as follows:

> "Introspective consciousness […] is perception-like awareness of current states and activities in our own mind. The current activities will include sense-perception: which latter is the awareness of current states and activities of our environment and our body."(Armstrong, 1997, p. 724)

According to Armstrong, introspection is an ability to gain information that is about certain of behavior-producing or potentially behavior-producing states:

> "[…] a mental state is a state of the person apt for bringing about certain of bodily behavior. So when I acquire by introspection the information that, for example, I am sad now or that I have a certain sort of perception now, this information is information about certain of my behavior-producing or potentially behavior-producing states. Now if introspection is conceived of as 'acquaintance' with mental states, or a searchlight that makes contacts with them, it is difficult to see how all it can yield is information of such highly abstract nature about inner causes or potential inner causes. But if introspection as well as perception is conceived of a mere flow of information or beliefs, then there is no difficulty." (Armstrong, 1968, p. 326)

A slightly deflated version of Armstrong's view has energetically been pursued by William Lycan. He holds that:

> "Consciousness is functioning of internal attention mechanism directed at lower-order psychological states and events […]. Attention mechanisms are devices that have the job of relying and/or coordinating information about ongoing psychological events and processes." (Lycan, 1996, p. 14)

HOT: Two distinct views can be accommodated in this entry. The first one is Rosenthal's view. He holds that introspection involves more than a mental state's being a conscious state:

> "Introspecting a mental state is deliberately and attentively focusing on that state. Non-introspective consciousness, by contrast, requires no special act of attention […]. A mental state is non-introspectively conscious when accompanied by a relevant HOT; introspection occurs when there is a third-order thought that makes the second order thought conscious." (Rosenthal, 1997, p. 745)

The proponent of the second idea is Dretske (1995).[16] According to him, introspective knowledge is mind's direct knowledge of itself and introspection is the process by means of which we come by such knowledge. Dretske holds that in introspection, mind turns its attention onto itself. In such a model, there is no need for internal scanners monitoring.

[16] See also Tye (1995, 2000)

> "On a representational theory of the mind, introspection becomes an instance of displaced perception – knowledge of internal (mental) facts via an awareness of external (physical) objects […]. One comes to know that *k* is F − sees that *k* is F − by seeing and hearing, not *k* itself, but *h*. The perceptual fact is displaced from the perceptual object." (Dretske, 1995, p. 41)

So, according Dretske, when I see, for example, the gas tank is nearly empty by seeing the gas gauge, I do not see the gas tank. My seeing-that is secondary or displaced.[17]

Narrative Consciousness. According to Dennett's (1991) Multiple Drafts Model (MDM), all varieties of perception are accomplished in the brain by parallel, multitrack processes of interpretation and elaboration of sensory inputs. In the nervous system, information is continuously under "editorial revision," during which various additions, incorporations, emendations, and overwritings of content can occur, in various orders. All we actually experience is a product of these processes taking place in the streams of activity in various parts of the brain. Once the observation of some feature has been made, the information content is fixed in a specialized, localized portion of the brain:

> "These distributed content-discriminations yield, over the course of time, something *rather like* a narrative stream or sequence, which can be thought of as subject to continual editing by many processes distributed around in the brain, and continuing indefinitely into the future." (Dennett, 1991, p. 113)

So, according to Dennett, a person's conscious mental states are considered equal to those, which appear in the stream of consciousness.

Access Consciousness. The proponent of this notion is Ned Block (1995) according to whom "a state is access-conscious if, in virtue of one's having the state, a representation of its content is (1) inferentially promiscuous, i.e. poised to be used to be as premise in reasoning, and (2) poised for [rational] control of action, and (3) poised for rational control of speech." (Block, 1995, pp. 230-231)

[17] Representational systems, according to Dretske, are examined in chapter 4.

Block holds that these conditions are together sufficient, but not all necessary. He sees A-consciousness (access consciousness) as a cluster concept in which (3) − roughly reportability − is the element of the cluster.[18]

As stated earlier, care should be taken that the phenomenal concept of mind is the same as conscious experience or phenomenal consciousness, but the psychological concept of mind is not the same as psychological consciousness because a psychological mental state responsible for the causation of behavior can be conscious or not. For instance, consider reportability as a variety of psychological consciousness. You may have a mental state, responsible for some behavior, that you are not able to report its content. So, it is not psychologically conscious. The moral, thus, is that phenomenal consciousness exhausts the phenomenal concept of mind but psychological consciousness does not exhaust the psychological concept of mind.

To sum up, it should be said that psychological and phenomenal consciousness exhaust the mental term "consciousness;" there are in general only two senses, kinds, concepts, or aspects of the term consciousness: phenomenal and psychological. The crucial question then will be: can cognitive models of consciousness successfully explain both aspects of it? It seems that as far as we are confronted with the processes that can reductively be explained, cognitive models are good candidates for giving appropriate accounts for consciousness, at least in principle. In other words, cognitive models can suitably give *functional* accounts of consciousness. Thus, cognitive models of consciousness can properly explain the psychological concept of consciousness. But the hard problem of consciousness is that of phenomenal consciousness, that of how phenomenal aspect of consciousness arises when the correspondent psychological aspect has been realized, as, according to several persuasive arguments which I will briefly mention in what follows, phenomenal consciousness is not reducible to the physical facts. Functional and causal accounts of psychological aspects of consciousness cannot shed light on phenomenal aspects. Experience cannot be explained by physical terms, while there is no problem, in

[18] This is why I do not mention "reportability" and "voluntary control" among varieties of psychological consciousness. Indeed, these two can be accommodated within the access consciousness. I also skipped explaining "attention" as one of the varieties of this concept of consciousness because of its similarity with transitive and intransitive consciousness.

principle, with explaining psychological consciousness physically. This idea is in line with Chalmers' proposal on the irreducibility of the phenomenal aspect of consciousness to its psychological aspect:

> "Cognitive models are well-suited to explaining psychological aspects of consciousness. There is no vast metaphysical problem in the idea that a physical system should be able to introspect its internal states, or that it should be able to rationally deal with information from its environment, or that it should be able to focus its attention first in one place and then in the next. It is clear enough that an appropriate functional account should be able to explain these abilities, even if discovering the correct account takes decades or centuries. But the really difficult problem is that of phenomenal consciousness, and this is left untouched by the explanations of psychological consciousness that have been put forward so far." (Chalmers, 1996, pp. 28 & 29)

So, the aforementioned hard problem of consciousness can be formulated in a new way, namely, how phenomenal aspect of consciousness can be linked to the psychological one. And this is the task of this essay to link these two aspects of consciousness by projecting a modified version of double-aspect theory of information first put forward by Chalmers as stated earlier in this section.

As said above, there are several convincing arguments against the reducibility of phenomenal consciousness to the physical facts. As I will refer to some of these arguments within different parts of this essay, in what follows, I will shortly review the main arguments for the irreducibility of consciousness to the physical facts. In doing so, I have used Meixner's (2004) way of organizing the arguments.

### 1.2.3 Recent Arguments against Physicalism: a Short Review

#### Nagel's Bat

"*Premise 1*: Bats have experiences.
*Premise 2*: If bats have experiences, then there are some things regarding the experiences of bats that one could only know if one were a bat.

*Premise 3*: If the experience of bats are – all of them – physical events, then there is not a thing regarding the experiences of bats that one could only know if one were a bat.
*Conclusion*: The experiences of bats are not – not all of them – physical events." (Meixner, 2004, pp. 160-161)

### Jackson's Mary

"*Premise 1*: Mary is, and always has been, unable to have color experiences.
*Premise 2*: If Mary is, and always has been, unable to have color experience, then there are some things regarding color experiences that Mary cannot now be able to know.
*Premise 3*: If color experiences – all of them – are physical events, then there is not a thing regarding color experiences that Mary cannot now be able to know.
*Conclusion*: Color experiences are not – not all of them – physical events." (ibid, p. 166)

### Searle's Chinese Room

"*Premise 1*: It is possible in principle that a Chinese Room does not understand (any) Chinese, but acts perfectly as if it understood (some) Chinese.
*Premise 2*: If understanding Chinese is a physical property, then it is not possible, not even in principle, that a Chinese Room acts (i.e., *functions*) perfectly as if it understood Chinese and does not understand Chinese.
*Conclusion*: Understanding Chinese is not a physical property.
*Corollary*: That Mao Tse-tung understood Chinese in 1955 is not a physical state of affairs." (ibid, p. 172)

### Chalmers' Zombie

"*Premise 1*: There is a possible world in which there are no (conscious, actual) experiences, but which is physically exactly like the actual world, in which I have experiences.
*Premise 2*: If having experiences is a physical property, then there is no possible world, which is physically exactly like the actual world, but in which there are no experiences.
*Conclusion*: Having experiences is not a physical property, and that I have experiences is a fact that is not physical." (ibid, p. 180)

**The Inverted Spectrum and Modified Colors**

“The following argument is an advantageous modification of an argument well-known under the name “The Inverted Spectrum”:

*Premise 1*: My color experiences can be different from the color experiences I actually have, the physical world remaining exactly as it (actually) is.
*Premise 2*: If my color experiences are – all of them – physical events, then my color experiences cannot be different form the color experiences I actually have, the physical world remaining exactly as it (actually) is.
*Conclusion*: My color experiences are not – not all of them – physical events.

This is more complex original:

*Premise 1*: The relationship between the qualitative character of your subjective color spectrum and the qualitative character of my subjective color spectrum can be, due to switching of subjective colors on my part (for example), other than it actually is, the physical world remaining exactly as it is.
*Premise 2*: If your and my color experiences are physical events, then the relationship between the qualitative character of your subjective color spectrum and the qualitative character of mine cannot be other than it actually is, the physical world remaining exactly as it is.
*Conclusion*: Your and my color experiences are not physical events.” (ibid, p. 186)

If, according to the above arguments, phenomenal consciousness is not reducible to the physical facts, we should adopt another position. As stated in the last section, the position adopted in this essay is naturalistic dualism. Naturalistic dualism is a sort of “fundamental property dualism.” According to the latter, the existence of conscious mental properties, as basic constituents of reality, is not ontologically dependent upon, nor derivative from any other properties such as those of physics. Naturalistic dualism claims that although consciousness is ontologically independent of physical properties, it arises from a physical substrate in virtue of certain contingent laws of nature. As I have mentioned frequently earlier, the claim of this essay is that it is possible that the most fundamental contingent law of nature, which links physical properties to phenomenal ones, be double-aspect theory of information, the modified version of which is projected in this essay.

To be able to delve into the above theory, and more importantly, to adopt the best possible notion of information for double-aspect theory, it is necessary to get familiar in more detail with the concept of information, its varieties, and the specifications, strengths, and weaknesses of each variety. Such a task will be fulfilled in the next section as well as the next chapter.

## 1.3 Information: a Selective History

Consider the words “information,” “deformation,” “conformation,” “transformation,” and “reformation.” They obviously derive from “formation,” the term, which in turn comes from “form.” Information is thus putting form into some previously unformed entity. Other terms mentioned above with the addition of prefixes de-, con-, trans-, and re- to formation, refer to undoing, copying, changing, and renewing of forms, respectively. (von Baeyer, 2003, p. 20) But, what is *form*?

The English word “form” may be used to translate two distinct concepts with which Plato was concerned—the outward “form” or appearance of something (Greek *eidos*, εἶδος, and *idea*, ιδέα, in their conventional, nontechnical senses, or other terms such as *morphē*, μορφή), and “Form” in a new, technical sense, apparently invented by Plato (esp. *eidos*, *idea*). These are often distinguished by the use of uncapitalized “form” and capitalized “Form,” respectively.

According to Plato, the forms that we see are not real, but literally *mimic* the real Forms. In the “Allegory of the Cave” expressed in *Republic*, forms are called the shadows of artificial replicas of real things. That which the observer understands when he views the mimics are the archetypes of the many types and properties (that is, of universals) of things we see all around us. They are not located in the object, which as far as Plato is concerned, is mere smoke and mirrors situated in space (which also is real). In this account, a horse, for example, is but a copy of the Form of horse-ness, the horse of horses.

Aristotle was not satisfied with Plato’s account of Forms. What proof do we have, Aristotle asked, that these things called Forms enjoy a separate existence? However, he did not reject the notion all together, but defined it as the sum total of the essential properties of a thing; Aristotle distinguished between accidental and essential form: a horse is characterized by its being quadrupedalism, not with its color, which is variable and consequently accidental; the latter is not part of its form. According to Aristotle, it is having forms in our intellects that leads to our understanding of the material world: it is not the stone, Aristotle claimed, that is present in the soul, but its form. He calls such mental forms ideas, abstractions, or concepts. Such an interpretation of the notion of form generally has some commonalities with the scientific analysis of mental representations in the brain including what is proposed in chapter 4 of this essay.

Let us jump to the modern time in the realm of biology. In this realm, the infinity of shapes of living organisms provides us with a spectacle of awesome profligacy. D'Arcy Thompson was the first modern biologist who tried to tame such profusion. He used mathematics as a tool to do such a task. In his book, *On Growth and Form* (first published in 1917), he wrote:

> "We have learned […] that our own study of organic form […] is but a portion of that wider Science of Form which deals with the forms assumed by matter under all aspects and conditions, and, in a still wider sense, with forms which are theoretically imaginable[…]. The mathematical definition of a 'form' has a quality of precision which […] is expressed in few words or in still briefer symbols, and these words and symbols are so pregnant with meaning that thought itself is economized; we are brought by means of it in touch with Galileo's aphorism (as old as Plato, as old as Pythagoras, as old perhaps as the wisdom of the Egyptians), that the book of Nature is written in characters of Geometry." (Thompson, 1961, p. 269)

Such a mathematical treatment with "form" carries over into the modern technical treatment of information.

What Thompson had in mind as form could be more clearly expressed by the term "shape." But what are other synonyms for the word form? Paul Young (1987) collected eight synonyms: arrangement, configuration, order, organization, pattern, shape, structure, and relationship. Taking care to interpret the word in the broadest possible way, he sees the term "relationship" as the most general concept that covers all possible applications in mathematics, physics, chemistry, biology, and neuroscience. (Young, 1987, p. 52)[19]

[19] For more information about types of relationships see: von Baeyer (2003, pp. 23-24 )

But information is not the same as form. Not whatever has form could be referred to as information. "Information carries a connotation of activity that is absent from mere form." (von Baeyer, 2003, p. 25) In short, information is the infusion of form, or the flow of relationships.

Getting back to the mathematical treatment of the notion of information, Claude Shannon can be introduced as the progenitor of the original burst of interest in the concept just after World War II. What he was concerned with was the amount of information as a function of one's freedom of choice in selecting a message (Shannon & Weaver, 1949). Although his Mathematical Theory of Communication (MTC) gained a great success with regard to solving engineering problems, as the theory ignores the meaning of the messages transmitted over the communication channel, [20] the theory cannot be regarded as a universal and comprehensive theory of information. So, one should be careful with using Shannon's thesis in the areas for which MTC is not originally formulated. Shannon himself was aware of the limitations of his thesis. He was concerned with the problem solving in control and communication systems, computers, and the like just regarding technical and engineering issues. Indeed, "[o]riginally, the theory mainly created by Shannon was intended as a theory of *information transmission*. As the linguist Bar-Hillel (1969) pointed out, the original term "theory of information transmission"[21] was abridged to "theory of information" and then converted into "information theory." This happened although the basic structure and the scope of the theory remained unchanged." (Gernert, 2006, p. 144)

In 1968, the widespread misuses of Shannon's theory in different areas led to a conference being held in Stranberg with the aim of showing the irrelevance of Shannon's theory in some fields such as psychology, biology, linguistics, and computer science. In the conference, Shannon's thesis was regarded as insufficient to explain everything in the mentioned fields.[22]

[20] So, Shannon's theory is regarded as the syntactic theory of information. I will deal with the triad syntactic, semantic, and pragmatic concepts of information in more detail in the next section.
[21] In the next section, I will go even further calling the theory as the "Mathematical Theory of Data Transmission."
[22] In writing the following paragraphs, till the last paragraph of this section, I have vastly benefited from Gernert (2006)

Concerning the shortcomings associated with the classical mathematical theory of communication, it was natural that some people start searching for the new accounts of information with regard to the meaning of the messages transmitted over the communication channel. So, semantic concepts of information were born. Semantic concepts of information deal with the signs together with the meaning they carry, but regardless of the user or receiving systems. Some people, like Fred Dretske, adopt a more extreme position claiming that not only is meaning independent of the receiver or user, which interprets and makes use of the signs, but also semantically of the producer. But semantic approaches are unsatisfying as, "[f]or instance, our letter "C" will be understood as the letter "S" by anyone using Cyrillic scripture." (Gernert, 2006, p. 147)

So, the concept of information still needs development in order to include recipient for whom (or for which) signs have meanings. Such a comprehensive notion of information, generally called pragmatic information includes other concepts, that is, syntactic and semantic notions of information as its special cases. Recall Gernert's example stated above. The "dependence of the meaning on the receiver's prior disposition can be disregarded only if there is a "tacit agreement", fixing a certain set of possible receivers. Hence, any concept of semantic information will be included in the concept of pragmatic information as the special case in which a class of possible receiving systems has been fixed, such that their uniform way of interpreting the signals is constant and predictable." (ibid)

An early attempt to formulate pragmatic information took place by P. Gäng (1967). Ernst and Christine von Weizsäcker (1972) continued Gäng's attempts proposing the notions *novelty* and *confirmation* in pragmatic information. (Gernert, 2006, p. 149) Before dealing with the latter notions, we should first have an idea of what pragmatic information is according to Weizsäcker family. Carl Friedrich von Weizsäcker (1985, p. 351) defines information as what follows:

"Information is only which is understood."

According to him, understanding is generally defined as the impact upon a receiving system or a recipient, which alters its behavior or structure and is "possible only in sufficiently large composite systems.” (ibid, p. 355f) Large composite systems, dealt with as complex systems in this work, are examined in chapter 3.

Let us get back to the notions of novelty and confirmation in what follows:

> “If a message only repeats material already known to the recipient, then it only conveys confirmation, but no novelty (as in the case of exact duplication). On the other hand, a “breaking news” in an unknown foreign language may contain a great deal of novelty, but because there is nothing in common with the receiver’s prior knowledge, nothing will be understood. This is expressed by the statement that “there is no confirmation”. The latter term is used here in a specific technical sense: It addresses the requirement that a message must be understandable to the receiver, refer to the receiver’s previous knowledge, and contribute at least partially to an existing information demand.” (Gernert, 2006, p. 143)

Ernst and Christine von Weizsäcker (1972) quantified pragmatic information through the following product formula: $P = N.C$

In the above formula, $P$ stands for the measure of pragmatic information and $N$ and $C$ stand for the measures of novelty and confirmation, respectively. As seen, if either the value of novelty or confirmation be equal to zero, the measure of pragmatic information will equal zero meaning that a message conveys no pragmatic information.

Kornwachs and Lucadou (1982) continued developing the above notion of pragmatic information starting with showing that classical Shannon theory is included in pragmatic information as a special case. As further detail of their work is out of the scope of the present essay, I refer the interested reader to Gernert (2006) for more information on the history and development of pragmatic information. For the time being, suffice it to say that P. Gäng, Ernst, Christine, and Carl Friedrich von Weizsäcker, Kornwachs, Lucadou, and Gernert are the protagonists of the notion of pragmatic information.

So, care should be taken that Roederer’s (2005) reading of pragmatic information adopted in this dissertation has its roots in the theses of the pioneers of the notion, although, unfortunately, in his book he never mentions this and does not give references to the outstanding works of the people stated above.[23] Moreover, he has given the notion of pragmatic information a physicalist drive, which has been criticized in the present work. However, Roederer's reading has some advantages as follows: as we will see in the next chapter, he recapitulates the notion of pragmatic information in a way that one can simply see clear-cut distinctions between the systems which work based on pragmatic information, namely, living systems and artifacts, and other systems, that is, natural abiotic systems. He also successfully applies his thesis to the different areas of science, such as biology, physics, and cognitive science, showing how productive the idea is in explaining many phenomena in the mentioned fields. In short, putting the physicalist nature of his interpretation of pragmatic information aside, his reading of pragmatic information is an explanatorily powerful thesis. In the next chapter, we will examine the latter thesis, together with syntactic, semantic conceptions of information, in more detail. Note that where I speak of pragmatic information hereafter, I mean Roederer's interpretation of the notion.

[23] Nor does Roederer mention David Bohm by whom he has inspired a lot as I have noticed.

## 1.4 Summary of the Chapter

There are two quite distinct concepts of mind, which exhaust it: phenomenal and psychological. There are also two distinct concepts of consciousness. The relation between concepts of mind and concepts of consciousness can be illustrated as what follows:

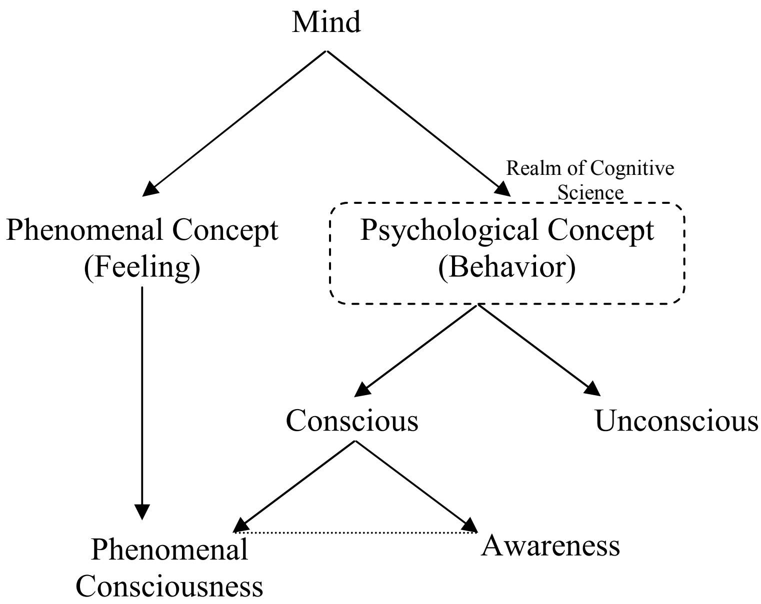


**Figure 1.1.** The interrelations between concepts of mind and concepts of consciousness.

As seen, the phenomenal concept of consciousness subsumes the phenomenal concept of mind, but the psychological concept of mind may be conscious or not. Awareness also exhausts all the various psychological concepts of consciousness enumerated in section 1.2. The above account has been projected based on Chalmers' (1996) account of the relations between notions of mind and consciousness. The hard problem of consciousness, then, can be reformulated as the following: "Even after we have explained the physical and computational functioning of a conscious system, we still need to explain why the system has conscious experiences." (Chalmers, 1996, p. 27) In other words, there is an explanatory gap[24] between awareness and the phenomenal

[24] The term first used by Levine (1983)

consciousness correspondent to it (in figure 1.1, this correspondence has been shown via a dotted line between them). This is a reformulation of the hard problem of consciousness according to which knowing of all facts about physics does not logically lead to the reason to postulate the existence of phenomenal consciousness.

So, a new fundamental psychophysical law is needed for specifying the dependence of the phenomenal facts to the corresponding physical ones. This essay suggests Chalmers' Double Aspect Theory of Information as a natural psychophysical law. Such a suggestion requires the notion of information to be explored in more detail.

For the start, in the course of a selective review of the history of the notion of information (section 1.3), it was argued that, generally, the notion of pragmatic information is the most universal and comprehensive notion of information, as it is concerned with both meaning of the signs transmitted over the communication channel and the receiving system or receiver. The meaning of the signs and the receiving system are ignored in the syntactic and semantic concepts of information, respectively. In the next chapter, I develop my above discussion on the different notions of information.

# Chapter 2
# Three Concepts of Information

## 2.1 Introduction

For the purpose of this essay, an appropriate way of classifying concepts of information can be accomplished due to the well-known triad syntactic, semantic, and pragmatic concepts of information a summary of which was presented in the last chapter. As we will further see in section 2.3, the terms have their roots in Morris' 1938 *Foundations of the Theory of Signs* in which he defined semiotics, the theory of signs, as grouping the triad syntactics, semantics, and pragmatics. In this chapter, I will deal with these three concepts in more detail.

The syntactic concept of information is concerned with the relationships between the signs. As stated earlier, the idea of syntactic concept of information stems from Claude Shannon's (1948) paper titled "A Mathematical Theory of Communication." In his paper, Shannon successfully gives a quantitative way of calculating how compressed a message can be, given the same amount of information to be encoded, and also how fast data can be transmitted over the communication channel, given the same accuracy of the transmitted data. In short, the syntactic concept of information is just concerned with the accuracy of the transmitted symbols in the process of communication. I will argue that Shannon's theory cannot be counted as a genuine theory of information, although it is quantitatively used as a tool for dealing with the amount of transmitted data. The thesis does not tell us anything at all about the concept of information itself. From his viewpoint, the question of whether the transmitted symbols have meaning is irrelevant. It will be shown that the theory can at best be counted as a *branch of probability theory* rather than an "information theory."

At the next step, general specifications of the semantic conceptions of information will be briefly examined. Semantic concepts of information comprise theories about signs and what they stand for. Among different proposals in this area, I will particularly investigate Dretske's 1981 *Knowledge and the Flow of Information* as his information-theoretic approach in explaining "knowledge," "perception," "meaning," and "belief" has had an intense impact in philosophical circles. We will see that Dretske's application of Shannon's Mathematical Theory of Communication results in an unfortunate account of the notions he attempts to explain. Here, the claim is that neither Dretske's account in particular, nor any other semantic concepts of information in general will give us a universal and satisfactory account of information potentially applicable to all sciences.

Finally, the concept of pragmatic information as depicted by J. G. Roederer (2005) will be projected. Generally, pragmatic notion of information comprises relationships between signs and their impacts on those who use them. Two fundamentally different classes of interactions will be identified, that is, "physical interactions" and "information-driven interactions." The former occurs between inanimate objects and the latter between certain kinds of complex systems that form the biological domain as well as some artifacts. The notion of "purpose" will be introduced and will play as a key discriminator between the two-mentioned classes of interactions. The idea, then, would be that pragmatic concept of information contains both semantic and syntactic concepts of information.

It should also be noted that the concept of information is intensely related to two other concepts, namely, "data" and "knowledge." Regardless of how we define the notions, it seems that we cannot have knowledge without information, and information without data. I will begin with the data, information, knowledge, wisdom hierarchy to have given the reader an impression on how they stand in relation to one another.

## 2.2 Data Information Knowledge Wisdom Hierarchy

*Where is the Life we have lost in living?*

*Where is the wisdom we have lost in knowledge?*

*Where is the knowledge we have lost in information?*

T.S. Eliot, "The Rock," *Faber & Faber* 1934.

As said, before dealing with the concepts of information, it seems worthwhile having a short review on the Data Information Knowledge Wisdom hierarchy (the DIKW hierarchy hereafter) without going into details and without trying to define the concepts. Interested reader may find further information in the literature I will mention. However, as far as space and time allow, the ideas of Shannon, Dretske, and Roederer on the above concepts and their relationships will be clarified.

The DIKW hierarchy has intensely attracted the attention of people in many domains, mostly in the three realms of Knowledge Management, Information Science, and Philosophy. "In most Knowledge Management literature the hierarchy is often referred to as the "Knowledge Hierarchy" or the "Knowledge Pyramid," while the Information Science domain refers to the same hierarchy as "Information Hierarchy" or "Information Pyramid" for obvious reasons. Often the choice between "Information" and "Knowledge" is based on what the particular profession believes to be manageable." (Sharma, 2005) Care should be taken that "while the domains of Information Science and Knowledge Management both refer to the DIKW hierarchy, they do not cross-reference." (ibid) Of course, philosophy successfully interacts with both.

In Knowledge Management, Zeleny (1987) and Ackoff (1989) are often considered the initiators of the DIKW hierarchy, while Harlan Cleveland's (1982) *Futurist* article is mentioned as the first projection of the DIKW hierarchy in the field of Information Science. However, as Cleveland himself emphasizes, the first ever mention of hierarchy was poetry. Cleveland cites T.S. Eliot as the one who first put forward such a hierarchy in his "The Rock" quoted above. As seen, the term "data" was not mentioned in Eliot's poem (nor in Cleveland's article). This term was added by others later. Also, in the

DIKW hierarchy, Ackoff believes that before reaching the level of wisdom, there should be some kind of "understanding" or some use of "intelligence." Zeleny adds "enlightenment" to the hierarchy as a higher level, which is gained after wisdom.[25] I deliberately will not go through the latter concepts, namely, understanding, enlightenment, and intelligence (and also wisdom) as the confusion regarding the concepts data, information, and knowledge appears to be enough. Moreover, there is a consensus on the existence of such a hierarchy, but not about other ingredients added to the theory. The major problems associated with the concepts of data, information, knowledge, and wisdom are that the distinctions between the stages are so blurred and there is no agreement on the definitions of the notions.[26] Thus, the best way of dealing with such terms in this essay, as stated earlier, would be to proceed with the people whose ideas on these concepts play a crucial role in the present work.

It is common to show the hierarchy in two ways, that is, by a linear chain and by knowledge (or information) pyramid. Figures 2.1 and 2.2 show these two common ways of depicting the relation between different stages of reaching the destination of wisdom from the raw level of data. In figure 2.1, the DIKW hierarchy is shown as a linear chain through "understanding." Figure 2.2 shows knowledge (Information) pyramid.

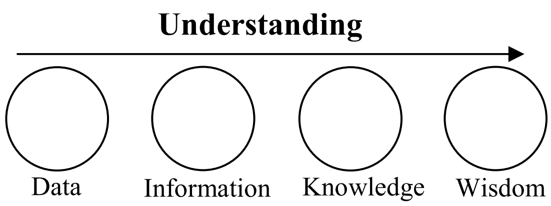


**Figure 2.1.** The DIKW hierarchy: linear chain

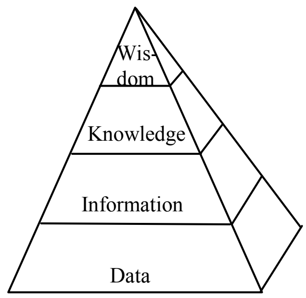


**Figure 2.2.** The DIKW hierarchy: knowledge pyramid

[25]See also Awad & Ghaziri (2004); Burry, Coulson, Preston, & Rutherford (2001); Chaffey & Wood (2005); Rowley (2006, 2007).

[26] Note that what may count "data" in some contexts or for one person may well count "information" in some other contexts or another person. The same goes for the concepts information and knowledge. However, among the above four concepts, it seems that there is less controversy on the definition of the concept "data" (or in singular form "datum"). As we will see in section 2.4, a datum is defined "as a putative fact regarding some difference or lack of uniformity within some context." (Floridi, 2005a) So, wherever the term "data" (or datum) is used in this essay, the reader can reliably bear the above definition in mind.

“Some claim that wisdom is specific knowledge and that information consists of data, but knowledge is not necessarily wisdom and data is not necessarily information. So, wisdom can be seen as a subset of knowledge, which is a subset of information, which is a subset of data.” (Abelli, 2007, p. 19) So, Abelli (ibid) modifies the knowledge pyramid as seen in figure 2.3.[27]

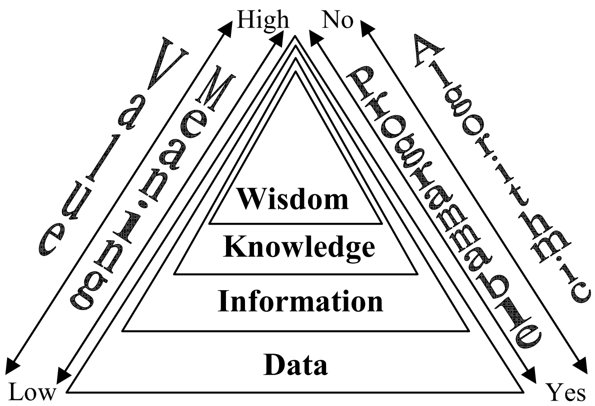


**Figure 2.3.** Abelli modifies knowledge pyramid based on the following works: Awad & Ghaziri (2004); Chaffey & Wood (2005); Rowley (2006).

Let us, in the DIKW hierarchy, consider just the first three concepts, i.e., Data, Information, and Knowledge (the DIK hierarchy). These are the concepts whose relations have been more concerned in the three realms alluded above. I will ignore Wisdom as even a more ambiguous concept. Also, let us, in our consideration of the DIK hierarchy focus on the concept “Information” as it plays a crucial role in the transmission from data to knowledge,[28] whatever they are. Then we will find many interesting theories, among others, Shannon’s Mathematical Theory of Communication discussed in what follows.

[27] I am not sure if the above analysis would be a legitimate one, since the above notions become more and more subjective as we move from “data” toward “wisdom,” becoming more difficult to determine the relations in a strict way. It is also out of the scope of this essay to analyze the most legitimate thesis, which examines the DIKW hierarchy. For the purpose of this essay, suffice it to be aware of the existence of such a hierarchy, regardless of what the definitions of the notions are and whether there is a distinct border between them.

[28] Note that what is considered as data by one can be considered information by another or vice versa. Likewise, what is considered knowledge by one can be considered information by another or vice versa. What counts here is the intermediary role of information. Thus, it seems natural to focus on some relevant theories of information considering the ideas of initiators of such theories on the above relation.

## 2.3 Mathematical Theory of Communication (MTC)

No doubt, the father of mathematical theories of data transmission is Claude Shannon. Such a field was born by his revolutionary paper, titled "A Mathematical Theory of Communication," contributed to the *Bell System Technical Journal* in July and October 1948. The paper attracted immediate worldwide attention such that, both theoretically and experimentally, many scientists and philosophers from different fields found enough material to be engaged in since that time onward. Later also came some other successful theories in this field such as Fisher Information (Frieden, 2004) and Algorithmic Information Theory (Chaitin, 1987). As mentioned before and will be discussed later, calling MTC as information theory, the label which became popular after the projection of MTC, is very misleading and even Shannon himself was not happy with the widespread popularity of such a label in this context. He finally felt urged to counter the widespread mistaken applications of his approach to the other fields, ranged from different areas of natural sciences to the realm of humanities such as society and philosophy.[29]

Initially, Shannon presented a quantitative model of communication as a statistical process. Symbolically, Shannon considers the communication system as illustrated in figure 2.4.

[29] In a paper titled *The Bandwagon* (1956, p. 3), Shannon declares his unwillingness on such a misuse. He writes: "INFORMATION theory has, in the last few years, become something of a scientific bandwagon. Starting as a technical tool for the communication engineer, it has received an extraordinary amount of publicity in the popular as well as the scientific press. In part, this has been due to connections with such fashionable fields as computing machines, cybernetics, and automation; and in part, to the novelty of its subject matter. As a consequence, it has perhaps been ballooned to an importance beyond its actual accomplishments. Our fellow scientists in many different fields, attracted by the fanfare and by the new avenues opened to scientific analysis, are using these ideas in their own problems. Applications are being made to biology, psychology, linguistics, fundamental physics, economics, the theory of organization, and many others. In short, information theory is currently partaking of a somewhat heady draught of general popularity.

Although this wave of popularity is certainly pleasant and exciting for those of us working in the field, it carries at the same time an element of danger. While we feel that information theory is indeed a valuable tool in providing fundamental insights into the nature of communication problems and will continue to grow in importance, it is certainly no panacea for the communication engineer or, a fortiori, for anyone else."

As shown in Fig. 2.4, a desired *message* is selected by the *information source* (*S*) out of a set of possible messages. Such a message may be anything such as words, pictures, music, etc. The task of converting this message into the signal is done by a *transmitter*. From the transmitter, the signal is sent to the *receiver* (*R*) over the *communication channel*. The receiver, as a sort of inverse transmitter, changes the signal back into a message and hands it on to the destination. When, say, Alice talks to Bob, Alice's brain is the information source, her voice mechanism which produces varying sound pressure (as the signal) is the transmitter and air is the channel. Bob's ear together with associated eighth nerve is the receiver and his brain is the destination.

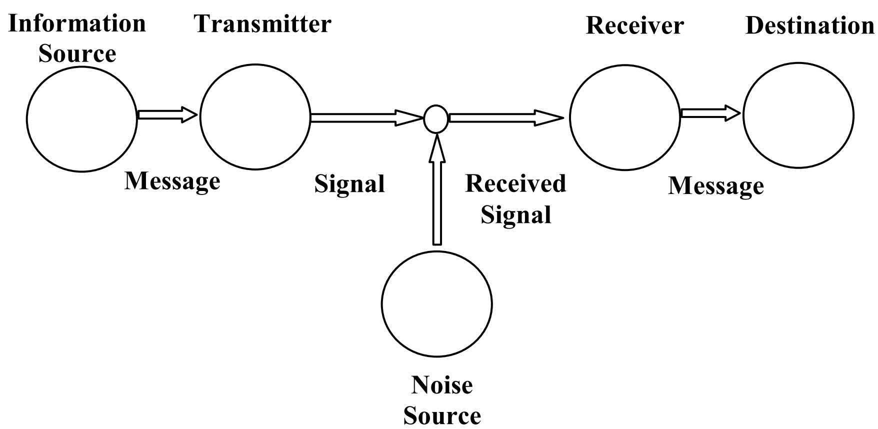


**Figure 2.4.** Shannon's symbolic communication system.

Suppose further that Alice is talking to Bob in a street, which is full of passing cars. There will be some unwanted sounds coming to Bob's ear preventing him from hearing Alice perfectly. These unwanted sounds are called *noise* (*N*), being produced by cars as the *noise source*. The sentences said by Alice but not heard by Bob comprise *equivocation* (*E*). I will deal with all these terms in detail later in this chapter.

As MTC is inherently statistical, it usually begins with probabilistic events like flipping coins, throwing dice, choosing cards, and the like. To provide such probabilistic events, I adopt "classical binary pinball machine," developing the way Roederer (2005) deals with it. Not only does such a machine have all features of the statistical events, but also some characteristics, which emulate other statistical systems stated above. For example, it has controllable components adjusting them will easily lead to the change in the probability of obtaining a certain result. Figure 2.5 symbolically shows such a machine.

The ball will hit the pin when it is released from the source bin. If the machine is perfectly symmetric and after many repetitions, we will find the ball to be equally in each bin 50% of the time. The possible paths, leading to the final result of finding the ball in bin 0 or bin 1 correspond to the realization of the states, say, |0> or |1> respectively.[30] Such a system functions exactly like a fair coin.

Using the binary number system, consisting of zeros and ones, to represent the occurrence of one of the possible states, i.e., to declare in which bin the ball has been found, we need one digit, say, *zero* where we find it in bin 0, and 1 where we find it in bin 1. Therefore, we say that the measure of information, where one state, out of the two equally possible states, is realized is 1 *bit* and we need 1 binary digit to state it.[31]

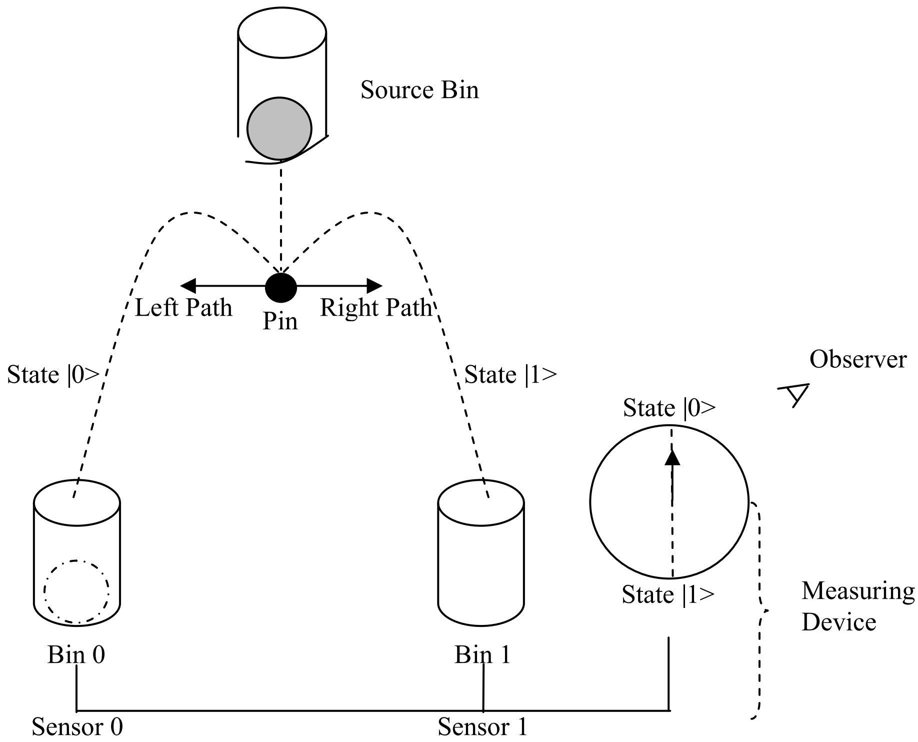


**Figure 2.5.** Standard pinball machine.

Suppose that there are two electronic sensors installed at the bottom of the bins, one of which is activated where the ball hits the correspondent bin. The sensors are connected to

[30] Instead of possible states to be realized, Shannon speaks of the possible *messages* to be selected. The difference is terminological. One can say that in the pinball machine the message is to tell us which final state is realized. I chose to use the term "state," instead of message, on purpose which will be clear later.
[31] In this case, we will see that not only does the occurrence of each of the individual states provide 1 bit of information, but also the information gained on the average.

a pointer whose position shows which of the possible states has been realized, i.e., the bin in which the ball falls. Sensors together with the pointer comprise our measuring device. The measuring device can be either mechanical or electronic or just a combination of both.[32] Whatever it is, the human brain has also to be involved in two stages: 1. *designing the whole system, consisting of the pinball machine and the measuring device, to achieve a certain goal, namely, showing which state has been realized* and [33] 2. *looking at the pointer.*[34] Finding the ball in one of the bins by looking at the record would provide us 1 bit of new information.

Suppose that the pin is shifted slightly to the left such that the state |1> occurs more frequently than |0>. We expect to see the ball in the bin 1 more than 50% of the time. That is, we are less uncertain about the result. To see how it works, let us shift the pin even further to the left such that the only possible state to occur be |1>. Then, finding the state |1> being realized is greeted with a bored "So what?" as we already *knew* the result. Therefore, we have gained no new information. On the other hand, the occurrence of the state |0> will lead to the exited "Oh my God!." The latter would be highly informative.

To sum up, the more likely the occurrence of an event is, the less informative it would be. This holds generally. For example, you find an article informative because it is unlikely to come up with such ideas yourself. Similarly, a genome, which exists in a context of proteins leading to the development of some creature contains a lot of information, if that end structure (the living creature) is very unlikely to have appeared in the absence of the genome (thus, human DNA contains more information than virus DNA, because a human is a more improbable creature than a virus).

So far, I have been concerned with the occurrence of the individual states: how much information is gained where each of the specific states |0> and |1> is realized. We saw that in case of setting the standard pinball machine symmetrically while running the

[32] I included measuring device because of a certain purpose clarified later. It could be that we just looked at the final state without using a measuring device. However, I emphasize that in both cases, the human brain has to be involved.

[33] We will soon see that if there is no goal and the measuring device works just by chance, there will be no information and information processing at all.

[34] There is also another stage in which the human brain is involved, i.e., resetting the system once each operation has taken place. It can be added as a separate stage to the mentioned two stages. However, I included it in the first stage: the pinball machine can be designed such that after each operation it resets itself.

system with sufficient repetition, we find the ball to be equally in each bin 50% of the time, leading to the conclusion that the probability of the realization of each state is 50%. Shifting the pin to the left decreases the number of times we may find the ball in bin 0, reducing the probability of the realization of the state |0>. When we speak of the probability of the realization of one of the states |0> and |1>, we are indeed speaking of the *particular* states being selected out of two possible states. However, MTC does not concern itself with the occurrence of specific or particular states, but with the average amount of information generated considering the system as a whole. It would be misleading to say each state *carries* 1 bit of information, because in MTC "the concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit amount." (Shannon and Weaver, 1949, p. 9) We will soon see that the measure of information, on the average, is maximum where the states are equiprobable to occur (in our pinball machine, when it is symmetric).

With the above prior knowledge in mind, we are ready to go through the *mathematical* description of what was mentioned above qualitatively, as Shannon's theory of communication is just a quantitative theory measuring the average amount of information transmitted from the source to the receiver being quite silent about what information is. This is why it is preferred to be called MTC, rather than Shannon Information, another misleading expression which is frequently used in relevant literature.

Suppose that we operate our standard pinball machine $N$ times $(N \to \infty)$ under exactly the same conditions. If $N_0$ and $N_1$ are the number of the realizations of the states |0> and |1> respectively, then $p_0 = N_0/N$ and $p_1 = N_1/N$ are defined as the correspondent probabilities of the realization of these states. Note that $p_0+p_1=1$. As we saw, if the machine functions symmetrically, then $p_0=p_1=1/2$, that is, both states are equiprobable to occur. Shifting the pin to the left, as described above, decreases $p_0$, giving rise to $p_1$, leading to the realization of |0> and |1> to be more and less informative respectively. In the most extreme case in which the pin is shifted such that the only possible state is |1>, we have $p_1=1$ and $p_0=0$. As mentioned, in this case, the realization of |1> corresponds to "So what?" as we already *expected* it to happen. The realization of the state |0> will also lead to "Oh my God!" as it is so surprising. To find a mathematical expression for the *informativeness* of the probabilistic events, we have to consider the general case, namely, $p_0 \neq p_1$. Note that here the term informativeness refers to the amount of information gained by the realization of one state out of the possible ones. [35]

Consider $I_0$ as a quantity representing the informativeness of the state |0> when it is realized. Likewise, $I_1$ corresponds to the amount of informativeness of the realization of |1>. Any desired mathematical expression for the calculation of $I_0$ and $I_1$ must fulfill the above-mentioned conditions, that is, for $p_0=p_1=1/2$ we should gain $I_0=I_1=1$b. Furthermore, as $p_0$ decreases, giving rise to $p_1$, since $p_0+p_1=1$, $I_0$ should increase whereas $I_1$ should decrease, both monotonically. Therefore, it seems that any mathematical expression for informativeness should be a function of probability, $I(p)$, such that if $p_i<p_j$, then $I_i>I_j$. In the most extreme case, where $p_i \to 0$, we should get $I_i \to \infty$, and where $p_i=1$, we have to obtain $I_i=0$. The most natural choice for $I(p)$, according to Shannon (1948), would be the logarithmic function[36], that is:

$$I_i = -K \ln p_i \tag{2.1}$$

The unit of $I_i$ corresponds to the choice of logarithmic base. In case of choosing base 2, the resulting units are called binary digits, or more briefly *bits,* a term first put forward by

[35] In Shannon's vocabulary, the amount of information gained by the selection of an individual message is called *surprisal* of that particular event. Since such a term has a psychologistic sense, following Floridi (2005), I chose informativeness for such a quantity.

[36] In what follows, Shannon explains why the logarithmic function is the most natural choice. Note that such a choice is just reasonable, without being emerged from a physical principle or any other natural law.

"1. It is practically more useful. Parameters of engineering importance such as time, bandwidth, number of relays, etc., tend to vary linearly with the logarithm of the number of possibilities. For example, adding one relay to a group doubles the number of possible states of the relays. It adds 1 to the base 2 logarithm of this number. Doubling the time roughly squares the number of possible messages, or doubles the logarithm, etc.
2. It is nearer to our intuitive feeling as to the proper measure. This is closely related to (1) since we intuitively measure entities by linear comparison with common standards. One feels, for example, that two punched cards should have twice the capacity of one for information storage, and two identical channels twice the capacity of one for transmitting information.
3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm but would require clumsy restatement in terms of the number of possibilities." (Shannon, 1948, p. 379)

J. W. Tukey. Thus, concerning our binary pinball machine in which for $p_i$=0.5 the informativeness is 1 bit ($I_i$=1 bit) for the occurrence of each of the individual states, equation (2.1) turns to:

$$I_i = -\log_2 p_i \text{ (bits)} \qquad (2.2)$$

As $K = 1/\ln 2$.

One of the most important advantages of choosing the logarithmic function is its additive property. Suppose that we operate the binary pinball machine three times and want to find out the total informativeness ($I_{Total}$) of the successive occurrences of the following states: |0>, with *a priori* probability $p_a$ in the first run; |1>, with *a priori* probability $p_b$ in the second run; and again |1>, with *a priori* probability $p_c$ in the third run. The overall probability of the realization of these three successive states is the product of the independent probabilities $p_a$, $p_b$, and $p_c$. Therefore, the total informativeness through the successive realizations of the mentioned states (|0>, |1>, and |1>) would be:

$$I_{Total} = -\log_2(p_a p_b p_c) = -\log_2 p_a - \log_2 p_b - \log_2 p_c = I_a + I_b + I_c.$$

So, the total amount of informativeness simply is the sum of the values of the informativeness in each run.

Note that "the word *information*, in this theory [MTC], is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning.

In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information." (Shannon and Weaver, 1949, p. 8, the bracketed term "MTC" above is mine) Why is this so? The answer is that Shannon is just concerned with solving the problems associated with the accuracy of the symbols of communication transmitted, not meaning, purpose, and the like. These concerns require the theory to address the *Average Informativeness*, not the information gained through the realization of the individual states. In what follows, I will address the average informativeness as used in MTC. But, first, some words about the "average value" in general.

Suppose a class of students takes a quiz that has six questions and that the students receive the following grades, out of 100: 10, 30, 30, 50, 50, 60, 60, 60, and100. Calculating the average grade, we get:

$$G = \frac{10+30+30+50+50+60+60+60+100}{9} = 50$$

In general, we can define the average value as the following:

$$I = \frac{\sum_{j=1}^{N} i_j}{N} \tag{2.3}$$

Where $I$ is the average value of (the large enough) $N$ discrete values $i_j$ obtained or possibly will be obtained.

Another way of putting this is to calculate the average grade by multiplying each grade by the number of times each grade is obtained. That is:

$$G = \frac{1(10)+2(30)+2(50)+3(60)+1(100)}{9} = 50$$

In general, equation (2.3) can be written as the following:

$$I = \frac{\sum_i N_i i}{N} = \sum_i \left(\frac{N_i}{N}\right) i$$

$N_i$ is the number of times $i$ is obtained or is probable to be obtained. If $N$ is large enough, $\frac{N_i}{N}$ is the probability ($p_i$) of obtaining the value $i$. Therefore, we can rewrite above formula as the following:

$$I = \sum_i p_i i$$

Coming back to our discussion on MTC, one can say that the average informativeness is the average amount of information generated by the device and corresponds to the average amount of information one expects to gain *before* the outcome is actually determined. It shows how much prior uncertainty (data deficit) we have about the outcome. In our pinball machine, consider $I_0$ as the informativeness associated with the individual state |0>. Similarly, $I_1$ is the informativeness associated with the state |1>. $p_0$ and $p_1$ are prior probabilities of the realization of the states |0> and |1> respectively. Note that $p_0+p_1=1$. The average informativeness ($H$) is calculated as the following:

$$H = +p_0 I_0 + p_1 I_1 = -p_0 \log_2 p_0 - p_1 \log_2 p_1$$

In general, we have:

$$H = -\sum_{i=0}^{n-1} p_i \log_2 p_i \qquad (2.4)$$

Where: $\sum p_i = 1$

Shannon calls $H$ as *entropy*[37]. He states that $H$ plays a central role in his theory as measures of information, choice, and uncertainty. (Shannon, 1948, p. 400) The entropy in case of just two possibilities $p_0=p$ and $p_1=1-p$ (as $p_0+p_1=1$) is illustrated above in figure 2.6.

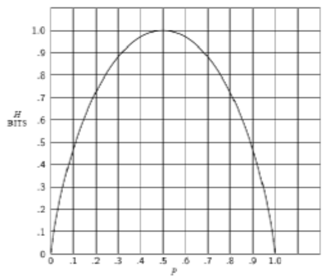


**Figure 2.6.** Entropy in case of two possibilities *p* and 1-*p*.

[37] It seems that this confusing label has been recommended to Shannon by von Newman. One should distinguish between the concept of entropy in MTC and in thermodynamics in order to avoid wrong analogies and conclusions. I will beware of using the term. Instead, I will use "average informativeness." Note that this quantity shows the average amount of our uncertainty (data deficit) about the result of the realization of one state out of possible states. This quantity is also the average amount of information generated at the source by the realization of such a state, or putting it in another way, it is the average amount of information the observer gains when she looks at the result of the realization of that state.

Shannon (1948) counts several interesting properties for $H$, some of which are examined here:

1. $H$ will be zero if and only if all probabilities but one are zero. Under such a condition, we *already know* the result of the operation of the machine, as there is just one possibility. In this case, the average informativeness is zero as there is no uncertainty about the result. Note that if there are just two possibilities $p$ and $1$-$p$, the average information *available* before the determination of the result is *1-H.* This quantity shows how much *prior knowledge* we do have about the result of the system under consideration. $H$ being zero means that we have 1 bit prior knowledge about the result of the operation of the machine. Thus, what remains as the quantity of the average informativenss gained after the operation would be zero.

2. Suppose that there are $n$ possible states one of which is to be realized after the operation of the machine. $H$ would be maximum when all possible states, or results, are equiprobable to occur.[38] Recall equation (2.4):

$H = -\sum_{i=0}^{n-1} p_i \log_2 p_i$, where: $\sum p_i = 1$. As all the $n$ probabilities of the realization of the states are equal: $H = -np_i \log_2 p_i$. But: $p_i = 1/n$. So: $H = \log_2 n$ (2.5)

This is also intuitively the most uncertain situation as we have no prior knowledge about the priority of any of the possible states to occur. As Brillouin (1962, p. 8) states, “every type of constraint, every additional condition imposed on the possible freedom of choice immediately results in a decrease of information,” as it gives us some prior knowledge enabling us to predict the result with more certainty. Consequently, the new information gained after machine’s operation would be less, comparing with the case in which we had no prior knowledge at all.

[38] The reader should bear it in mind that although the greatest *average* informativeness is gained when the possibilities are equally probable to occur, the greatest informativeness values, that is, the amount of information gained through the realization of *individual* states, are to be gained when the possibilities are not equiprobable. We saw that in the extreme case of binary pinball machine in which $p_0$ approached zero, the informativeness $I_0$ approached infinity.

To see how the constraints affect the free choice, decreasing *H*, consider English language. To build up English sentences, we need 27 symbols: 26 letters plus the "blank" or spacing. Further, suppose that these symbols could be used without any constraints with *equal a priori* probability. Suppose that we have used *G* symbols in our sentence. Using equation (2.5), the average informativeness associated with the sentence could be calculated as:

$H = G \log_2 27 = 4.76G$, or just 4.76 per letter.

However, we know that different letters do not happen with equal *a priori* probability. For example, according to Brillouin (1962, p. 5) the *a priori* probability of the occurrence of space or "blank" is 0.2, letter E: 0.105, letter Z: 0.001, and so forth. So, we expect to have *H* to be less than 4.76. Actual *H* in English language is 4.03. (ibid, p. 8) Moreover, we should also take it into account that there is always correlation between letters in languages, that is, when a specific letter in a word occurs, the probability of other certain letters to follow is not the same as the *a priori* probabilities that these letters occur. For example, the probability of the letter "p" being followed by "h" is much greater than that it is followed by "z." Also, given the complex "sio," the probability that "n" will follow is much higher than "h." Shannon defines this kind of correlations as *redundanc*y.[39] "Redundancy refers to the difference between the physical representation of a message and the mathematical representation of the same message that uses no more bits than necessary." (Floridi, 2005) I will again come back to this point.

So far, I have discussed some of the most important elements of MTC. However, to know how all these elements relate to each other, we need some clarifying examples. Before doing that, we need to deal with the term "source" as used in MTC. For it plays a crucial

[39] It is perhaps interesting to quote Weaver's remarks on redundancy in English language: "It is most interesting to note that the redundancy of English is just about 50 per cent, so that about half of the letters or words we choose in writing or speaking are under our free choice, and about half (although we are not ordinarily aware of it) are really controlled by the statistical structure of the language. Apart from more serious implications, which again we will postpone to our final discussion, it is interesting to note that a language must have at least 50 per cent of real freedom (or relative entropy) in the choice of letters if one is to be able to construct satisfactory crossword puzzles. If it has complete freedom, then every array of letters is a crossword puzzle. If it has only 20 per cent of freedom, then it would be impossible to construct crossword in such complexity and number as would make game popular. Shannon has estimated that if the English language had only about 30 per cent redundancy, then it would be possible to construct three-dimensional crossword puzzles." (Shannon and Weaver, 1949, p. 12)

role in the following examples. Roughly speaking, a source is some mechanism or process the result of which is the reduction of *n* possible states to one through the realization of that chosen state.

Suppose that we have a binary pinball machine, this time with more pins, as illustrated in figure 2.7. If the machine is completely symmetric, there will be four possible states, each with the probability of 0.25 to occur. We also need at least two digits to label each bin or state. Concerning our definition of the source, the pinball machine can be considered a source. We also have a measuring device, consisting of sensors and a pointer. The measuring device *is wired* to the pinball machine such that the realization of each state in the pinball will result in the aim of the pointer to the correspondent state. Again, *regardless* of whether the outcome in the measuring device is dependent on the pinball machine or not, the measuring device, too, can be regarded as a source of information. For it consists in the reduction of alternatives to just one realized state. Consider $I(s)$ as the average informativeness at the pinball ($S$). Likewise, $I(r)$ is the average informativeness at the measuring device ($R$). As in our case, $S$ and $R$ are connected to each other, there is an interdependence between them. The information generated at $R$ is the information *about* or *from* $S$.[40] Thus, it seems reasonable to rearrange our terminology, considering $R$ as the *receiver* of information and defining $I_s(r)$ as the average amount of information (average informativeness) generated at $S$ and received at $R$.

In the source, we have four equiprobable states. The informativeness associated with each individual state, according to equation (2.2), is 2 bits. Using equation (2.5), we can simply calculate $I(s)$, which would again be 2 bits. Thus, comparing with the pinball machine with just two possible states, there is more initial uncertainty about the occurrence of one of the states and more information is gained (2 bits) when the result is determined. Note that when the states are equiprobable to occur, the informativeness associated with each individual state is equal to the average informativeness. This holds generally. Recall our example about the average grade of the students in page 50. If all grades were the same, each individual grade would be equal to the average grade.

[40] $I(s)$ and $I(r)$ should still be regarded as distinct quantities.

The situation is the same in the receiver. We have four equiprobable states, one of which will be realized by aiming the pointer to that state when the ball is found in the correspondent bin. Again, the informativeness associated with each individual state (2 bits) is the same for all four states and equals to *I(r)*, that is, the average informativeness at the receiver. As all information in the receiver is the information received from the source, $I_s(r)$ would also be 2 bits. In such a case, all information generated in the source has been received by the receiver. Figure 2.8 shows the whole story.

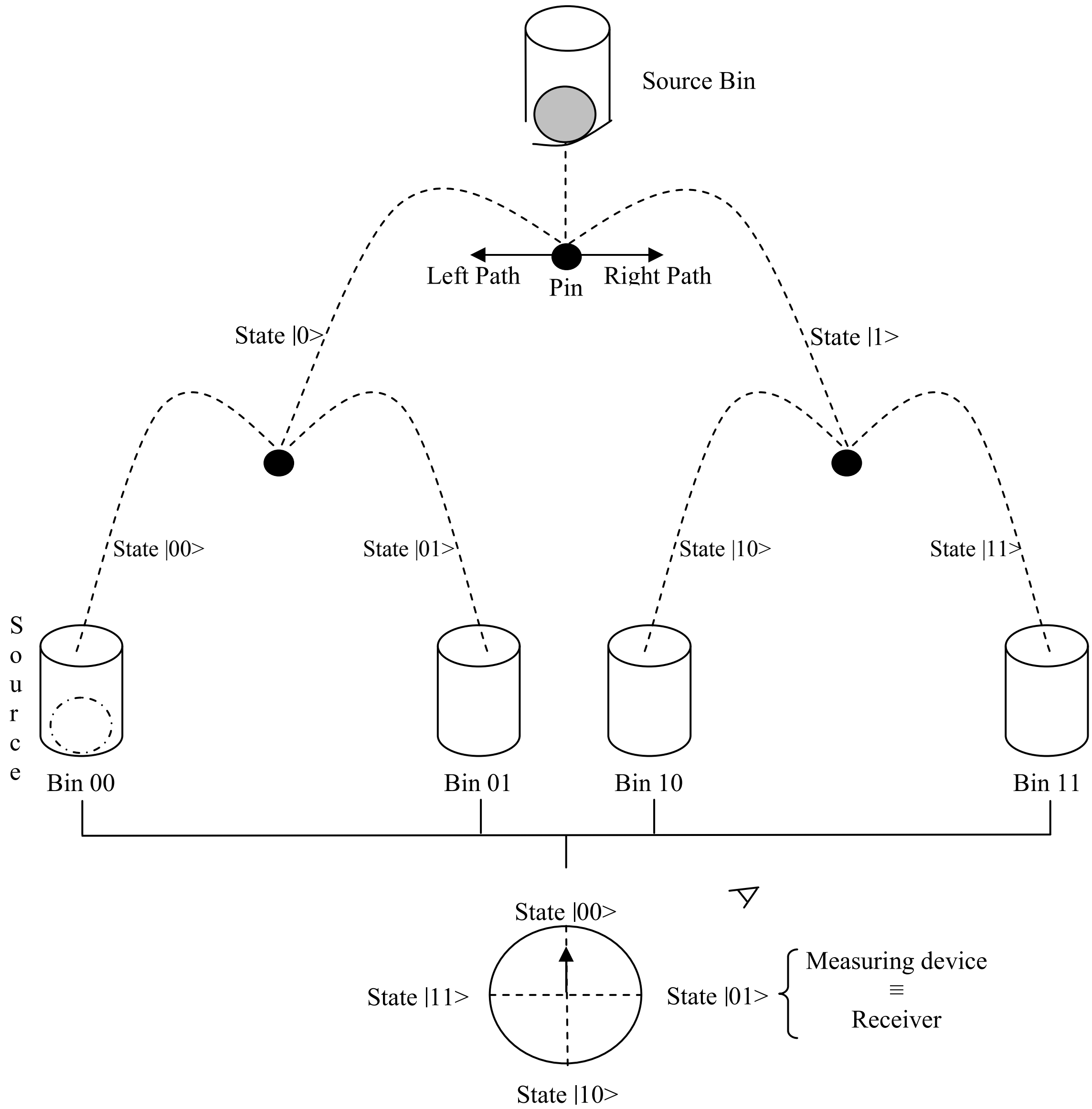


**Figure 2.7.** Pinball machine with four equiprobable states connected to the measuring device with correspondent states to be realized when the ball is found in one of the bins. The machine is considered as the source of information and the measuring device is counted as the receiver of this information.

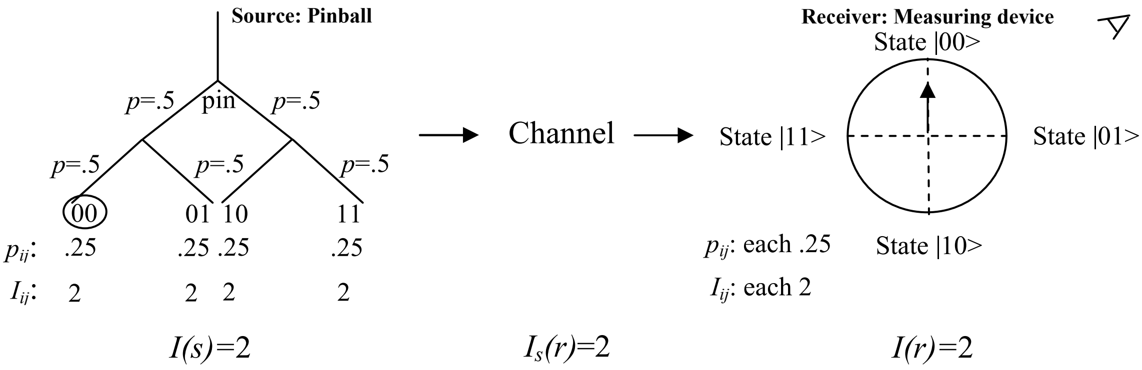


**Figure 2.8.** Pinball machine with four equiprobable states as the source, being connected to the measuring device as the receiver. The probability of the occurrence of each individual state in the source is .25. The informativeness associated with each individual state in the source would be 2 bits. *I(s)* as the average informativeness in the source is also equal to 2 bits. The same goes for the receiver. The informativeness associated with each individual state, as well as the average informativeness, is 2 bits. All this information is the information that has been transmitted from the source. Thus, $I_s(r)$ is also 2 bits.

There can be another situation in which the source is the same as the previous example, but the setup of the measuring device has changed such that it has just three states: either the ball is found in the bin 00 or in 01, the pointer aims at the same state, that is, the state I would label as "|00> or |01>." In case of the realization of state |10> or |11> in the source, the pointer in the measuring device will aim at the correspondent state, namely, |10> or |11> respectively. Figure 2.9 illustrates such a situation. Again, as far as the source is concerned, the average informativeness would be 2 bits. That is, the average informativeness generated by the realization of one of the states, out the four equiprobable states, is 2 bits. However, not all the information generated in the source is transmitted to the receiver. Suppose the pointer aims at the state "|00> or |01>." If the observer just looked at the measuring device without being informed of the state realized in the bin which actually contained the ball, she would still have some uncertainty about the result as she would not know which of the following states was realized, |00> or |01>. So, although all individual states are equally probable to happen in the source, resulting in the generation of 2 bits of information as the average informativeness, the prior probabilities of observing the pointer aiming at each of the states are not the same. The *a priori* probability of finding the pointer aiming at the specific state "|00> or |01>" is 0.5, while the prior probability of finding the pointer aiming at the state |10>, as well as the state |11>, is 0.25. In this case, as mentioned above, the average informativeness would not be maximum as the states do not have equal prior probability to occur. Equation (2.4) tells us that the average informativeness at the receiver is 1.5 bits. Recall that the average informativeness shows the average amount of our uncertainty about the result of the realization of one state out of the possible states before the operation of the machine. This quantity also corresponds to the average amount of information generated at the source of information (either pinball or the receiver) by the realization of such a state.

While, on the average, 2 bits of information is generated at the source, 1.5 bits of this information is transmitted to the receiver. I labeled the latter as $I_s(r)$. The remainder, namely, 0.5 bit is the average amount of information generated at the source but not received at the receiver. This amount is defined as *equivocation*.

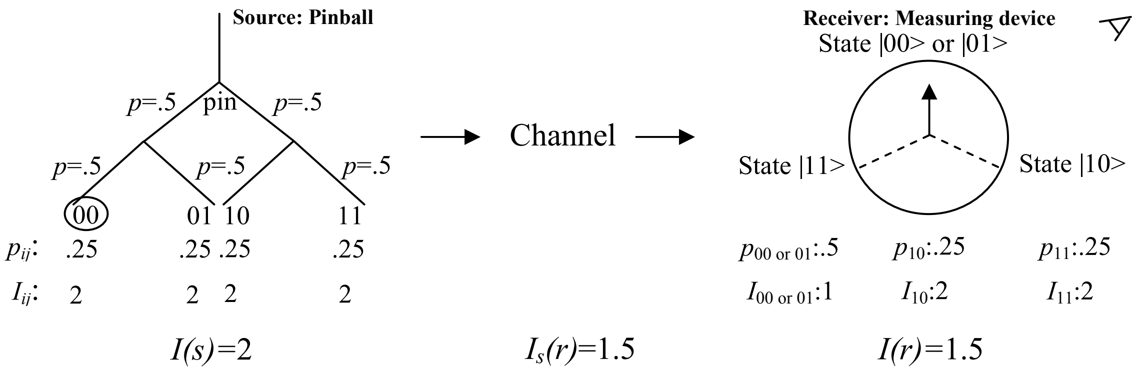


**Figure 2.9.** Pinball machine with four equiprobable states as the source, being connected to the measuring device as the receiver. States in the measuring device do not have equal prior probability. So, while 2 bits of information is generated at the source on the average, the information received at the receiver is 1.5 bits. 0.5 bit of information has been generated at the source but not transmitted to the receiver. This amount of information is called equivocation.

Finally, consider our system as illustrated in figure 2.10, again, consisting of the pinball machine as the source and the measuring device as the receiver. Suppose further that there are four equally probable states, both in the source and receiver, but this time the connection between them is cut off. What happens in the source has nothing to do with the receiver, that is, information received at receiver is not the information generated at and transmitted from the source to the receiver. We let the pinball machine operate. Each individual state has *a priori* probability of 0.25 to occur. The same goes for the measuring device. Suppose that we find the ball at bin 00. Exactly at the same time, the measuring device aims at the state |00>, under the influence of some third system, say, an electromagnetic device. In this way, the average amount of information generated at the source (the average informativeness), as well as the receiver, is 2 bits. But the average amount of information generated at the receiver is not the information generated at the source and transmitted to it. The average amount of information received at the receiver but not generated at the source is called *noise*. Here, we have 2 bits of average informativeness as equivocation. $I_s(r)$ in this case is zero. Figure 2.11 also depicts the situation associated with figure 2.10 in more detail.

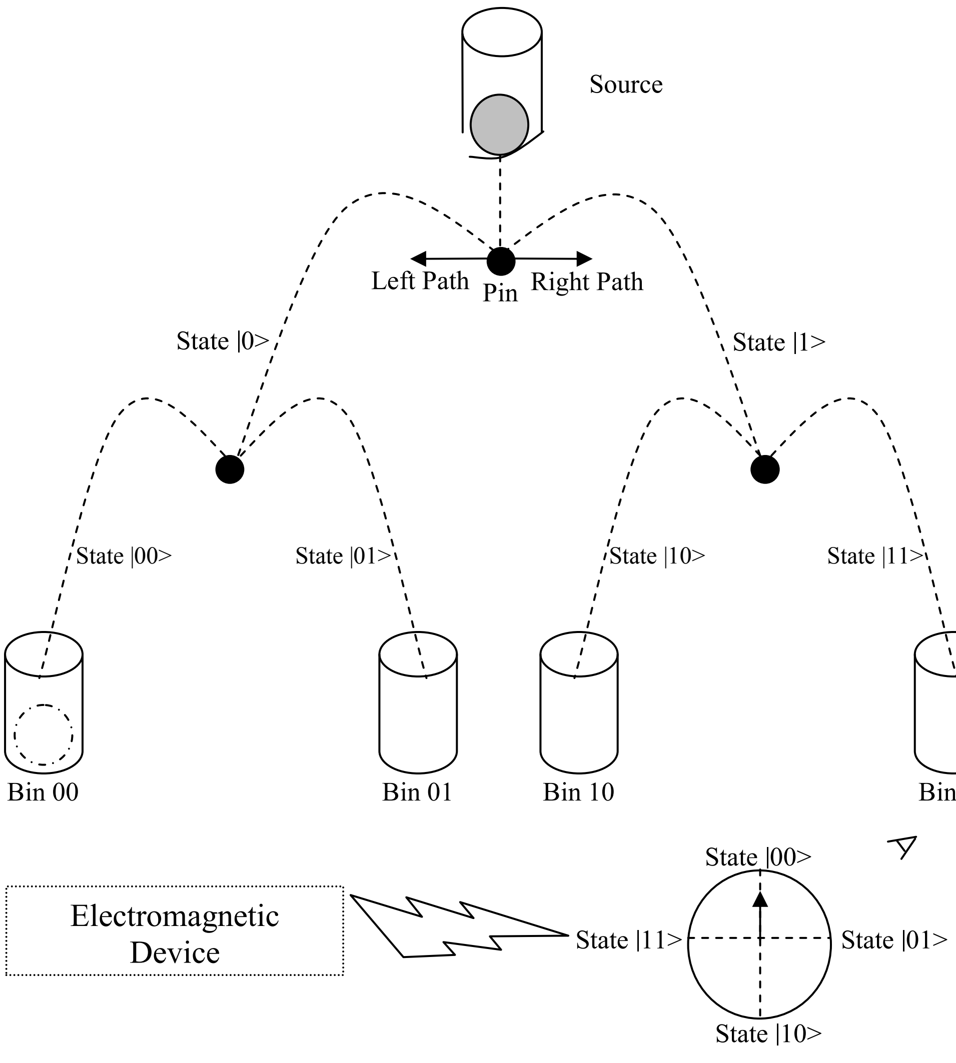


**Figure 2.10.** There is no connection between the pinball machine and the measuring device. At the same time that the pinball machine operates resulting in having the ball in bin 00, the measuring device, under the effect of some third electromagnetic device, shows the state |00>. We have 2 bits of average information generated at the source. But this amount of information has not been received by the receiver. It is equivocation. The average amount of information generated at the receiver is also 2 bits. But it is not the amount of information generated at and transmitted from the source to the receiver. It is noise. $I_s(r)$ in this case would be zero.

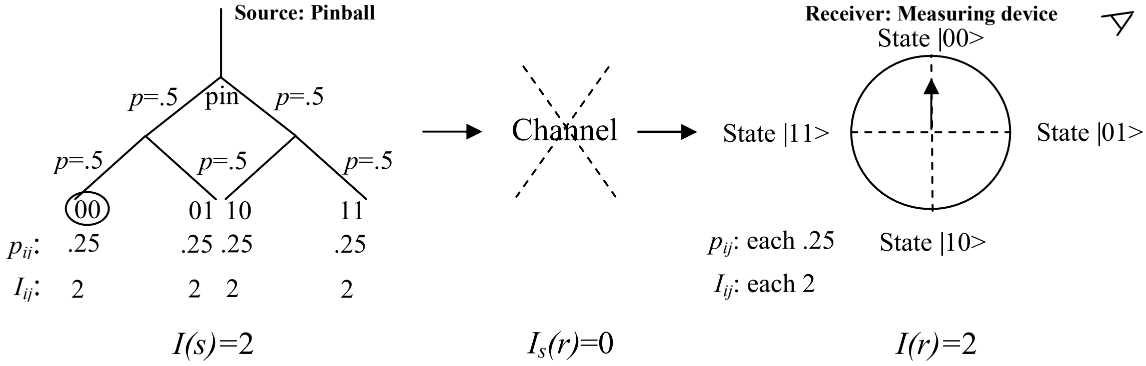


**Figure 2.11.** The situation associated with figure 2.10 in more detail.

It seems worthwhile summarizing the last three examples through appropriate diagrams showing the relation between the following quantities. Consider *I(s)* as the average informativeness generated at the source, *I(r)* as the average informativeness generated at the receiver, $I_s(r)$ as the average informativeness generated at the source and received at the receiver, *E* as equivocation, and *N* as noise. In the first example associated with figures 2.7 and 2.8, equivocation as well as noise are zero and all information generated at the source is received at the receiver. Figure 2.12 illustrates the above-mentioned situation. The closed areas without any label are devoid of information.

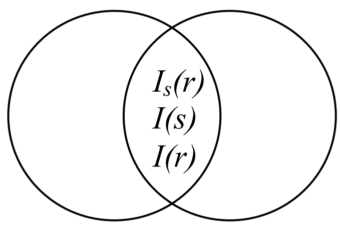


**Figure 2.12.** The appropriate diagram associated with figures 2.7 and 2.8. All average amount of information generated at the source is received at the receiver. *E* and *N* are zero.

The diagram associated with the situation depicted in figure 2.9 can be seen below in figure 2.13. There remains 0.5 bit of average informativeness, which has been generated at the source but not received at the receiver. This amount is equivocation (*E*). The average informativeness generated at the source and received at the receiver ($I_s(r)$) would be as the following:

$$I_s(r) = I(s) - E \qquad (2.6)$$

Similarly, one can write equation 2.7 below in which *I(r)* is the average informativeness generated at the receiver and *N* is noise.

$$I_s(r) = I(r) - N \tag{2.7}$$

In the equation 2.7, *I(s)* is the average informativeness generated at the source. It amounts to the sum of *E* and $I_s(r)$. Note that in this example, the average informativeness transmitted to the receiver corresponds to the average informativeness generated at the receiver. Therefore, there is no noise.

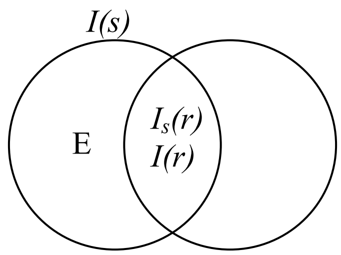


**Figure 2.13.** The appropriate diagram associated with figure 2.9. Not all of the average information generated at the source is transmitted to the receiver. There remains some as equivocation. In this case, noise is zero.

The reader should bear it in mind that in all previous examples concerning the pinball machine *connected* to the measuring device, we assumed the communication channel to be noiseless. In other words, the communication channel does not contribute any noise to the system.

The last example I dealt with was the situation shown in figures 2.10 and 2.11 in which no information from the source is transmitted to the receiver. All information generated at the source remains there as equivocation. On the other hand, all the average information generated at the receiver is received from a system other than what was introduced as the source. Thus, all information generated at the receiver is noise. Figure 2.14 shows such a situation.

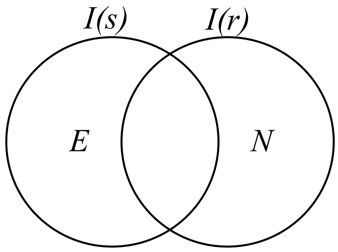


**Figure 2.14.** The diagram associated with figures 2.10 and 2.11. All information generated at the source remains there as the equivocation. All information generated at the receiver corresponds to a third system other than what was introduced as the source. So, it is noise. $I_s(r)$ is zero.

As seen in above examples, the amount of dependence between source and receiver can simply been explored by determining the values of $E$ and $N$. We saw that if the dependence between source and receiver was maximum (figures 2.7, 2.8, and 2.12), then the values of $E$ and $N$ were minimum ($E$=$N$=0) while $I_s(r)$ was maximum ($I_s(r)=I(s)=I(r)$). On the other hand, if they are totally independent, the values of $E$ and $N$ are maximum, that is, $E=I(s)$ and $N=I(r)$. Thus, the value of $I_s(r)$ would be zero.

Figure 2.15 shows how the quantities $I(s)$, $I(r)$, $I_s(r)$, $E$, and $N$ relate to each other in general.

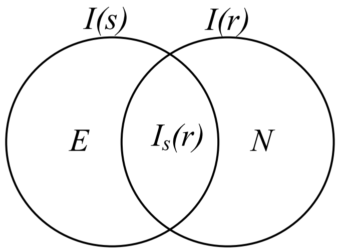


**Figure 2.15** shows how the quantities $I(s)$, $I(r)$, $I_s(r)$, $E$, and $N$ relate to each other in general.

It is important here to mention that "$E$ and $N$ are not only the function of the source and the receiver, but also of the communication channel. The introduction of the communication channel leads directly to the possibility of errors arising in the process of transmission: channel $CH$ is defined by the matrix $[p(r_j / s_i)]$, where $p(r_j / s_i)$ is the conditional probability of the occurrence of $r_j$ given that $s_i$ occurred, and the elements in any row must sum to 1." (Lombardi, 2004, p. 108) Note that $s_i$ and $r_j$ are individual states at the source and receiver respectively. Mathematically, $E$, $N$, and Channel Capacity (C) are defined as the following:

$$E = \sum p(r_j) \sum p(s_i / r_j) \log_2 1 / p(s_i / r_j)$$
$$E = \sum \sum p(r_j, s_i) \log_2 1 / p(s_i / r_j) \tag{2.8}$$

$$N = \sum p(s_i) \sum p(r_j / s_i) \log_2 1 / p(r_j / s_i)$$
$$N = \sum \sum p(s_i, r_j) \log_2 1 / p(r_j / s_i) \tag{2.9}$$

$$C = \max I_s(r) \tag{2.10}$$

In equation (2.10) the maximum should be taken over all the possible distribution $p(s_i)$ at the source.[41]

Concerning all this, one can find that MTC's concern is to answer the following question: "how accurately can the symbols of communication be transmitted?" (Shannon and Weaver, 1949, p. 4) To answer such a question, one should concern herself with two sub-questions:

1. How compressed can a message be, given the same amount of information to be encoded?

2. How fast can data be transmitted over the communication channel, given the same accuracy of the transmitted data?

The first question has to do with data compression, as a pure technical issue, and its solution has to do with the "average informativeness." In the second question, "channel capacity" matters.

Compression (question 1) comes about by decreasing redundancy in the system. At first glance, it may seem that redundancy is an unwanted thing. However, it does not hold

[41] Reader with no taste for mathematical description of the communication theory stated here can simply skip these relations. Suffice it to know that $E$ and $N$ are functions of the communication channel.

necessarily, as redundancy can help to counteract equivocation and noise, securing the accuracy of the original message from sender to receiver (question 2). In other words, some degree of redundancy can possibly counterbalance the noise and equivocation arisen by the physical process of communication and the environment.

Let us now explore the place of MTC in the DIKW hierarchy. Consider we are using our pinball machine as illustrated in figure 2.7 in order to choose the winner of one million dollars out of four persons who have correctly answered a question posed on a TV show and reached the final. Each state to occur corresponds to one of these four finalists as the lucky winner. According to MTC, the occurrence of each state generates 2 bits of new information, regardless of how important the realization of such a state is for the informee (and for other audience watching the show) and thus regardless of the meaning contained in the realization of that state. All matters in the mathematical theory of communication is the message (or symbol) which is to be selected out of the possible ones. For MTC, the information generated through writing this essay does not differ from the information generated by walking my dog on the keyboard of my computer typing sequences of meaningless letters, if the number of the symbols typed is the same as it is in what I have written here! MTC is a theory about information without concerning meaning.

> "[…] information relates not so much to what you do say, as to what you could say. The mathematical theory of communication deals with the carriers of information, symbols and signals, not with information itself. That is, information is the measure of your freedom of choice when you select a message." (Shannon and Weaver, 1949, p. 8-9)

Therefore, as Floridi (2005) emphasizes, MTC can be counted as a *branch of probability theory* rather than "information theory." Dretske (1981, p. 1) also asserts that MTC is concerned with the amounts of information transmitted to the other desirable points, not with the information that comes in those amounts. One can even go further if she explores the place of MTC in the DIKW hierarchy: MTC is just concerned with the transmission of signs or massages regardless of their meaning. Such a treatment with signs mostly equalizes them with the concept of "data" in the DIKW hierarchy, as the data part is the most objective and physical part of the mentioned hierarchy, being less entangled with meaning in principle. It is acceptable to have meaningless data, but it is nonsense to speak of meaningless knowledge. So, as we move towards the higher notions of the hierarchy, namely, information, knowledge, and wisdom, the concepts get more pregnant with meaning.[42] Hence, in the aforementioned hierarchy, Shannon's thesis only gives an acceptable account only regarding the notion of data, not the other three notions. Thus, perhaps changing the label of the theory from the "mathematical theory of communication" to the "mathematical theory of data transmission" be even more expressive.

At this stage, another problem arises: according to the above, Shannon's thesis ignores the notion of knowledge as irrelevant in his theory while the notion of prior probability lies at the roots of his theory. Prior probability is a sort of *metadata*, that is, the *knowledge* of the device or paradigm used for the acquisition of data and also the knowledge of the units, instrumental errors, codes and software used, and particular circumstances of the acquisition of such data. (Roederer, 2005, p. 7) So, while the notion of knowledge plays an essential role in the mathematical theory of communication, the theory does not give any account of it. This makes MTC an incomplete theory.

Concerning above, we can reliably exclude MTC from theories dealing with "information" itself. And this is what many scientists and philosophers who are aware of the above problems do. What are the other options? There are two other concepts of information, which seem to be appropriate substitutes: Semantic conceptions of information and pragmatic ones. Such a terminology, syntax, semantics, and pragmatics, traces back to the nineteenth century, when Charles Sanders Peirce defined what he termed "semiotics" as the "quasi-necessary, or formal doctrine of signs" (Peirce, 1932, paragraph 227). It abstracts "what must be the characters of all signs used by [...] an intelligence capable of learning by experience," (ibid) and which is philosophical logic pursued in terms of signs and sign processes. (Peirce, 1902) Charles Morris followed Peirce in using the term "semiotics" and in extending the discipline beyond human communication to animal learning and use of signals. He defined semiotics as grouping the triad syntax, semantics, and pragmatics. (Morris, 1938) Syntax studies the interrelation of the signs, without regard to meaning. Semantics studies the relation

[42] This claim is in line with figure 2.3 in page 42.

between the signs and the objects to which they apply. Pragmatics studies the relation between the sign system and its human (or animal) user.[43]

To get back to our discussion on MTC, my claim is that if one still insists on considering MTC as a theory of information, it should be counted as a syntactic conception of information. However, as emphasized, it can just be a label. In the rest of this chapter, I first briefly delineate the main specifications of semantic theories of information in general[44] and Fred Dretske's semantic theory of information (1981) in particular. There are some reasons for choosing Dretske's ideas as the representative of semantic information, the most important of which are as the following:

1. Dretske's semantic theory of information, especially his information-theoretic account of knowledge, has been the most widely known among philosophers working in the relevant fields. The impact of his semantic theory of information in the realm of philosophy is safely comparable with the profound effect of MTC in this area.

2. In the relevant fields, the ideas Dretske projects have some commonalities (at least terminologically) with what I will claim to be a convincing concept of information physically, namely, pragmatic information as posed by Roederer (2005).

[43] Unlike his mentor George Herbert Mead, Morris was a behaviorist and sympathetic to the Vienna Circle positivism of his colleague Rudolf Carnap. Morris has been accused of misreading Peirce.
[44] Bar-Hillel and Carnap (1953) were the first developers of a genuine semantic theory of information. See also Hintikka's "On Semantic Information" in *Information and inference* (1970).

## 2.4 Semantic Conceptions of Information

Suppose that, through these lines, I acknowledge that what I have written up to this point has all been wrong. The question then would be that whether the reader has gained any information by reading the previous lines. MTC's answer would be positive. For, this theory, as I have repeatedly mentioned, is just interested in data transmission, regardless of the meaning these data may carry or whether they are capable of yielding *knowledge*. Note that when we speak of knowledge, we should concern ourselves with *truth.*

Dretske's semantic theory of information is aware of the problems associated with MTC. Therefore, Dretske's answer to the above question would be negative. According to Dretske, "information is that commodity capable of yielding knowledge, and what information a signal carries is what we can learn from it. If everything I say to you is false, then I have given you no information." (Dretske, 1981, p. 44) Roughly speaking, Dretske believes that information can be seen as *true semantic content* capable of yielding knowledge. Care should be taken that not all philosophers dealing with semantic theories of information agree with such a definition. Some argue that being *true* for the semantic information is not a necessary condition.

Let me start with a less controversial definition of semantic information as *meaningful data*. I would not claim that there is no disagreement among people over such a definition. However, one can say that it is less controversial. As it is apparent in the above definition, the notion of *data* is considered the building blocks of information.[45] Without data, there cannot be any information. According to this view, in the simplest case, information must at least consist of a datum. Furthermore, to be qualified as semantic information, the constituents of information, i.e., data, must have (at least potentially) meaning. "Over the last three decades, several analyses in Information Science, in Information Systems Theory, Methodology, Analysis and Design, in Information (Systems) Management, in Database Design and in Decision Theory have adopted a *General Definition of Information* (GDI) in terms of *data* + *meaning* (see Floridi (2005b) for an extended bibliography). GDI has become an operational standard,

[45] Such a treatment with information is called data-based definition of information.

especially in fields that treat data and information as reified entities (consider, for example, the now common expressions “data mining” and “information management”). Recently, GDI has begun to influence the philosophy of computing and information (Floridi (1999) and Mingers (1997)).” (Floridi, 2005a)

How is the notion of data (or a datum) defined in a data-based definition of information? “A datum is a putative fact regarding some difference or lack of uniformity within some context.” (ibid) For example, any lacks of uniformity between the perception of two physical states (like higher and lower pressure shown by a barometer) or between two symbols (like between C and D in the Latin alphabet) represent data.[46] Such a definition of data as the difference or lack of uniformity within some context naturally leads us to consider data to be a relational entity.

Furthermore, general definition of information supports data to be semantically independent of the informee. If a German cannot read a piece of poem in Persian and does not know what it means, *according to the general definition of information*, it does not necessitate that it contains no information. General definition of information holds the idea of the possibility of information without informed subject. This is still a weak form of the general definition of information. There is a stronger sense according to which not only are data semantically *independent* of the informee, but also of intentional, intelligent producer of the information, that is, informer. Dretske (1981) is one of the proponents of such an idea.[47]

Having all this in mind as some of the specifications of the general definition of information, we are ready to deal with Dretske’s semantic theory of information as one of the most influential theories in the relevant fields. As mentioned before, Dretske is one of the defenders of the truth-based definition of information.[48] But Dretske perambulates a long way to get to this point. Let us have a short review on the way which has led him to this theory.

[46] For a comprehensive discussion on “data” see Floridi (2005a)

[47] In presenting the general specifications of semantic theories of information, I have vastly benefited from Floridi (2005a)

[48] Other defenders of such theories of truth-based definition of semantic information are Grice (1989), Barwise and Seligman (1997), Graham (1999) and Floridi (2004 and 2005b) as opposed to Fetzer (2004), Dodig-Crnkovic (2005), Colburn (2000), and Fox (1983). They do not see truthfulness as a necessary condition for the definition of semantic information.

Dretske begins his book, *Knowledge and the Flow of Information*, with rehearsing some of the elementary ideas of MTC. He proposes examples through which equations (2.2) and (2.3) to (2.9) are obtained. Then, he *adapts* and extends MTC's account of an information source and channel into an account of *how much* information a *particular* signal carries about a source and *what* information it is. What-ness of information has to do with the *content* of a signal or message, with *what-it-is-we-can-learn* from that signal or message. As such, it does not make sense to speak of the average amount of the content of a signal. When through reading my emails I find that one of my papers has been accepted to appear in the journal *Science* and my wife calls me for dinner, it is meaningless to talk about the average *content* of these two messages. One cannot talk of the average of *what* one has *learned*. However, there may be some average to the amount of information through receiving particular messages. One may talk of *how much* on the average one has learned.

Dretske correctly proposes that information as it is ordinarily understood and as it should appear in semantic and cognitive studies has only to do with individual events (signals, structures, conditions). For "it is only the particular signal (utterance, track, print, gesture, sequence of neural discharges) that has a content that can be given propositional expression (the content, message, or information carried by a signal)." He constructs his theory of semantic information on the basis of the adaptation of MTC's formulas, applying them to the particular signals. But doing so makes the theory problematic at the very beginning. He mentions equation (2.11) as the amount of information generated at the source by the realization of a particular state $s_a$ and also equation (2.12) as the amount of information carried by a particular signal $r_a$ about $s_a$ by analogy with equation (2.6).

$$I(s_a) = \log_2 {}^{1}\!/\!{}_{p(s_a)} \qquad (2.11)$$

$I_{s_a}(r_a) = I(s_a) - E(r_a)$, or equally and for more convenience:

$$I(s_a, r_a) = I(s_a) - E(r_a)$$

(2.12)

Where: $E(r_a) = \sum p(s_i / r_a) \log_2 {}^{1}\!/\!{}_{p(s_i / r_a)}$

$E(r_a)$, according to Dretske, is the contribution of $r_a$ to the equivocation $E$. But, as Lombardi (2004) correctly states, some problems arise even at the very beginning of the proposal of such formulas. In equation (2.12), Dretske considers $r_a$ as a particular *signal* about $s_a$. But $r_a$ is not a signal. It is one of the possible states at the receiver. Another problem is associated with the notation Dretske uses for his formalism: he applies the same subindex ($a$) for both the state of the source ($s_a$) and the state of the receiver ($r_a$). It does not hold generally. Lombardi (ibid, p. 111) modifies the notation as the following:

$$I(s_i, r_j) = I(s_i) - E(r_j) \tag{2.13}$$

Where:

$$E(r_j) = \sum p(s_i / r_j) \log_2 \frac{1}{p(s_i / r_j)} \tag{2.14}$$

Summing up the individual equivocations $E(r_j)$ and weighting them according to their respective probabilities of occurrence, we will get equation (2.8), that is:

$$E = \sum p(r_j) \sum p(s_i / r_j) \log_2 1 / p(s_i / r_j)$$ [49]

Dretske's claim here is that the idea of defining the amount of information associated with the particular events and signals is completely consistent with MTC's formulas:

> "It should be emphasized (if only for the benefit of those who will accuse me of *mis*representing, or *mis*understanding, communication theory) that the above formulas are now being assigned a significance, given an interpretation, that they do not have in standard applications of communication theory. They are now being used to define the amount of information associated with particular events and signals. Such an interpretation is foreign to (but, I would urge, perfectly consistent with) orthodox uses of these formulas." (Dretske, 1981, p. 52)

[49] Note that, up to this point, the formulas proposed by Dretske have been modified just formally.

However, adaptation of MTC's formulas to the particular events and signals is not perfectly consistent with MTC's formulas. If it were, we could obtain formula (2.7) by averaging $I(s_i, r_j)$. But we cannot.[50] Such an incompatibility arises form an implicit supposition, namely that the individual equivocation $E(r_j)$ is only a function of $r_j$ at the receiver. But it is also a function of communication channel. "[…] it is not the state $r_j$ which individually contributes to the equivocation *E*, but the pair $(s_i, r_j)$ with its associated probability $p(s_i)$ and $p(r_j)$ and the corresponding conditional probability $p(r_j / s_i)$ of the channel." (Lombardi, 2004, p. 113)

Dretske expresses the theoretical definition of a *signal's informational content* as what follows:

> "A signal *r* carries the information that *s* is *F* = the conditional probability of *s*'s being F, given *r* (and *k*), is 1(but, given *k* alone, less than 1)." (Dretske, 1981, p. 65)

Consider the pinball machine (figure 2.7) with four equiprobable states, connected to the measuring device. The aim of the pointer in our measuring device to the state |00> will carry the information that the ball is in bin 00, if the probability of finding the ball in 00, given the reading on the pointer aiming at |00> is 1.

In the above definition, *k* stands for the prior knowledge of the receiver (informee) about the existing possibilities at the source. Coming back to our pinball machine, suppose that one knows *a priori* (before machine's operation) that it is only possible to find the ball in bin 00 or 01. The occurrence of the message "the ball is not in bin 01" which eliminates the possibility of the ball being found in bin 01 will carry the information that "the ball is in bin 00." For someone who does not know that there are just two possibilities, however, the message "the ball is not in bin 01" does not carry the information that "the ball is in bin 00." The message "the ball is not in bin 01" carries 1 bit of information for someone who already knows about the impossibility of finding the ball in bins 10 and 11. However, for the one without such *a priori* knowledge it just cancels one of the possibilities, carrying 0.41 bit of information.

[50] For the proof, see: Lombardi (2004, pp. 111-112)

The definition of informational content, according to Dretske, results in the following:

> "If the conditional probability of *s*'s being *F* (given *r*) is 1, then the equivocation of this signal must be 0 and [in accordance with formula (1.5)] the signal must carry as much information about *s*, $I_s(r)$, as is generated by *s*'s being *F*, $I(s_F)$." (ibid, p. 65)

The formula he refers to as (1.5) above is: $I_s(r) = I(s) - E$. Suppose that the occurrence of the message "the ball is in bin 00" generates 2 bits of information. Then, according to Dretske, the signal carrying the information that "the ball is in bin 00" ($I_s(r)$) must carry *as much* information about the ball (source) as is generated by the message "the ball is in bin 00 (*s*'s being *F*)," that is, 2 bits. In other words, the transmission of a particular message obeys an all-or-none principle: either all or none of the information associated with the content of a message will be transmitted.

To summarize, Dretske claims that a particular signal $r_B$[51] carries the information $s_A$ (which stands for "*s* is *F*") iff $p(s_A / r_B, k) = 1$ (*k* stands for receiver's prior knowledge of the possibilities). He proposes that the satisfaction of such a criterion guaranties the contribution to the equivocation *E* to be zero and consequently: $I(s_A, r_B) = I(s_A)$

However, this claim is wrong. The criterion $p(s_A / r_B, k) = 1$ does not imply that the individual contribution to the equivocation *E* is equal to zero and $I(s_A, r_B) = I(s_a)$. According to the equation (2.8), the equivocation of the complete channel depends on all the probabilities of the channel, and not only on $s_A$ and $r_B$. In fact, *E* measures the equivocation of the entire channel. Even if the individual contribution $E(s_A, r_B)$ to *E* is zero, because $p(s_A / r_B, k) = 1$, there may be other pairs $(s_i, r_j)$ such that their probabilities $p(s_i, r_j)$ introduced in equation (2.8) make *E* not be zero. In other words, we may have other pairs $(s_i, r_j)$ such that their individual contribution $E(s_i, r_j)$ to *E* is not zero.

[51] Again, I have to mention that $r_B$ is not a signal, but one of the possible states at the receiver. Dretske wrongly calls it signal. However, during explaining Dretske's ideas, I will use this term where he has used it.

With respect to the above discussion, as far as the concept "information" itself is concerned, it seems to me that *not only has Dretske's semantic theory of information nothing to do with his definition of informational content, but also with MTC formulas.* Perhaps one can reliably consider Dretske as one of the addressees of the Shannon's paper "The Bandwagon" mentioned in the previous section.

What about Dretske's ideas on knowledge? In Dretske's view, information shares a feature with cognitive attitudes dependent on it, e.g., beliefs and knowledge. This feature is that "the informational content of a signal is a function of the *nomic* (law-governed) it bears to other conditions. […] The reason my thermometer carries information about the temperature of *my* room (the information that it is 72˚ F. in the room), but not about your room though both rooms are at the same temperature, is that (given its location) the registration of my thermometer is such that it *would not* read 72˚ F. *unless* my room was at this temperature. This isn't true of your room." (Dretske, 2000, pp. 108-109) A signal may carry the information that *s* is *F* without carrying the information that *s* is *G*, even if all *F*s are *G*. So, the informational content of a structure has *intentional* properties. Intentionality in Dretske's vocabulary should be understood the same as what philosophers mean by it: not only does the informational content of a signal or structure depend on the reference of the terms in its sentential expression, but also on their meaning. Thus, substitution of the coreferring expressions in the sentential expression of a structure's informational content will change the content.

Beliefs, desires, hopes, judgments, intentions, love, knowledge, etc., as well as signals' informational content, exhibit intentional properties. This led Dretske to the hope for finding an information-theoretic account of knowledge.

> "Perhaps, that is, one can know that *s* is *F* without knowing that *s* is *G*, despite the fact that all *F*s are *G*, *because* knowledge requires information, and one *can* get the information that *s* is *F* without getting the information that it is *G*. If intentionality is 'the mark of the mental', then we already have, in the physically objective notion of information defined above (even without *k*), the traces of mentality. And we have it in a form that voltmeters, thermometers, and radios have. What distinguishes us from these more pedestrian processors of information

> is not our occupation of intentional states, but the sophisticated way we process, encode, and utilize the information we receive. It is our *degree* of intentionality." (ibid, p. 109)

The above quote contains important claims one has to probe:

1. The hope for finding an information-theoretic account of knowledge stems from the fact that both have intentional property.

2. Dretske considers information to be a physical, objective commodity, independent of the producer and the receiver of it. Information is defined in terms of the network of lawful relationships that hold between distinct events and structures.

3. According to Dretske, as far as information processing and intentionality are concerned, the only thing that distinguishes us from pedestrian processors, like artifacts, is that we are more complex in processing information. I will argue, in detail, that this cannot be the case.[52]

According to Dretske, the traditional answer to the question "what is knowledge?" as justified, true belief should be replaced with an information-theoretic analysis. He characterizes perceptual knowledge as the following:

> "*K* knows that *s* is *F* = *K*'s belief that *s* is *F* is caused (or causally sustained) by the information that *s* is *F*." (Dretske, 1981, p. 86)

By "perceptual knowledge," he means the "knowledge about an item *s* that is picked out or determined by factors other than what *K* happens to know (or believe) about it. That is [...] we are concerned with knowing of something that it is *F* where the something known to *F* is fixed by perceptual (noncognitive) factors."[53] (ibid)

[52] In the later chapters, I will argue that one should distinguish between natural living systems as "originally intentional" systems and artifacts as "derived intentional" systems.

[53] Note that in his account, Dretske distinguishes between perception and recognition: "to describe what object you see is to describe what object you are getting information about; to describe what you recognize it as (see it to be) is to describe what information (about the object) you have succeeded in cognitively processing (e.g., that it is a duck). You can see a duck, get information *about* a duck, without getting, let alone cognitively processing, the information that it is a duck. Try looking at one in dim light at such a distance that you can barely see it. To confuse seeing a duck with recognizing it (either as a duck or as something else) is simply to confuse sentience with sapience." (Dretske, 2000, p. 114)

In the above view, knowledge is an information-caused belief. To know something perceptually, one need not justification, reason, evidence, etc. To know that the glass in front of me is half-full of water I do not need the latter concepts. Suffice it to *see* that it is half-full. In this example, the information that the glass is half-full causes the belief that the glass is half-full. Dretske explains how it is possible.

> "Assuming that belief is some kind of internal state with a content expressible as *s* is *F*, this is said to be caused by the information that *s* is *F*, if and only if those physical properties of the signal by virtue of which it carries this information are the ones which are causally efficacious in the production of the belief." (Dretske, 2000, p. 110)

The light shone to my eyes from the pattern of the half-full glass carries the information that the glass is half-full. Only if *this* light (correspondent to the pattern of half-full glass) causes me to believe that the glass is half-full, the information that the glass is half-full has caused me to believe that the glass is half-full.

We saw that although Dretske claims that his semantic theory of information has its roots in MTC, the formulas he projects for individual messages are not consistent with Shannon's theory of communication. So, one would not know what Dretske has in mind when he speaks of information-theoretic account of other mental, intentional states such as knowledge and belief. It seems that his notion of information refers to nothing but just conditional probabilities. I doubt that one can call his account of other mental, intentional states as information-theoretic.

Consider Drestske's account of knowledge. Foley (1987) replaces the reference to information in such an account with Dretske's probabilistic account of informational content[54]. We will get the following:

> "*S* knows that *a* is *F* = An event *e* causes *S* to believe that *a* is *F* and is such that the probability of *a*'s being F, given *e* and the rest of what *S*

[54] As I showed, Dretske's semantic theory of information has also nothing to do with his account of signal's informational content. Here I wanted to say that even if it had, Dretske's account of knowledge could not be counted as information-theoretic.

> knows, =1, where the probability of *a*'s being *F*, given the rest of what *S* knows (without e), is not 1." (Foley, 1987, p. 165)

Can one see any traces of an information-theoretic account of knowledge in the above definition? Where is the reference to the inverse probabilities (equation 2.11), bits of information or other individualized formulas of MTC? Knowledge here is explained in terms of causality, belief, and conditional probability.

To clarify the above point, Dretske's proposal on the relation between the amount of information a signal carries and the informational content will help.

> "What communication theory tells us is that the amount of information a signal carries about a source sets an upper bound on what information a signal can carry about the source." (Dretske, 1981, p. 62)

The above claim leads to the following proposal: *S* must have received sufficient amount of information to know that *a* is *F*. We have two ways of interpreting such a proposal. First, to think of information and amount of it as what Dretske claims they are and are based on (that is, Shannon's communication theory). But we saw that his account of information is not consistent with such a theory. Another way of interpreting the above claim is to think of information as ordinarily understood, that is, without any reference to a special mathematical theory. In this case, there is nothing special and distinctive about Dretske's semantic theory of information to make it superior to other accounts of knowledge. For "almost any account of knowledge can be interpreted as endorsing the idea that *S* knows that *a* is *F* only if he has enough information about *a* to know this (Foley, 1987, p. 166)," even justified true belief account of knowledge.

To sum up, Dretske's semantic theory of information has both advantages and disadvantages. The advantages include recognition of (at least terminologically) "information" as a relational, intentional, and objective commodity. The disadvantages are as follows:

1. Dretske's semantic theory of information dangles from MTC's formulas, while it has no consistency with them.

2. In Dretske's account of intentionality, as a property of information, no clear-cut distinction is made between living and nonliving systems. This leads to panpsychism in the following way: If intentionality is the mark of mental, and if the mere difference between us and more pedestrian processors of information is just in the degree of intentionality, then as we move to less complex systems, entities just get less intentional, no matter they are living or non-living systems. He sees no reason to consider intentionality winking out altogether as one moves down the scale of complexity. In the later chapters, I will argue against such an idea. I am not inclined to call pansychism as a problem, but it sounds to me, like many others, very counterintuitive. My claim is that if a theory of information can successfully give us a clear-cut distinction between living and nonliving systems, then such a theory would be a better choice to adopt.

3. The worst problem associated with all semantic theories of information, in general, and Dretske's semantic theory of information, in particular, is that meaningful data, as semantic information, is out there independently of the informee. Such a thesis is dead from the very beginning. Recall Gernert's example stated in chapter 1, page 32, quoted again in the following:

> "[O]ur letter "C" will be understood as the letter "S" by anyone using Cyrillic scripture. This dependence of the meaning on the receiver's prior disposition can be disregarded only if there is a "tacit agreement", fixing a certain set of possible receivers. Hence, any concept of semantic information will be included in the concept of pragmatic information as the special case in which a class of possible receiving systems has been fixed, such that their uniform way of interpreting the signals is constant and predictable." (Gernert, 2006, p. 147)

The above example simply shows that, generally, meaning is dependent on the receiver or the receiving system and semantic conceptions of information (and also syntactic conceptions) are just special cases of the concept of pragmatic information in which the interpreter of the messages plays an essential role in information processing.

As stated earlier, Dretske goes even further and considers semantic information to be independent of the producer or informer of it. In the next section, I will discuss this in very detail showing that such a claim cannot be satisfactory. To summarize, the claim of this essay, contrary to Dretske, is that information comes into play when both the *purpose* of the informer and the *meaning* such a purpose has for the informee are concerned.

So, the idea advocated in this essay is that, in general, the concept of pragmatic information fulfils the requirements of a universal theory of information as it is concerned both with the meaning and with the interpreter of the received signals or messages. As seen in the next section, among different interpretations of the notion of pragmatic information, I have chosen and adopted Roederer's reading of the notion because of the reasons stated in page 34.

## 2.5 Pragmatic Information

As stated before, the claim of this essay is that the notion of pragmatic information represents the most universal and comprehensive concept of information of which other concepts are just special cases. In this essay, Roederer's (2005) reading of the notion of pragmatic information has also been chosen and adopted as the most appropriate version of pragmatic information to apply to the "double aspect theory of information," the idea that information has two aspects, namely, physical and phenomenal.

This section contains a summary and a restatement of some of Roederer's ideas carefully selected in favor of the ideas I will propose. An important point is that choosing Roederer's pragmatic information as a suitable concept of information for double-aspect theory does not imply that I agree with all ideas he puts forward in his works. In using his theory in later chapters, I will contribute two different kinds of modifications: minor and major. The minor modification would be applied to the terminology used in his thesis in order to provide a suitable context for the major modification: considering pragmatic information to have two aspects, physical and phenomenal.[55]

Let me start with Roederer's account of the term "interaction" as an underlying basic, primordial concept in pragmatic information. Without attempting to define it formally, we can say that when two bodies interact, there happens a change in any of their properties such as shape, motion, constitution, and so on. Roederer identifies "two different classes of interactions between the bodies that make up the universe as we know

[55] I already know that he would not be happy with such a usage of his theory. He frequently states that he is a physicalist and does not want to deal with philosophical, dualistic accounts of mind: "Therefore I will take a reductionist, physicalist, "Copenhaguenist," biodeterministic and linguistic deterministic- stand. As a practicing space plasma physicist I am no expert in philosophy – which means that I have (thank heavens) no preconceived opinions in related matters. Philosophers have been asking questions during millennia about Nature, human beings and their place in it. Today, however, we must recognize that answers can only be found by following in the strictest way all tenets of the scientific method – logical analysis alone cannot lead to a quantitative understanding of the Universe. Surely, philosophers should continue pressing ahead with poignant questions; this will serve as a powerful stimulant for us "hard" scientists in the pursuit of answers!"(Roederer, 2005, p. 5) Here, I have to mention some points. First, all these "-ist" or "-istic" attributes mark philosophical points of view, not scientific. Second, the task of contemporary philosopher has not been to give just logical analysis or to produce preconceived opinions, although I think Roederer himself already has some of those opinions. Recent philosophy has been intertwined with science such that most of the time it is not possible to determine the border between them. Finally, in contrast to what Roederer wished, in the recent literature, his works on pragmatic information are counted as philosophical rather than scientific.

it, with the concepts of information and information processing as the key discriminators between the two" (Roederer, 2003, p. 5): Physical Interactions[56] and Information-Driven Interactions.[57]

### 2.5.1 Physical Interactions

Both in quantum domain and in macroscopic world, when two bodies approach each other, they may interact. Under some *initial conditions,* such interactions may lead to stable, bound structures, while under some other initial circumstances they may lead to no new structures or unstable ones, which decay after a time.

When two bodies 1 and 2 in the classical non-relativistic realm,[58] isolated from the rest of the Universe[59], interact, Newtonian mechanics tells us that there are two scalars $m_1$ and $m_2$ such that:

$$m_1 \vec{a}_1 + m_2 \vec{a}_2 = 0 \tag{2.15}$$

$\vec{a}_1$ and $\vec{a}_2$ are instantaneous acceleration vectors. Each term in the above equation is called *force*: $\vec{f} = m\vec{a}$ .

Hamiltonian mechanics is a re-formulation of classical mechanics whose entire framework can be derived from (2.15) together with the fact that the acceleration of a body which is simultaneously subjected to different interaction mechanisms is

[56] Depending on the context, Roederer uses other expressions for "physical interactions," such as "force-field driven," "force-driven," "classical interactions" (in the macroscopic domain), and "quantum interactions" (in quantum domain). To avoid confusion, I only use the term "physical interactions" for this sort of relationships.

[57] Note that the relationship between bodies in an *interaction* must be bidirectional. As we shall see, not all relationships called "physical interactions" by Roederer are thorough interactions. Some of them are not bidirectional and there *is* a privileged direction of time in their relationships. They are indeed actions. This also holds for "information-driven" relationships in which direction is *always* from sender to receiver. Such relartionships are unidirectional. Roederer is aware of this. However, according to him, as in the so-called information-driven interactions there must be a previous relationship or "understanding" between the sender and recipient, they can be called interaction. Moreover, in quantum systems, we are confronted with the "feedback" perturbation of the system being measured, which makes the system bidirectional. (Roederer, 2005, p. 117)

[58] In order to consider the bodies to be in such a realm, the following conditions must be fulfilled: 1. the size of bodies must be negligible compared to their mutual distance; 2. their rotational energy must also be negligible with respect to translational energy, and 3. The velocities of the bodies must be negligible with respect to that of light.

[59] Strictly isolated systems (closed systems) do not exist since the gravitational field, as a geometric property of space-time, cannot be canceled in any given region of space.

proportional to the vector sum of all forces acting on it. Hamiltonian mechanics deals with a system of interacting mass points with geometric limitations on their motion imposed by given constraints. For a given classical closed system consisting of $N$ mutually interacting material points, classical Hamiltonian mechanics allows us to determine the system's coordinates $q_k$ and momenta $p_k$ at any time $t$ once we have known the initial conditions, i.e., the coordinates $q_k^0$ of the bodies and the associated momenta $p_k^0$ at the initial time $t^0$. Hamiltonian mechanics provides a correspondence between initial state of the system at initial time $t^0$ and final state at time $t$. Such a system would be perfectly reversible as far as the interaction force is conserved. We can run the system backwards (for a time $t$) reaching the initial states again, this time with reverse velocities $-p_k^0$. This reversibility states that there is no preferred direction of the time. Thus, the relationships between the points are *inter*-actions not cause-and-effect. A very crucial point here is that it is *we* humans who set the initial conditions of the given mass points and also the time $t$ of the final state. However, the conditions of the system at the final state are out of our choice. I will come back to this point in the next chapter.

As far as the concept of force field in the classical, macroscopic, nonrelativistic domain is concerned, we have to turn our attention to gravitational, electromagnetic interactions. Gravitational interactions are available between all material bodies. Electromagnetic interactions act between some types of elementary particles as well as charged or magnetized bodies. Note that frictional, chemical and thermal interactions in the macroscopic domain are also collective interactions based on electromagnetic interactions at atomic and molecular level. Finally, there are also elastic interactions. They occur when a medium acts between the interacting bodies or a special physical device intervenes the interaction with the condition that its mass is negligible or can be considered an integral part of the system.

If all interactions mentioned above occur between bodies isolated from the rest of the Universe, and if the force acting on the mutual interacting bodies is conserved, then they will remain *inter*-actions. "Such interactions are bidirectional in the sense that neither of the two interacting bodies has a hierarchical ranking over the other. There is no "cause-and-effect" relationship as long as there is no external interference: no irreversibility or asymmetry, and no privileged direction of time." (Roederer, 2005, p. 98)

Let us turn our attention to the quantum domain. At the microscopic, subatomic level, everything is reducible to the four fundamental interactions: gravitation, electromagnetism, strong, and weak interactions.

We are already familiar with the gravitational and electromagnetic interactions as they are present in our everyday life in the macroscopic world. Gravitation force is the weakest of the four fundamental forces, yet it is the dominant force in the universe for shaping the large-scale structure of galaxies, stars, etc.

Fundamental electromagnetic interactions occur between any two particles that have electric charge. These interactions involve the exchange or production of photons. Thus, photons are the carrier particles of electromagnetic interactions. Electromagnetic interactions are responsible for the binding force that causes negatively charged electrons to combine with positively charged nuclei to form atoms.

As stated earlier, in classical domain, for a closed system consisting of interacting particles in which force is conserved, gravitational and electromagnetic interactions are reversible and bidirectional and there is no privileged direction of time. This holds for quantum domain. Strong interactions in the quantum world are bidirectional. They are understood to represent the interactions between quarks and gluons. The strong force is mediated by gluons, acting upon quarks, anti-quarks, and the gluons themselves. Fundamental weak interactions occur for all fundamental particles except gluons and photons. Weak interactions involve the exchange or production of W or Z bosons. Weak interaction was first recognized in cataloging the types of nuclear radioactive decay chains, as alpha, beta, and gamma decays. Alpha and gamma decays can be understood in terms of other known interactions (residual strong and electromagnetic, respectively), but to explain beta decay, it required the introduction of an additional rare type of interaction, that is, the weak interaction. Quark flavor never changes *except* by weak interactions, like beta decay, that involve W bosons. Weak interactions are the only fundamental ones, which are not reversible.

Let me summarize the idea projected in this section. A classical closed system consisting of interacting mass points in which the force is conserved is bidirectional, symmetric, and reversible. There is no cause-and-effect relationship between the interacting bodies as far as there is no external interference. However, if a system of interacting pointes is under the effect of some external forces, it will no longer be reversible. Reversing all the velocities at a time $t$ does not lead to the same initial state at time $t_0$ with reversed initial velocities. For such a system to be reversible, all external interactions must follow exactly the same pattern in reverse. To fulfill such a condition, we have to have total information on and control of the external part. So, any, even very weak, perturbation would make a limited classical system of interacting bodies irreversible. (ibid, p. 173) I will come back to this point in the next chapter.

In the quantum domain, all kinds of interactions, but weak interaction, are reversible and bidirectional.

All above-mentioned physical interactions, occurring in nature, have some fundamental characteristics, i.e.:

> "The end effect of a physical interaction will always depend on some "initial conditions," such as the initial configuration (positions, velocities) of the interacting bodies. A most fundamental characteristic is the fact that during the interaction, there is a direct transfer of energy from one body to the other, or to and from the *interaction mechanism* itself (in case of fundamental interactions force field; in more complex cases, some process linking the two bodies)." (Roederer, 2003, p. 6).

A clearer understanding of physical interactions will be possible after the other main class of interactions, i.e., information-driven interactions are dealt with. So, I postpone going through the details of the former until the reader has been familiar with the latter.

### 2.5.2 Information-Driven Interactions

"Consider a ship on automatic pilot guided by radar waves. The ship is not pushed and pulled mechanically by these waves. Rather, the form of the waves is picked up, and with the aid of the whole system, this gives a corresponding shape and form to the movement of the ship under its own power. Similarly, the form of radio waves as broadcast from a station can carry the form of music or speech. The energy of the sound that we hear comes from the relatively unformed energy in the power plug, but its form comes from the activity of the form of the radio wave; a similar process occurs with a computer which is guiding machinery. The 'information' is in the program, but its activity gives shape and form to the movement of the machinery. Likewise, in a living cell, current theories say that the form of the DNA molecule acts to give shape and form to the synthesis of proteins (by being transferred to molecules of RNA)." (Bohm, 1990)

These examples of information-driven interactions, given by David Bohm, clearly show the specifications of such interactions.[60] The first and the second examples express the realization of information-driven interactions in artifacts and the third in biological systems. The reason I have quoted Bohm's examples at the very beginning of this section is that in chapters five and six, Bohm's ideas will play a crucial role in my thesis.

Consider Bohm's second example, that is, the radio waves as broadcast from a station and received by our radios. The electromagnetic waves, as *patterns*, are sent by a transmission antenna, as the *sender*, with the *purpose* of eliciting a change in the recipient, namely, our radios, as the *receiver* and, ultimately, in our cognitive state. The radio will receive these specific waves only if it is tuned correctly with respect to the specific wavelength or frequency of the waves. This means that there must be a *one-to-one correspondence* between the sender and the receiver: only a specific wavelength will trigger a specific change in the radio. To guarantee such a one-to-one correspondence once the sender is present, there must be a *common code* between the sender and receiver. Common code is a kind of memory operating between sender and recipient assuring a univocal change to occur in the recipient in presence of a specific pattern. In other words, to ensure a specific change in the recipient, the presence of a specific pattern must have a *meaning* for the recipient. Common code is responsible for such a univocal correspondence. If common code is changed, we will no longer have the same

[60] Bohm does not use the term "pragmatic information," but "active information." However, care should be taken that in Bohm's interpretation of "active information," information is out there, regardless of being processed in natural non-living systems, living systems, or artifacts. In this respect, I will be a critic of his ideas on "active information." My claim, inspired by Roederer, is that natural non-living systems are not information processing systems. Information plays no role in the interactions occurred in the natural inanimate objects.

correspondence and the purpose of the sender will have no meaning or another meaning for the recipient. The last and the most important point is that the energy of the radio as the receiver must be provided locally, i.e., by power plug. It is not supplied by the energy of the electromagnetic waves themselves. In other words, the sender and the receiver are *decoupled energywise*. Note that the correspondence between sender and receiver has not been established by chance. It is *repeatable*.

The terms written in italic in the above paragraph represent the fundamental elements of information-driven interactions occurring in pragmatic information. Concerning above, the very first question for a curious reader would be: what really is the difference between information-driven interactions and physical interactions? At first glance, one can see some anthropomorphic new terms such as “purpose” and “meaning.” So another important question would be: where do these terms come from? Is it us who attribute these terms? Isn’t it possible to use the same vocabulary for physical interactions? In what follows, I will try to answer these questions partly. However, a full understanding of the idea requires patience until the reader comes to the end of the next chapter. For the start, Roederer’s (2005, pp. 111-112) example as explained as follows will help.

Consider some billiard balls on a frictionless table as shown in Fig. 2.16. Three layers are distinguishable: balls in input layer, output layer, and a few intermediate balls. At initial time $t_0$, a given distribution of velocities at the input layer triggers a chain of collisions in the intermediate layer, which in turn impart velocities to the output layer. All collisions are *mutually* elastic and thus reversible. In such a case, Hamiltonian mechanics enables us to calculate the state of the system at any other time $t$. So, the final velocities of the balls at any time $t$ in the output layer will be computable. Now we increase the number of the intermediate balls making the interaction mechanism occurring in the intermediate level more *complex*. This brings about two important consequences: the system becomes more *irreversible*[61] and the input-output relationship gets more *decoupled energywise*. By

[61] This phenomenon is called the paradox of reversibility. The roots trace back to a similar problem in thermodynamics. Consider a tank of gas. It is full of gas molecules, the motion of each one describing a trajectory in both physical space, and in phase space. Each particle moves in accordance with the laws of Newtonian mechanics, which are fully reversible in time; a "time-backward" trajectory cannot be distinguished from a "time-forward" trajectory, in the absence of some outside context used to define "backward" and "forward.” However, when those particles are all gathered together in the ensemble of a gas tank, their combined motion is not time reversible! We know this because the second law of thermodynamics enforces it. Such a system mentioned above is quite similar to our example, i.e., balls on

adding more balls to the intermediate level, the energy transferred to the output layer becomes more dependent on the interaction mechanism than input layer.

Let us further assume that we put the intermediate balls in a black box such that we have no access to the balls comprising the interaction mechanism. Now, the entire process can be seen as system in which the input is coupled to the output by an unknown complex interaction mechanism.

Up to this point, everything was purely physical. However, the wish to construct a univocal relationship between input and output layer every time the system operates requires an extra device to reset the system exactly to the same initial position it was before the operation. In other words, a system with the *purpose* of establishing a *one-to-one correspondence* between input and output layer through a complex interaction must be reset every time it operates. There are three ways of fulfilling such a task:

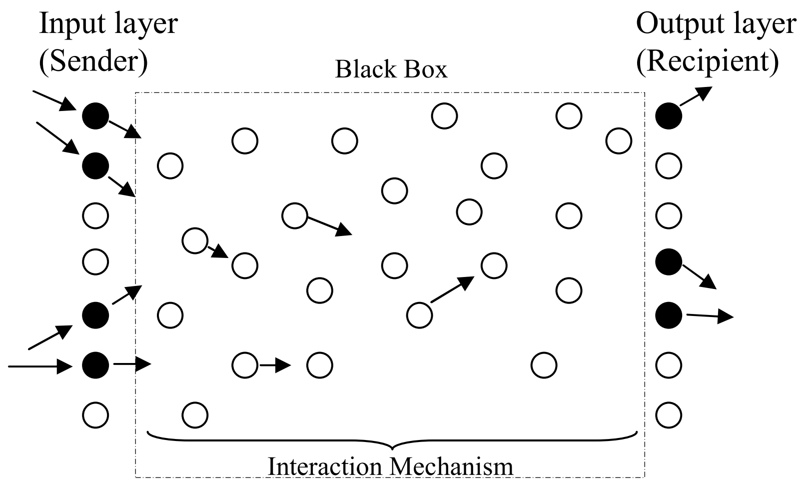


**Figure 2.16.** Billiard balls on a frictionless table. See text for details.

1. Purposefully designing a resetting device with the ability of recording all the initial conditions in its memory and adding the device to the system to reset itself once it is activated;

---

the frictionless table. Notice that this way of treating thermodynamics is consistent with Boltzmann's account according to which thermodynamics is reducible to dynamics. I have discussed the notion of irreversibility in chapter 3.

2. the intervention of a purposeful, intelligent biological system, which has recorded all the initial conditions in mind, and

3. by chance in any natural system. Here, any purpose will disappear. As we shall see, being reset in this way excludes the system to be informational.

Among the three ways of establishing an input-output univocal correspondence mentioned above, 1 and 2 do not need any further explanation, but 3 seems to be a bit ambiguous. How would Nature do such a task just due to stochastic processes? "Given a large number of the identical copies of the system with the same initial setup of the balls in the black box and some process which introduces small random variations ("errors") in the reset mechanism, only those systems which evolve in the right directions, i.e., whose end states come closer and closer to the wanted values, would be able to survive while those less fit would disappear." (Roederer, 2005, p. 113)

Let me summarize the fundamental characteristics of a system depicted in Fig. 2.16, which also benefit from a reset system. The following properties specify *information-driven interactions* versus physical interactions.

1. There exists a specific *pattern* of velocity distribution at the input layer. This specific pattern makes a *specific change* at the output layer.

2. The system has a *complex interaction mechanism*, which leads to the irreversibility and unidirectionality of it. Physical interactions can also be complex. Thus, mere complexity does not characterize an interaction to be information-driven.[62]

3. To achieve a certain goal by every run, that is, to obtain purposefully (not just by chance) the same result every time the system is operated, the system is attributed a *purpose*. In physical interactions, no purpose can be identified.[63]

[62] Here, complexity has simply been characterized by the number of participating objects in an interaction mechanism. In the next chapter, I will delineate more on the concept of complexity.
[63] In chapter 4, I will distinguish between relative and observer-dependent intentionality. At most, physical interactions are purposeful in an observer-relative way.

4. There is a *one-to-one correspondence* between the input and output layers. In other words, there is a univocal relationship between the distribution of velocities in the input layer and the velocity of out-flying balls in the output layer. This correspondence is guaranteed by the role "purpose" plays in such interactions.

5. The input and output layers are *decoupled energywise*. There may be many input constellations totally equivalent energywise, yet triggering different outgoing distributions. This shows that information-driven interactions are dependent on the initial conditions very little, while in physical interactions the end effect of the system always depends on the initial conditions.

Of course a system as illustrated in Fig. 2.16 does not exist in nature. However, as mentioned earlier, there are similar natural mechanisms. "Natural mechanisms responsible for this class of interactions do not arise spontaneously: They must *evolve* – in fact, Darwinian evolution itself embodies a gradual, species-specific information extraction from the environment. This is why *natural* information-driven interactions are all *biological* interactions." (Roederer, 2005, p. 116) While physical interactions happen in natural inanimate systems, information-driven interactions occur in living systems.

Concerning above, a system like Fig. 2.16 can be a good metaphor for explaining two important cases: neural networks and genome. In the former, the role of initial positions is played by synaptic connections and in the latter, the position of the bases in the DNA have the role of initial positions. (ibid, p.113) But recall that there were two ways of achieving purposefully the goal of resetting the system of the billiard balls on a frictionless table, the goal that makes the system information-driven: resetting the system ourselves and designing a device to fulfill the task. The latter shows that artifacts[64] can also be information-driven systems (recall Bohm's example of the radio and radio station).

Having all this in mind, we are ready to define information-driven (or information-based) interactions formally:

[64] Note that when I speak of artifacts in this essay, I am speaking of artifacts which can in a way fulfill the conditions of being information-driven systems.

> "System A is in information-based interaction with system B if the configuration of A, or, more precisely, the presence of a certain spatial or temporal pattern in system A (the sender of source) causes a specific alteration in the structure or the dynamic of system B (the recipient), whose final state depends *only* on whether that particular pattern was present in A. The interaction mechanism responsible for the intervening dynamic physical processes may be an integral part of B, and/or a part of A, or separate from either. Furthermore: 1. both A and B are *decoupled energywise* (meaning that the energy needed to effect the changes in the system B must come from some external source); 2. *no lasting* changes occur as a result of this interaction in system A (which thus plays a catalytic role in the interaction process – unless the pattern is erased by a separate mechanism, or a temporal pattern is obliterated or disappears); and 3. the interaction process must be able to occur *repeatedly* in consistent manner (one-time events do not qualify) – the repetition, however, requires that the mechanism and the changes in the recipient B reset to the same initial state." (ibid, p. 117)

Complexity of the interaction mechanism, unidirectionality of the whole system, and the establishment of a one-to-one correspondence between the spatial or temporal feature (or pattern) in system A and a specific alteration in system B are other main features of such interactions. Concerning the latter, three concepts should be defined, namely, purpose, meaning and common code. Let us start with the concept of purpose: In an open system comprising pattern A and system B, pattern A has a "purpose" where its presence triggers a change in system B so that such a change would not happen if pattern A were not present, or would happen just by chance. In such an interaction, energy must be provided locally. If the correspondence between the sender and the recipient has already been established, that is, if every time pattern A is present, the same change happens in system B in a repeatable manner, we say the purpose of pattern A has a "meaning" for system B. This requires the interaction mechanism to contain some memory device. This memory device, which guaranties the purpose of the sender to have the same meaning for the recipient in every run, is called *common code*. Changing the common code will result in the purpose of the sender to have no meaning or another meaning for the recipient.

What is information itself according to this view? "Information is the agent that embodies the above-described correspondence: It is what links the particular features of the pattern in a source system or sender A with the specific changes caused in the structure of the recipient B." (ibid, p. 118) The above-described features of an information-driven interaction can reliably be illustrated in Figure 2.17.

Notice that the notions such as "purpose" and "meaning" used in pragmatic information are *perspective notions*. "Perspective notions are terms which – beyond the well-known context-dependence of word meanings in general – require an *explicit* statement of the context." (Gernert, 2006, p. 156) Consider the term "categorization." Knowing the meaning of the term *per se* does not suffice to fulfill the task of categorization. The rules or criteria according to which the task is supposed to be done should have already been mentioned explicitly. In order to categorize things, one should already be aware of the purpose of the task. Things can be categorized based on many different properties they have such as shape, size, color, etc. Terms like "purpose" and "meaning" used in information-driven interactions are also perspective notions. They depend on the involved sender and recipient in the interaction. The presence of a pattern, as the sender, has meaning just for the correspondent recipient in a specific context.

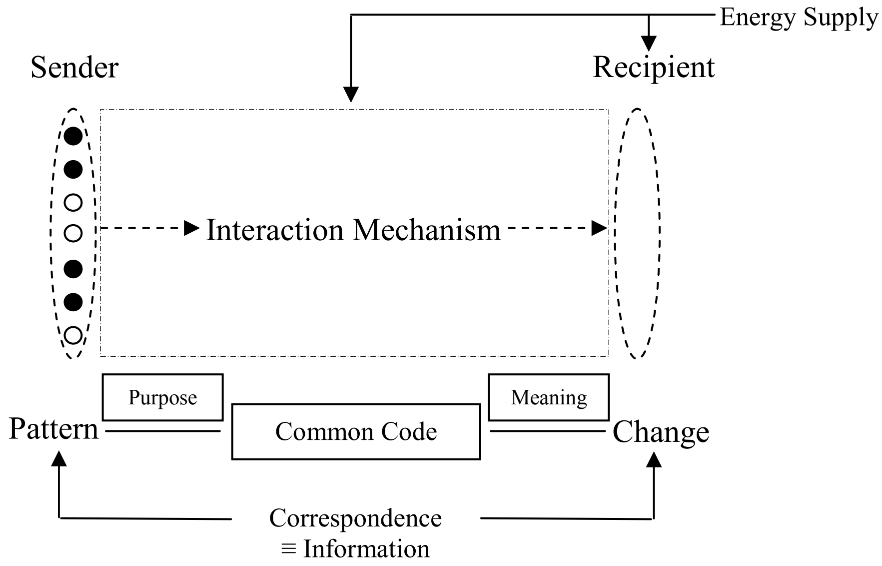


**Figure 2.17.** Elements of an information-driven interaction and the way such elements are connected to each other. See text for more details.

Two expressive examples will show the major differences between information-driven and physical interactions. Consider a moon of a planet orbiting around it. Given the initial conditions at time $t_0$ together with Newton's law of universal gravitation, according to which the force of gravitational interaction $f$ is a function of positions and masses of the mutually interacting bodies, one can know of the position of the moon at any later time $t$. Thus, the specifications of such an interaction can be summarized as the following: 1. for each given initial condition of position and velocity at the initial time $t_0$, there exists a well-defined orbit for the motion of the moon. The position and velocity of the moon at any later time is totally dependent on the initial conditions; and 2. the central body and the moon are coupled energywise. There is mutual energy give-and-take between the gravitational field of the planet and the moon spinning around it. Roughly, such a system consisting of the two interacting bodies can be considered a classical closed reversible system. In such a case, no "purpose" can be identified.

Consider now the motion of a butterfly in orbit around a lit candle. Again, there exists a well-defined path for the motion of the butterfly, this time through a very complex interaction mechanism. The butterfly's motions of the wings determine the path. In this case, contrary to the case of the interaction between the planet and the moon, the energy needed for the movement of the wings comes from the relatively unformed energy provided by butterfly's metabolism, not from the electromagnetic energy emitted by the candle. However, the pattern of the electromagnetic waves emitted by candle plays its role as a kind of *pilot wave* resulting in regulating the force of the wings of the butterfly. This is the pattern of the light emitted by the candle, not its energy, which is considered a controlling factor. The purpose in this case does not lie in the lit candle as the original pattern, but in the mechanism of the perception in the butterfly itself. However, if one still insists on the role of the candle as the sender, it can be said that in this case, both purpose and meaning lie at the recipient's side. In other words, "purpose is not given by the symbols in the input […], rather it is given by the physiological and neural mechanisms that are seeking out certain input patterns […]. This is a prototype of the interactions that involve *information extraction*: it is a fundamental aspect of the interaction of any organism with the environment." (Roederer, 2005, p. 115) According to Roederer, the most fundamental purpose of any measurement process in physics is information extraction. (ibid, p. 116)

Another important difference between the latter and the case of planet-moon system is that the configuration of the butterfly's orbits around the candle depends very little on the injection point. Knowing the initial conditions at the starting point of the butterfly's motion at time $t_0$ would not help in knowing its position and velocity at any later time.

As my last example of information-driven interactions, I would like to describe very briefly (and in a nontechnical manner) how the brain regulates the heart based on information-driven interactions. In doing so, the frontal cortical control over cardiovascular functioning is explored based upon the model presented by Skinner and extended by Greetje (1997). The interested reader can find some details of such a mechanism in appendix 1. Figure 2.18 schematically illustrates such an interaction between the brain and the heart. The claim here is that not only is the interaction between the brain and the heart, as two wholes, information-driven, but also between every two individual parts, which interact locally. Each two parts, which have been connected by an arrow, interact informationally. In each interaction, the sender is the one from which the arrow starts and the receiver is the one to which it ends. In what follows, as a sample, I will analyze one of these interactions, that is, the interaction between the nucleus of the solitary tract (NTS) in the brainstem and nucleus ambiguus (vagal nuclei – NA) that regulates the parasympathetic nervous system during conditions of elevated blood pressure. The context of this mechanism is the following:

The heart and circulation give output to the lateral part of the NTS. The output of the NTS projects to the nucleus ambiguus, and also through the caudal ventrolateral medulla (CVLM) to the rostral ventrolateral medulla (RVLM). The nucleus ambiguus (NA) receives input from various other areas as well. The component in which we are interested here is the cardioinhibitory part of the nucleus ambiguus (located ventrolaterally), which receives input primarily from the medial part of the NTS. The cardioinhibitory part of the NA has cholinergic (vagal) projections to the heart. The ventral external formation (longitudinal division of the NA) encompasses parasympathetic preganglionic neurons. According to Taylor et al. (1999), except for in the external division ventrolaterally to the principal column, vagal-projecting neurons are also situated in the dorsal NA. The parasympathetic nervous system (the vagal projections) is activated with the purpose of depressing heart rate and contractility, with the final purpose of reducing blood pressure.

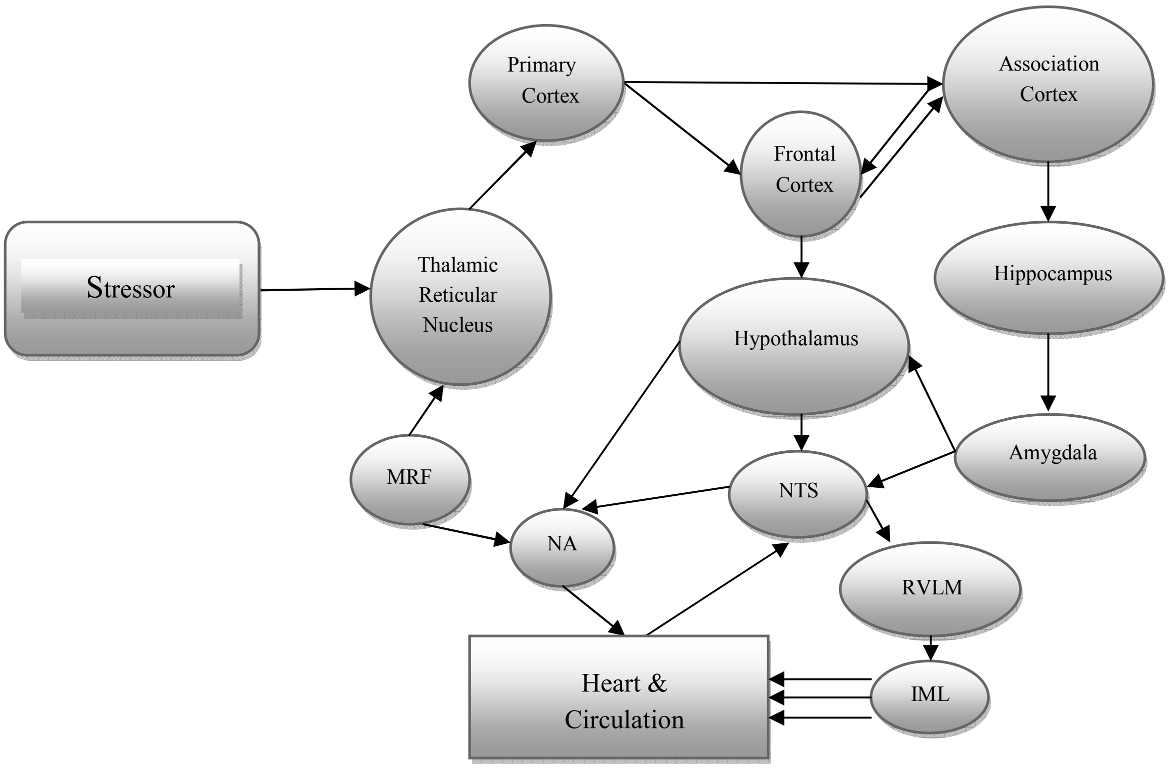


**Figure 2.18.** Schematic illustration of the frontal cortical control over cardiovascular functioning. MRF stands for "mesencephalic reticular formation," NA stands for "nucleus ambiguus," NTS stands for "nucleus of the solitary tract," RVLM stands for rostral ventrolateral medulla, and IML stands for intermediolateral cell column (Greetje, 1997, p. 8). See appendix 1 for further details.

Evidence has been put forward for polysynaptic convergence and interactions of afferent inputs on postsynaptic NTS neurons. It is still a matter of debate how considerable these interactions are.

Having this in mind, we can draw a diagram that shows the major interactions between the NTS and the NA in terms of information-driven interactions.

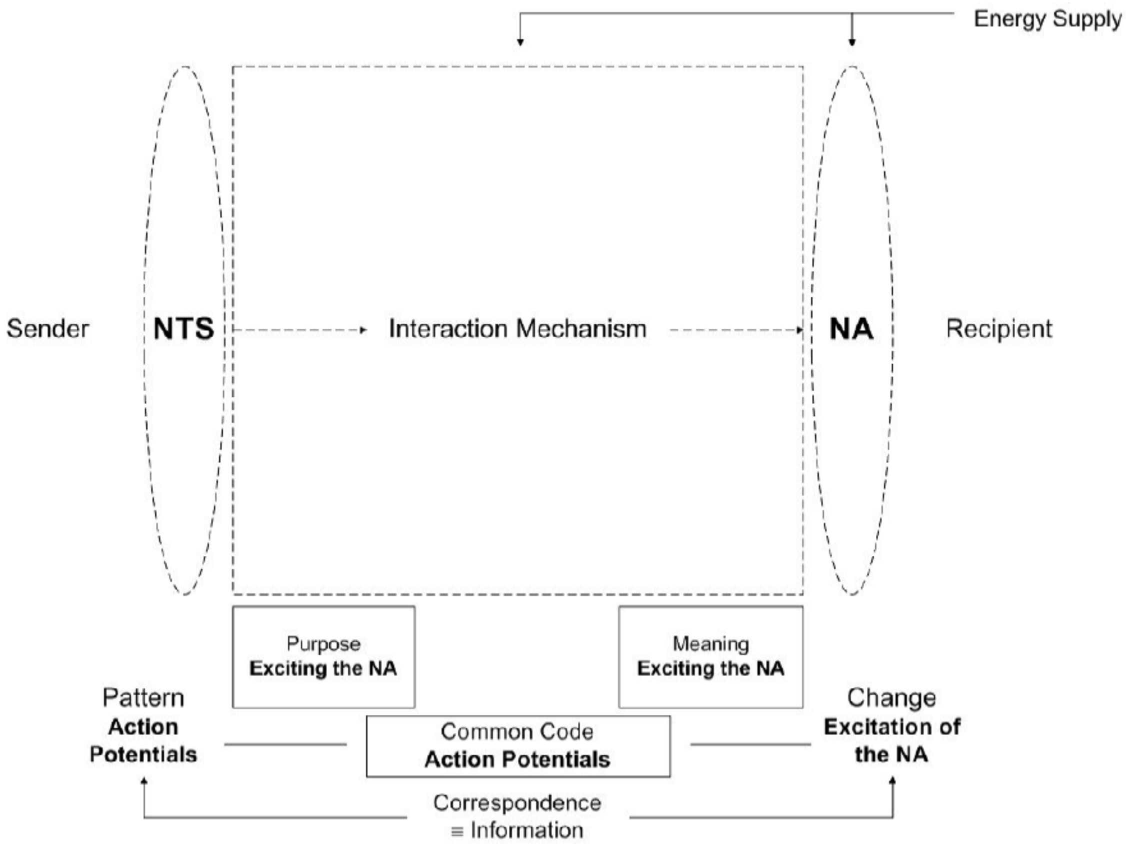


**Figure 2.19.** The interaction between the nucleus of the solitary tract (NTS) in the brainstem and nucleus ambiguus (vagal nuclei – NA) according to pragmatic information

I hope my introduction to Roederer's reading of pragmatic information has shed light to the main differences between physical and information-driven interactions. In the upcoming chapter, the elements of pragmatic information and the application of the notion will be investigated in more details. In the next chapter, we will see that each of these two different classes of interactions belongs to a certain domain of the universe, that is, to the different domains of living and non-living systems.

# Chapter 3
# Pragmatic Information: Elements and Realms

## 3.1 Introduction

In chapter two, it was claimed that the concept of pragmatic information fulfils the requirements of a universal theory of information. Moreover, as a unifying, comprehensive framework, the concept of pragmatic information includes two earlier notions, namely, syntactic and semantic concepts of information.

The present chapter discusses some of the important elements of pragmatic information, as interpreted by Roederer, in more detail. Information-driven interactions occur in complex systems. So, the notion of complexity together with some related notions such as "self-organization" and "emergence" are necessary to be discussed. The claim is that complexity and self-organization can happen as the result of both information-driven and physical interactions, but Information is just represented in self-organizing systems, which enjoy information-driven interactions between their constituent parts, interactions that only happen in the realm of biology.

As one of the main features of self-organizing systems, the notion of "emergence" has also been examined in this chapter. Varieties of emergentism have been generally categorized as "weak" and "strong." It is claimed that every phenomenon is weakly emergent but consciousness.

There is also another notion, which plays an essential role in the account of pragmatic information as far as the purpose of this essay is concerned: the notion of initial conditions. Whether an interaction is physical or information-driven depends on how much the final state in an interaction is dependent on the initial conditions. The notion will be discussed in section 3.3.

## 3.2 Complexity

In the last chapter, we saw that information-driven interactions occur between systems the complexity of which exceeds a certain minimum degree. There, the complexity of a system was characterized by the number of the components in the interaction mechanism. This is one of the features of complex systems, that is, they consist of many parts interconnected via their interactions. However, there is more to say on complex systems. In studying complex systems, the very first thing to start with is the distinction between linear and non-linear dynamics. The latter in turn cries for some remarks in advance about classical mechanics.

The rise of classical mechanics traces back to the end of seventeenth century being started with Newton's *Philosophiae naturalis principia mathematica* (1687). Right from the start, Newtonian mechanics enjoyed considerable success being thought as able to provide complete descriptions of reality, and more importantly, its future evolution. According to the vision of Newtonian mechanics, natural processes are strictly deterministic such that the knowledge of the present state of a system enables calculation of the system's state in any later time. We dealt with this view in detail in section 2.5.1 while speaking of Hamiltonian mechanics as a re-formulation of classical mechanics. In what follows, I will briefly mention the fundamental characteristics of the ontological vision of classical mechanics, dominated from Newton's era onwards, reaching its most complete expression by Laplace with whom I will deal in the next section.

1. Classical mechanics is a *reductionist* view, according to which the behavior of the system is completely describable on the basis of the behavior of the elements it is composed of.

2. The models[65] that classical mechanics provides are in principle *reversible* in time and space.

[65] In chapter 4, we will see that the notion of model is defined as a mental representation that, in a certain sense, replaces the world, or the universe.

3. As mentioned, the thesis is *deterministic*: the equations of motion used for the calculation of dynamics have unique solutions that permanently depend on the initial state and on the values of the parameters that describe the system's state.

The above characteristics are in line with the linear models of nature according to which the dynamic processes typical of systems are close to a configuration of stable equilibrium. The description of the evolution of the linear systems towards equilibrium can be approximated by means of a calculable evolutive trajectory (if the initial conditions are known), using a law of motion expressed in the form of a linear differential equation.[66] A physical system modeled by a linear differential equation enjoys linear causality meaning that there is a direct link between the cause that acts on the system, originating from an environment external to the system, and the change in the system's structure.

The enormous success of classical mechanics led scientists and experts in variety of areas, other than mechanics, to use the most important elements of Newtonian mechanics, especially his law of universal gravitation, as the most appropriate paradigm in their filed: sociologists, philosophers, economists, etc. applied the thesis in their studies. "This took place immediately following the publication of Newton's *Principia*, starting with Archbishop George Berkeley, who, although a harsh critic of the mathematical methods used by Newton […] was the first to attempt to construct a social analogy of Newton's law (Boyer, 1968; Cohen, 1993). Later, various studies by well-known philosophers, economists and sociologists such as David Hume, John Craig, Léon Walras and Émile

[66] Linear differential equations are characterized by equations in which the unknown function and its derivatives appear *only* summed together or multiplied by continuous functions of the independent variable. The general form of such equations would be as the following: $x^{(n)} + a_{n-1}(t)x^{n-1} + a_{n-2}(t)x^{n-2} + \cdots + a_1(t)x' + a_{0(t)}x = b(t)$ where the $a_i$ and $b(t)$ are continuous functions of the sole independent variable *t* and the $x^{(i)}$ are the *i*-th derivatives of the unknown function $x(t)$. If $b(t) = 0$ the above equation is called the homogeneous equation which has an interesting property: if two solutions, $x_1(t)$ and $x_2(t)$, are determined in some way, then $c_1x_1(t) + c_2x_2(t)$ is still a solution to the above equation. $c_1$ and $c_2$ in the latter expression are arbitrary coefficients. In general, any linear combination of the solutions also satisfies the homogeneous differential equation. Such a property is called the principle of superposition. By virtue of the principle of superposition, the general solution describing the dynamics of a system can be broken down into parts, all of which evolve individually over time. The principle indeed is a mathematical expression of reductionist point of view, linking it to the concept of linearity. It is always possible to solve linear differential equations when we have complete knowledge of either the boundary conditions (the coordinates at different times) or initial values of the dynamics (the coordinates and their derivatives) in exact terms.

Durkheim were made, right up to the construction, in more recent years, of gravitational models of spatial interaction (Sen and Smith, 1995)." (Bertuglia, & Vaio, 2005, p. 20)

Even during the period of the greatest achievements of Newtonian paradigm, physics began to develop in new directions, tending to distance from the general conceptual framework that had been associated with the original form of the classical mechanics. The most important of such developments are the formulation of the basic laws of electromagnetim, the elaboration of the kinetic theory of gases, the initiation of the use of statistical explanations for the laws of thermodynamics, which stemmed from statistical mechanics dealing with some macroscopic properties of matter statistically, rather than in the completely determinate way.

The above developments were not in complete contradiction with the framework of Newton's mechanistic philosophy. Rather, they showed the need for further enrichment of basic concepts of the vision. In *Causality & Chance in Modern Physics* (1957), David Bohm gives a detailed account of these later developments, showing how Newtonian philosophy could accommodate itself to deal with them. As an example, take the emergence of field theory. During the nineteenth century, it was found that the Newtonian scheme was not adequate to account for some electromagnetic phenomena. However, it was possible to enrich the scheme with the addition of a new set of entities known as the electric and magnetic fields. So, Newtonian mechanics, which postulated that things were constituted of bodies interacting according to specified forces, was enriched by the assumption that the whole of nature can be reduced to nothing more than the motions of a few kinds of bodies and a few kinds of fields, or according to Einstein, just nothing more than fields alone.[67] (Bohm, 1957, p. 31)

Later on, in the first decades of twentieth century, there appeared quantum mechanics,[68] which dealt with entities whose dimensions were close to or below the atomic scale, such as molecules, atoms, electrons, protons and other subatomic particles. It seemed that Newtonian paradigm was not able in principal to explain some phenomena at quantum domain. So, the appearance of such a field divided mechanics discipline into two sub-

[67] The case of statistical thermodynamics will be discussed in section 3.3.
[68] Note that I use the term quantum mechanics as a general expression in this context. Quantum mechanics here does not mean the quantum theory as formulated by Copenhagen school.

domains of classical mechanics and quantum mechanics. Initially, there were mainly two rival underlying philosophies behind the most important contributions to quantum mechanics: one was called standard (or orthodox) or Copenhagen interpretation of quantum mechanics, which considered quantum mechanics as intrinsically probabilistic, and so indeterministic in nature, and the other, which refused to accept quantum indeterminism and sought to demonstrate that the principle of indeterminacy could be violated, suggesting some thought experiments, which were assumed to permit the accurate determination of incompatible variables, such as position and velocity, or to explicitly reveal simultaneously the wave and the particle aspects of the same process. Heisenberg, Bohr, and Born were the main advocates of the former view, while Schrödinger and Einstein of the latter.

So far in this section, I have historically mentioned the new conceptions arisen following Newtonian mechanism the summary of which comes as what follows: electromagnetism and adding the concept of fields to classical mechanics, statistical thermodynamics as the application of statistical mechanics in thermodynamics, and quantum mechanics as an area that splits mechanics into two sub-disciplines, i.e., classical and quantum mechanics.[69] The latter case, namely, quantum mechanics will be discussed in the next section as this section has been devoted to the notion of complexity as far as it sheds light to the notion of pragmatic information.

Another discipline that emerged following the construction of Newtonian mechanism was "deterministic chaos," first proposed by Henri Poincaré at the end of nineteenth century. The notion was abandoned after its projection being renewed from the 1960s onwards by works of Lorenz, Smale, Yorke, and Prigogine, among others, and is still the subject of interest today. Chaotic behavior has been observed in the laboratory in a variety of systems including electrical circuits, lasers, oscillating chemical reactions, fluid dynamics, and mechanical and magneto-mechanical devices. Observations of chaotic behavior in nature include the dynamics of satellites in the solar system, the time

[69] This is a controversial statement to claim like this. We have, on the one hand, "correspondence principle" according to which the behavior of quantum mechanical systems reproduce classical physics in the limit of large quantum numbers. On the other hand, the specifications of individual quantum mechanical systems, like indeterminism according to Copenhagen interpretation of quantum mechanics, or nonlocality, in deterministic Bohmian interpretation of quantum mechanics (as we will see in chapter 5) are specifications which cannot be found in the macroscopic domain. In the next section, I will delve into the above discussion in more detail.

evolution of the magnetic field of celestial bodies, population growth in ecology, the dynamics of the action potentials in neurons, and molecular vibrations. Everyday examples of chaotic systems include weather and climate.

In this section, I will focus on deterministic chaos,[70] in a nontechnical manner, and on one of the elements that is part of a more general phenomenology that characterizes many natural and social systems, namely, complexity. As stated in the last paragraph, only those aspects of complexity, which contribute more clarifications to the notion of pragmatic information in general, and the distinction between information-driven and physical interactions in particular will be discussed in this section.

The very first specification of chaotic systems is that they are mathematically modeled nonlinearly. This means that causes and effects in real chaotic systems are not linked linearly. In mathematics, a nonlinear system is a system, which is not linear, i.e., a system that does not satisfy the superposition principle. Less technically, a nonlinear system is any problem wherein the variable(s) to be solved for cannot be written as a linear sum of independent components. Nonlinear differential equations, as the mathematical expression of nonlinear systems, are extremely diverse and methods of solution or analysis are very problem dependent. The violation from superposition principle is the greatest difficulty associated with solving nonlinear problems leading to impossibility to combine known solutions into new solutions. Only a limited number of nonlinear differential equations can be solved in exact term, but the discussion of whether such a property of differential equations is an essential one is still in progress.

Chaotic dynamical systems, that as said must be modeled nonlinearly, enjoy another property, that is, sensitivity to initial conditions: an infinitesimally small change of the initial conditions of the system changes the new final state drastically compared to the original one.[71] Sensitivity to initial conditions makes such systems difficult to predict. The feature of unpredictability of real chaotic systems is in line with the fact that they are mathematically modeled nonlinearly. Although controversial and still a matter of debate,

[70] I will confine this section to the examination of deterministic chaos in the classical, non-relativistic domain.
[71] It should be noted that although sensitivity to initial conditions is a necessary condition for a system to be chaotic, it is not sufficient; for a system to be chaotic, it must not evolve towards infinity for any variable.

the idea advocated in this essay is that such unpredictability arises *practically* because of errors *we* make in measuring or setting initial conditions. I will develop the latter idea, while speaking of emergent properties (in the next subsection) and the role of initial conditions in differentiating information-driven from physical interactions.

The above properties of chaotic systems, namely, sensitivity to initial conditions, unpredictability, nonlinear links between causes and their effects arise out of an essential property for these systems, i.e., the exponential growth of perturbations in the initial conditions, which make the behavior of chaotic systems *appear* to be random. The latter phenomenon is called positive feedback. Feedback, in general, is a process whereby some proportion of the output signal of a system is passed (fed back) to the input. While negative feedbacks tend to reduce the output, by amplifying the effects of the perturbations, positive feedbacks increase it.

Up to this point, I have mentioned some specifications of chaotic systems one of which is being far from equilibrium as opposed to systems that tend towards equilibrium. During the evolution of the system, there is an intermediate state between the two mentioned states usually indicated by the expression "complexity" in which the system displays a different behavior, both from stable equilibrium tendencies that are unchangeable in the future, and from chaotic tendencies. Complexity thus has to do with transitions between states. Understanding how transitions can happen has to do with the concept of "control parameters," which describe the different ways a system is coupled to its environment and affected by it. With certain values of control parameters, the system displays its tendency to create new structures and new functional properties that did not exist for the other values of control parameters. To sum up, there is an intermediate area between order and chaos, the area at the edge of chaos[72], in which we witness sudden emergence of unexpected properties. In such an area, the dynamical system displays a very particular type of behavior that is called complexity.

"The fundamental characteristic of complexity is the fact that, in the study of the evolution of dynamical systems that complexity deals with, the nature of the system in

[72] This is the name given to the critical point of the system, where a small change can either push the system into chaotic behavior or lock the system into a fixed behavior. It is also regarded as a phase change. It is at this point where all the really interesting behavior occurs in a "complex" system.

question is usually irrelevant. In the vision proposed by complexity, we can identify forms and evolutive characteristics common to all, or almost all, systems that are made up of numerous elements, between which there are reciprocal, nonlinear interactions and positive feedback mechanisms. These systems, precisely for this reason, are generally called complex systems." (Bertuglia, & Vaio, 2005, p. 269)

Science of complexity takes systemic (collective) properties of a system into account and shows how a complex system can spontaneously self-organize when such a dynamical system finds itself in a state distant from equilibrium while there is no external force acting on the system. With self-organization, a new order of the system "emerges." This is an order of the non-equilibrium, non-static one where the whole evolutive dynamics tends to dissolve into a stable equilibrium. In self-organization, a new system structure appears without any external pressure, apart from a regular inflow of free energy from its environment and a disposal of entropy to the outside; the system shows certain properties called "emergent properties," as such properties are not intrinsically identifiable in any of its parts individually, appearing when the system is observed as a whole. I will discuss the concept of emergence in more length in the next subsection. There, we will see that the unexpected properties arisen from self-organizing systems are weakly emergent. This section has been organized to shed light to the notion of complexity (and self-organization as one of the characteristics of complex systems) as the notion appears as one of the basic requirements of an interaction to be information-driven.

But more importantly, the aim of this section is to emphasize the point that *complexity and self-organization alone do not represent information*. In the latter phenomena, local interactions between the individual agents[73] are explainable based on purely physical interactions. They are all the result of the forces responsible for the respective systems' formation. But "since we humans can introspect our own brain function […], we view self-organizing systems – whose form and structure makes them more "predictable" – somewhat akin to ourselves, even when they pertain to the physical, inanimate domain. We marvel about a growing crystal and the singing tone of a whistling wind – yet these are all natural, logical and mathematical results of physical laws." (Roederer, 2005, p. 80)

[73] Each part of a complex system is modeled as an *agent*.

As said earlier, in nature, the emergent property of regularity in self-organizing systems is the result of the forces responsible for the systems' formation. At the macroscopic level, entropy decrease resulting from organization and regularity requires regular inflow of free energy from surrounding and a disposal of energy to the outside (Chaisson, 2001). Self-organizing systems are thermodynamically open. This means that, as far as the second law of thermodynamics is concerned, the localized entropy at the system's side decreases while a greater entropy increase is witnessed in the rest of the Universe. This always leads to the sum-total increase of the entropy. So, the second law of thermodynamics is naturally upheld without any need to include the concept of "purpose" in the process. However, during the evolution of the Universe some sorts of organized structures based on carbon atom chemistry did emerge that were capable of maintaining a metastable, nonequilibrium state of low-entropy and high organization in a changing and, often, hostile environment, namely, the *living organisms*. To make this "consistently beating the elements" possible, Roederer counts three conditions to be fulfilled concurrently:

> "1. [*E]ncapsulation* in a protective but permeable boundary that allows for a highly selective interchange of matter and energy with the environment; 2. *self-adaptation* to environmental change; and 3. *reproduction*." The puzzling thing is that this triad, which can be summarized in reverse order as "genetics, metabolism and containment," must have developed simultaneously in tightly coordinated steps, not one after the other. Condition 2 requires the operation of mechanisms that establish correspondences between some critical spatial and temporal patterns in the environment and dynamic actions within the organism needed for metabolism, adjustment and regeneration; […] we are talking about the capability of entertaining *information-driven interactions* with the environment. (Roederer, 2005, p. 126)

According to Roederer, it is possible for inanimate natural bodies to fulfill condition 1, as happens, for example, in some geologic minerals, and condition 3, as in prebiotic polymers, but condition 2 is fulfilled only in living organisms.

> "[O]nly living organisms can entertain information-driven interactions to systematically maintain a low-entropy, highly organized

> quasiequilibrium state despite nonperiodic, stochastic degrading influences from outside. Inert systems able to do this would necessarily be human-made, or fabricated by human-made machines." (ibid)

But self-organization can happen both in living and nonliving systems. In other words, self-organization can appear as the result of both information-driven interactions and physical ones. In chapter 5 of his book, Roederer (2005) shows how this is possible, using the notion of entropy in thermodynamics. He, then, concludes:

> "If an ordering process happens naturally in a system without the intervention of a living system (i.e., exclusively on the basis of force-based interactions leading to self-organization, […]), the system must have been open to a high-grade/low-grade energy exchange and the entropy decrease must have occurred at the expense of a larger entropy increase elsewhere […]. If, on the other hand, the ordering process is the result of information-based interactions ([…], the synthesis of a protein), the entropy decrease must occur at the expense of a greater entropy increase within the interaction mechanism itself (living system or a human-designed machine) that mediates the information-driven interaction (pattern recognition mechanism, metabolism of the brain, radiated heat, etc.)." (ibid, pp. 185-186)

As stated earlier in this section, a self-organizing system shows certain properties called "emergent properties," which are not intrinsically identifiable in any of its parts individually, appearing when the system is observed as a whole. In the next subsection, I will examine the concept of "emergence" in more detail and find the place of some important phenomena, as regards the purpose of this essay, in the varieties of emergentism.

### 3.2.2 Emergence

As mentioned in the last section, there are varieties of emergentism each of which appropriate for classifying a certain group of phenomena. In this section, I will go more through details in this respect.

British emergentists of the late-nineteenth and early-twentieth centuries were the first who worked comprehensively on the emergentist picture, although they were not the first who posed the idea in general. Chemistry and Biology were two fields around which much of the discussions of emergentism in this era were centered. Concerning the question of whether or not the constitutive principles and features of the mentioned sciences were reducible to those of the corresponding "lower level" sciences of physics and chemistry respectively, two competitive groups are discernable. "Reduction-minded 'mechanists', who supposed that the processes of life were governed wholly by physical-chemical principles, contended with the extreme anti-reductionist 'vitalists,' who posited an entelechy, a primitive substance or directing principle embodied in the organism which guided such characteristic vital processes as embryonic development and the regeneration of lost parts. Emergentists sought to develop a middle way, eschewing vital substances but retaining — in *some* sense — irreducibly vital qualities or processes." (O'Connor & Yu Wong, 2006)

Strong emergence is a notion put forward by the British emergentists of the 1920s in contrast with weak emergence. To understand the differences briefly, Stephan's (1999b) classification of the varieties of emergentism, as shown in figure 3.1 will help.[74]

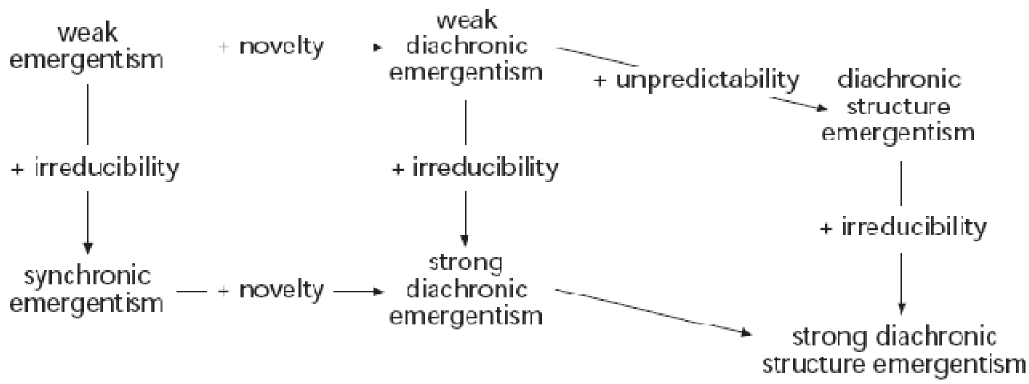


**Figure 3.1.** Varieties of emergentism as classified by Stephan (1999b)

*Weak emergentism*, according to contemporary theories of emergence, has three features:

1. *Physical monism*, according to which "entities existing or coming into being in the universe consist solely of material parts. Likewise, properties, dispositions,

[74] See also his other works: Stephan (1992, 1997, 1998, 1999a, 2004)

behaviors, or structures classified as emergent are instantiated by systems consisting exclusively of physical parts." (Stephan, 1999, p. 50)

2. *Systemic properties*, according to which "emergent properties are systemic properties. A property is a systemic property if and only if a system possesses it, but no part of the system possesses it." (ibid)

3. *Synchronic determination*, according to which "a system's properties and dispositions to behave depend nomologically on its micro-structure, that is to say, on its parts' properties and their arrangement. There can be no difference in the systemic properties without there being some differences in the properties of the system's parts or their arrangement." (ibid, pp. 50-51)

As the systemic properties arise out of the feature of collectivity of a system, they are considered to be "novel." Concerning whether systemic properties are timeless or temporal, Stephan distinguishes between synchronic and diachronic novelties and prefers to characterize only diachronic novelties as being "novel."

Adding the thesis of *irreducibility* to weak emergentism will result in *synchronic emergentism* as a strong type of emergentism. Stephan defines irreducibility as what follows:

> "A systemic property is irreducible if (a) it is neither micro-nor macro-scopically behaviorally analyzable, or if (b) the specific behavior of the system's components, over which the systemic property supervenes, does not follow from the component's behavior in isolation or in other (simpler) constellations." (ibid, p. 53)

Strong diachronic emergentism results from adding the temporal dimension of the thesis of novelty to the synchronic emergentism.

By adding a temporal dimension in the form of the thesis of novelty to weak emergentism, *weak diachronic emergentism* will result (see figure 3.1). Diachronic structure emergentism is the result of weak emergentism and *structure unpredictability*, whose definition comes in the following:

> "The rise of novel structures is unpredictable in principle, if their formation is governed by laws of deterministic chaos. Likewise, any novel properties that are instantiated by those structures are unpredictable in principle." (ibid, p. 54)

Diachronic structure emergentism is weaker than synchronic emergentism being compatible with reductive physicalism.

Finally, strong diachronic emergentism and strong diachronic structure emergentism result from adding the thesis of irreducibility to weak diachronic and weak diachronic structure emergentism respectively.

Quite many phenomena can be categorized as weakly emergent. Among others, the emergence of high-level patterns in cellular automata, the way we deal with cognitive sciences currently, especially in characterizing the collective properties of connectionist networks, and some properties of self-organizing systems are some phenomena in which weak theories of emergentism can be used. Weak diachronic structure emergentism is also a good candidate for evolutionary research as Stephan declares.

The above-mentioned phenomena are categorized as being weakly emergent as far as one can stay faithful to "physical monism" as one of the features of the thesis of weak emergentism. However, if we think of consciousness as a feature of the world ontologically over and above physical features, then the notion of strongly emergent properties comes into play. Take the phenomenon self-organization for example. As said, the phenomenon can occur because of both physical interactions (in natural inanimate systems) and information-driven interactions (in living systems). In the former, the properties arisen from totality of the system are compatible with the criteria for being weakly emergent. But as regards the organization emerged based on information-driven interactions, we may face with two properties: physical and phenomenal. Although, compared with physical interactions, information-driven interactions acquire more properties, which make them occur just in living systems, they are *reducible* to the laws of physics and chemistry according to Roederer:

> "*Is information reducible to the laws of physics and chemistry?* Our answer now should be: *If defined in an objective way it is reducible, indeed – but only in the biological realm.*" (Roederer, 2005, p. 123)

But information, as mentioned in chapter 1 and will be discussed in more detail in chapter 6, may have another aspect other than physical one that is not reducible to the laws of physics and chemistry, namely, a *phenomenal aspect*. This is the latter aspect that may give rise to phenomenal consciousness. Here the claim of this essay is that there exists at least one clear case of strongly emergent phenomenon, that is, the phenomenon of consciousness. This idea is in line with Chalmers' (2006) thesis on consciousness. In a less technical, however still very expressive, article he introduces phenomenal consciousness as a strongly emergent phenomenon. Chalmers defines strongly emergent properties as what follows:

> "We can say that a high-level phenomenon is *strongly emergent* with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are not *deducible* even in principle from the truths in the low-level domain." (Chalmers, 2006)

He also defines weakly emergent properties as what follows:

> "A weakly emergent property of a system is an interesting property that is unexpected, given the underlying principles governing the system." (ibid)

The attribute "interesting" above suggests that weakly emergent properties are observer-relative. Likewise, the attribute "unexpected" measures "how difficult it is for an observer to deduce the high-level property from low-level properties." (ibid) In this sense, weakly emergent properties are characterized as *interesting*, *non-obvious consequences* of low-level properties. In order to be understood in multiple ways, Chalmers brings the notion of "underlying principles" deliberately vaguely. Some examples may clarify the idea. In case of evolution, the underlying principles are considered operating at the level of the gene, while the complex, interesting, high-level properties, such as intelligence, are considered unexpected, relative to the underlying principles. Thus, such high-level properties count as emergent. On the other hand, if we consider the underlying principles as operating at the level of the organism, then a property like intelligence is less unexpected, and thus less clearly emergent (ibid).

Let us get back to the notion of strong emergence as discussed by Chalmers (2006) as his ideas on this notion play a crucial role in this essay.

As said above, in Chalmers' opinion, perhaps the phenomenon of consciousness is the only case of strong emergence. In some works (Chalmers, 1996 & 2003), he argues this position at length reasoning that the facts about consciousness are not deducible from any number of physical facts and "if there are phenomena that are strongly emergent with respect to the domain of physics, then our conception of nature needs to be expanded to accommodate them. That is, if there are phenomena whose existence is not deducible from the facts about the exact distribution of particles and fields throughout space and time (along with the laws of physics), then this suggests that new fundamental laws of nature are needed to explain these phenomena."(Chalmers, 2006)

In chapter 1, I mentioned the most influential arguments against the deducibility of consciousness from physical facts. But it is also worth recalling two general lines of reasoning against physicalism shortly. The first argues that a colorblind neurophysiologist with all physical information about what goes on when one sees red color will never be able to deduce what it is like to have the conscious experience of red.[75] The second contends that it is logically coherent in principle to conceive a world physically identical to this one, but lacking consciousness entirely or different conscious experiences from what we experience.[76]

The non-deducibility of consciousness from physical facts suggests two things: first, it suggests that consciousness is a strongly emergent property arisen from physical processes. In other words, although consciousness is not deducible from physical facts, it seems quite plausible for the states of consciousness to be systematically correlated with the physical states. Second, it suggests that consciousness itself can be considered a

[75] The argument was mentioned in chapter 1 entitled Jackson's Marry.
[76] The argument was mentioned in chapter 1 entitled Chalmers' zombie.

fundamental property. If this is the case, then we need a further basic law or laws, what we might call fundamental psychophysical laws, to ground the connection between the physical properties and consciousness. Notice that according to this view, "the lawful connection between physical processes and consciousness is not itself derivable from the laws of physics but is instead a further basic law or laws of its own." (ibid)

To sum up, Chalmers distinguishes two clear theses of emergence, weak and strong emergence, together with an intermediate level. According to him, almost all phenomena but consciousness are weakly emergent. However, one might think of an intermediate "but still radical sort of emergence in which high-level facts and laws are not deducible from low-level *laws* (combined with initial conditions). If this intermediate sort of emergence exists, then if our Laplacean super-being is armed only with low-level laws and initial conditions (as opposed to all the low-level facts throughout space and time), it will be unable to deduce the facts about some high-level phenomena." (ibid)

The latter suggests the possibility of *downward causation* according to which higher-level phenomena, which are also irreducible, exert a causal efficacy of some sort. If so, one needs "configurational laws," as McLaughlin (1992) calls it, in order to formulate the principles, which state what happens to the lower levels when certain high-level configurations occur.

One should also distinguish *strong* downward causation from *weak* downward causation whose definitions come as the following:

> "With strong downward causation, the causal impact of a high-level phenomenon on low-level processes is not deducible even in principle from initial conditions and low-level laws. With weak downward causation, the causal impact of the high-level phenomenon is deducible in principle, but is nevertheless unexpected." (Chalmers, 2006)

Are there any examples of strong downward causation? As stated before, except for the case of consciousness, Chalmers does not see any compelling evidence for high-level facts and laws not to be deducible in principle. However, it is not impossible to encounter some. For example, the collapse of the wave function in conventional quantum mechanics can be a good candidate to be analyzed by the thesis of downward causation.

The evolution of the quantum wave function undergoes two distinct processes: before measurement takes place, linear Schrödinger equation governs the evolution of the wave function. Schrödinger equation is an equation that describes how the quantum state of a physical system varies. The equation evolves deterministically, that is, if one knows the initial conditions at the initial time $t^0$ , she can predict the state of the quantum system at any other time *t*. However, when measurement takes place, the wave function undergoes a sort of nonlinear "quantum jump." It might be likely such a phenomenon not to be deterministic (for example according to Copenhagen interpretation) at all; no one knows of the rules governing the process of measurement indeed. Yet, clearly, measurement must involve certain causal events, as Chalmers states, most likely at a high-level. If this is the case, then measurement involves downward causation.

> "Both consciousness and the quantum measurement case can be seen as strong varieties of emergence in that they involve in-principle non-deducibility and novel fundamental laws. But they are quite different in character. If I am right about consciousness, then it is a case of a strongly emergent quality, while if the relevant interpretations of quantum mechanics are correct, then it is more like a case of strong downward causation.
>
> In principle, one can have one sort of radical emergence without the other. If one has strongly emergent qualities without strong downward causation, one has an epiphenomenalist picture on which there is a new fundamental quality that plays no causal role with respect to the lower level. If one has strong downward causation without strongly emergent qualities, one has a picture of the world on which the only fundamental properties are physical, but on which their evolution is governed in part by high-level configurational laws." (ibid)

In case of having both strongly emergent qualities and strong downward causation, according to Chalmers, we reach the orthodox interpretation of quantum mechanics, that is, Copenhagen interpretation. According to Copenhagen school, consciousness is not epiphenomenal, but responsible for the collapse of the wave function. In this interpretation, consciousness plays a crucial causal role during measurement.

No one knows that whether or not strong downward causation takes place at all. However, if it does, Chalmers sees quantum mechanics as an appropriate realm for its occurrence. On the other hand, *if there exists a strongly emergent phenomenon, consciousness will be the best – and most likely the only – candidate.* The latter idea is what is advocated in this essay.

As seen in this section, whether a phenomenon is considered strongly emergent or weakly emergent highly depends on whether high-level facts and laws are deducible in principle from low-level laws and initial conditions. In section 2.5.1, while speaking of physical interactions, I also dealt with "initial conditions" as playing a crucial role in differentiating physical interactions from information-driven ones. There, it was said that the end effect of a physical interaction would always depend on some initial conditions, while the final state of a system with information-driven interactions is much less dependent on the initial conditions. As initial conditions play such an important role in a phenomenon to be strongly or weakly emergent as well as an interaction to be physical or information-driven, I have devoted the next section to the notion of initial conditions.

## 3.3 Initial Conditions

Consider the following three ensembles: flock of birds, solar system, and billiard balls on a frictionless table as a system, which *we* assemble *purposefully* in the laboratory as discussed in section 2.5.2. These three systems are open systems, that is, the components of any ensemble interact with the rest of the Universe as well as with each other. At some initial time $t^0$ , these three ensembles are in a certain initial state. There are two ways of determining the initial states of the ensembles: in the first two cases, namely the flock of birds and the solar system, it can be done by appropriate measurements as these systems run independent of us and we have no direct control over them. Besides doing measurement, however, we can deliberately set the initial conditions of the third case, that is, in the case of billiard balls on a frictionless table. The question now is that "given the initial states of these three systems, can we find mathematical algorithms, which allow us to predict the state of each system at a later time *t*?" Let us check the answer case by case.

In the case of the flock of birds, it is not possible to predict the state of the ensemble at any later time only by knowing of its initial state. The reason is that interactions between the birds and the environment on the one hand and between the birds themselves on the other are information-driven and fall into the realm of biology.

In the case of planets in the solar system, however, it is completely possible to predict the final state of the system once we know the initial state of it,[77] though the system is not under our control. The interactions in such a system are purely physical.

One is also tempted to think of the case of the billiard balls as entertaining purely physical interactions as the system resembles the latter case of the solar system. However, there is one big difference, which changes the characteristics of the

[77] This would be the case when we model such systems, in first approximation, linearly according to Newtonian paradigm. It is also possible to see the solar system, for example, to behave chaotically. If we take the interaction between the Sun and the various planets of the solar system not in pairs, but as a whole, then the system may be modeled nonlinearly. For example, the mutual attraction between Jupiter and Saturn causes periodic disturbances to the movement of the two planets. These disturbances cause increasingly significant effects over time, in line with the typical behavior of chaotic dynamics.

interactions in the system of billiard balls turning them into information-driven. Contrary to the case of the solar system, the system is completely under our control. This means that it is us who determine the initial conditions. This makes the system of the billiard balls purposeful. As stated earlier, such a purpose has to do with the ability of eliciting the same output change in every run. Without such a purpose, there is no difference between the solar system and the system of billiard balls on a frictionless table; without purpose, both would entertain physical interactions. An interesting point about the solar system is that when biological systems interact with such systems, the interaction would count as information-driven. This occurs, for example, when we model physical systems. I will deal with this in the next chapter in more details.

> "When *we* set the initial conditions of a physical system, whether it is by actually manipulating it or whether it is only in thought by manipulating a model of it, the system becomes *information-driven*, […]. If on the other hand the initial conditions for and time $t_0$ are "given by nature," as is the case with the planets […], any purpose disappears, the correspondence between initial and final state is a purely physical one, and information plays no active role in the process. All that counts are the individual physical interactions between the mass points." (Roederer, 2005, p. 170)

Recall what was said in section 2.5.1 about a classical closed Hamiltonian system with conservative forces. Such a system is reversible and predictable in principle. Complete knowledge of the initial state of these systems leads to the predictability of the final state with certainty. Any lack of *our knowledge* of initial conditions or errors in setting or measuring them (occurred by us) leads to the uncertainty in the predictability of the final state. However, the system is still strictly deterministic. In such a case, we are faced with the probability distributions of possible evolutions of the system.

This is also what happens in chaotic systems with which I dealt in the last section. As regards the aim of this section, the claim is that interactions which make systems behave chaotically in nature obey the rules counted for physical interactions. The evolution of a classical closed chaotic system is not predictable and reversible for merely "practical" reasons as *we* will never be able to determine physically the *mathematically exact* initial conditions required to repeat exactly the motion. If we were able to do that, the system would be completely predictable. Such a system is completely deterministic. Bricmont (1995), among others, is one of the proponents of such a thesis.

According to him, the classical deterministic world-view is not invalidated by the thesis of chaos. He defends Boltzmann's thesis (and also Laplace's determinism), as opposed to Prigogine and Stengers, according to which one can account perfectly for macroscopic irreversibility on the basis of deterministic, reversible, microscopic laws.[78] So, contrary to Prigogine's thesis, for the chaotic systems, the statistical description is not irreducible.

According to Bricmont (1995), during the course of scientific development, it has widely been believed that chaos implies the rejection of classical, deterministic worldview projected by Laplace in his *A Philosophical Essay on Probabilities*. But this is just a confusion which stems from the confusion between *determinism* and *predictability*. "Determinism has to do with how nature behaves, and predictability is related to what we, human beings, are able to observe, analyse and compute." (Bricmont, 1995) As seen, this account considers predictability as subjective and determinism as objective.

What does determinism exactly mean? According to Bricmont, the most complete answer has been given by Laplace himself:

> "An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in one single formula the motions of the largest bodies as well as the lightest atoms in the world, provided that its intellect were sufficiently powerful to subject all data to analysis; to it nothing would be uncertain, the future as well as the past would be present to its eyes." (Laplace, 1825)

In the above quotation, determinism, according to Laplace, depends on what the laws of nature are. Given the state of a system at a certain time, a formula which describes the laws of nature will give us the state of the system at any later time. This formula can be a differential equation or a map. When it comes to the notion of predictability, however, one has to *know* of the state of the system by fulfilling measurements with enough

[78] See numerous works of Prigogine (1980, 1983, 1994, 1995, and 1996) and Prigogine & Stengers (1979, 1984, and 1988) for their ideas on chaos and irreversibility.

precision at the initial time. Then, using the formula, she has to compute the state of the system at any later time. If there is incomplete knowledge of initial conditions or inability in setting them, concepts of probability as well as unpredictability appear. So, an important point in construing the above quotation is that one must not attribute the thesis of complete predictability to the Laplace's account of determinism. There are always limitations in our access to the complete knowledge of initial conditions or in setting them, leading to a sort of probabilistic account.

> "The idea of 'absolute precision' in determining a system's state is an abstract concept (or even intrinsically wrong, completely lacking meaning in practice, absolutely unachievable in any measurement […].
>
> We can never have *thorough* knowledge of reality, because we cannot measure any of the state variables with absolute precision, even if we assume that we have identified all of the state variables needed to fully define a system (which, if a system is complicated, is by no means necessarily easy, or may even be impossible." (Bertuglia, & Vaio, 2005, p. 136)

Our inability to make measurement (or to know or to set initial conditions) with absolute precision leads to the thesis of unpredictability. However, unpredictability or our probabilistic accounts are due to the lack of *our* knowledge and do not contradict with determinism which is due to the laws of nature. This is also true for the chaotic dynamical systems.

> "Chaotic dynamical systems are of course unpredictable in practice, at least for long enough times, since there will always be some error in our measurement of the initial conditions. But this does not have any impact on our discussion of determinism, since we are assuming from the beginning that the system obeys some deterministic law. It is only by analysing this deterministic system that one shows that a small error in the initial conditions may lead to a large error after some time. If the system did not obey any law, or if it followed a stochastic law, then the situation would be very different. For a stochastic law, two systems with the same initial condition could be in two very different states after a short time." (Bricmont, 1995)

Such an idea on chaos confirms the idea according to which the properties arisen from chaotic dynamical systems are weakly emergent.[79]

Some words should also be said about irreversibility in chaotic systems. The main question here is: if the basic laws of nature are reversible, why are motions at a chaotic closed system unidirectional, not bidirectional? In other words, why do not we see time-reversed motions at macroscopic level? First, I have to emphasize that there is no contradiction between the basic laws of physics (which are reversible inherently) and irreversibility. So, Prigogine and Stengers' claim that irreversibility is either true at all levels, that is, at micro and macro levels, or at none, does not seem to be reasonable. Second, there is also no contradiction in considering an irreversible chaotic closed system to be reversible in principle. Whether a motion evolves in one direction or in the reversed one depends on the initial conditions. In addition to the notion of initial conditions, another notion comes into play, as regards the classical explanation of irreversibility, namely, many degrees of freedom. Remember billiard balls on a frictionless table discussed in section 2.5.2. As far as there are just few balls, the system is totally reversible. It gets irreversible when we increase the number of balls. We are faced with irreversibility in the latter system because to have the system in the exactly reverse way, it will not be enough all balls to start with reversed velocities: "*Many* other conditions must be fulfilled exactly (e.g., each collision partner inside the "black box" must arrive at the right place and time with the right (reversed) velocity)." (Roederer, 2005, p. 111)

Finally, another question concerning irreversibility is whether the notion is subjective or not. Here, I have to emphasize that macroscopic variables are objectively determined by microscopic ones. Irreversibility at macroscopic level is completely objective, not a sort of mind projection fallacy. Macroscopic variables behave as they do, whether or not we look at them. Now let us turn our attention to a single molecule or at a collection of molecules represented by a point in phase space. Doing so, if we are not to consider some of the macroscopic variables determined by such a collection of molecules, irreversibility disappears. The issue is thus subjective as far as our concentration from one level to the

[79] I have to caution the reader that all I have written on universal determinism in Laplacean sense, which can also be called universal reductionism, holds just for the classical physical world. As I argued in the last section as well as chapter 1, Laplacean determinism does not hold for consciousness. The case of quantum domain has also its own construal, which will be discussed soon.

other matters: although macroscopic variables are determined by microscopic ones, we see irreversibility in the former, but not in the latter. Note that irreversible phenomena could be reversible if we had the knowledge of initial condition. In this context, "knowledge" means having control over initial conditions.

What about open systems? It is known that except for the Universe as a whole, there exists no closed system in practice. An open system consisting of interacting mass points would be a system subjected to external forces. This makes the system not to be reversible in general any more. If we reverse all momenta and velocities of the mass points at time *t*, we do not reach the same initial state at time $t^0$. The reason is that to get to the same initial state at time $t^0$, we also need to reverse all external interactions, a task which is normally not possible in practice.

Let us turn our attention to the quantum domain. To fulfill such a task, some definitions and descriptions are needed. First, I pose the definition of a *quantum bit* (Qbit) as opposed to a classical bit. In section 2.3, we saw that a classical bit is the physical realization of a system, which has acquired the following specifications after appropriate preparation and operation:

1. There exist two clearly distinct, mutually exclusive states such as 0 – 1, yes – no, true – false, on – off, etc.

2. The system can be measured in order to determine which state, out of two, has been realized. If the probability of the realization of the two states is the same *a priori*, the realization of each state provides one bit of information.

The difference between a Qbit and Cbit is that for the former the first specification of Cbit does not hold. In other words, a Qbit is a physical system with two basis states without the requirement of being mutually exclusive.

> "A Qbit is a register which can be prepared in a superposition of two different component states – representing an infinite number of

possibilities with different relative amplitude and phase." (Roederer, 2005, p. 65)

We can simply show a Qbit as what follows:

$$|\psi\rangle = \cos\alpha|0\rangle + e^{i\varphi}\sin\alpha|1\rangle \qquad (3.1)$$

In the equation (3.1), $|\psi\rangle$ stands for the general state of the quantum system, $\cos\alpha$ and $e^{i\varphi}\sin\alpha$ are probability amplitudes, which fulfill the normalization condition, i.e.:

$p_0 = \cos^2\alpha$, $p_1 = \sin^2\alpha$, and $p_0 + p_1 = 1$.

$\varphi$, in equation (3.1), is the phase.[80]

There is also a fundamental difference between a Cbit and a Qbit: the measurement process does not change the state of the Cbit, while it does not hold for the Qbit. The final state in a Cbit is defined according to the interaction between the ball and the pin. This has nothing to do with the measuring device or measurement processes. In quantum systems, however, the final state is the result of the interaction between the quantum system and the measuring device (or, generally, environment, which counts as a macroscopic entity).

When a quantum system interacts with the measuring device, the information extracted out of such an interaction has to be accessible to our sensory system. So, we are confronted with two different ends: at one end, the quantum system in which quantum rules are dominant and the measuring device, at the other end, whose interaction with the quantum system leads to the translation of nonlocal quantum information into the classical signals with finite informational content.

Let us deal with the above subject in further detail. To do so, first I have to discuss some notions in quantum mechanics. The very first and the most important notion is "wave

[80] See (Roederer, 2005, chapter 2) for detailed description of Qbits and also Quantum Pinball Machines.

function." A wave function is a mathematical tool used in quantum mechanics to describe any physical system. It is a function from a space that maps the possible states of the system into the complex numbers. The time-dependent Schrödinger equation describes how the wave function evolves over time. The values of the wave function are probability amplitudes — complex numbers — the squares of the absolute values of which, give the probability distribution that the system will be in any of the possible states.

Let $\psi$ stand for the wave function. Then, the time-dependent Schrödinger equation will be as what follows:

$$i\hbar \, {}^{\partial\psi}\!/_{\partial t} = H\psi \tag{3.2}$$

Where $\psi$ is the wave function, $\hbar$ is Planck's constant and $i$ is the imaginary unit. Equation (3.2) is valid for a system of $N$ material points. $H$ in the above equation is the quantum operator equivalent to the Hamiltonian in classical mechanics. As a standard example, a non-relativistic particle with no electric charge and zero spin has a Hamiltonian which is the sum of the kinetic (T) and potential (V) energies. In one dimension, say $x$, $H$ equals the following:

$$H = {}^{p^2}\!/_{2m} + V = -({}^{\hbar^2}\!/_{2m})\,{}^{d^2}\!/_{dx^2} + V(x) \tag{3.3}$$

The first term in equation (3.3) corresponds to the kinetic energy of the particle and the second to the potential energy.

There are two important points about the time-dependent Schrödinger equation (3.2). First, the equation is valid only for truly *closed systems*. As stated in the last section, the equation evolves deterministically, that is, if one knows the initial conditions at the initial time $t^0$, she can predict the state of the quantum system at any other time $t$.

However, as we will see, the way we use the notions of determinism, and also causality, in quantum world are totally different from the what we consider in classical world. And this is the second point. We have no direct access *in principle* to the quantum domain and all information we acquire comes from the interactions between the external influences of the macroscopic world (measuring devices) and the quantum system. In such a case, mutually entangled states will break down into the individually superposed states but mutually disconnected states (see Roederer, 2005, chapters 2 and 5 for more detail) and the latter will also break down (collapse) into a statistical mixture of zillions of independent states. “If this were not to happen with the ensembles of particles in macroscopic bodies, they would be able to retain entangled properties – individual objects would lose topological identity and form, macroscopic variables could not be defined, our brains would be in a permanent state of multiple personality, and the “macroscopic” world would be very strange.” (Roederer, 2005, p. 193)

Let me develop the discussion posed above. Regarding the interpretations of quantum theory, as also mentioned in section 3.2, there exist mainly two rival theories: Realist and Copenhagen interpretations. Schrödinger and Einstein were two main proponents of the former while Heisenberg, Bohr and Born are counted as the main advocates of the latter.

As a realist, Schrödinger believed that his wave mechanics describes part of an independent underlying reality in quantum world. On the contrary, Heisenberg took a positivist stance, developing his matrix mechanics, which, according to him, is nothing more than an algorithm through which the results of experimental observations could be correlated and new predictions made. Interestingly, Schrödinger demonstrated that the two approaches are mathematically equivalent.

How could it be? How is it possible that two theories, which are constructed based on two opposing underlying philosophies, one based on realism and the other based on anti-realism, lead to the same results? I think the answer lies in what I posed in the above paragraph: quantum domain is not directly accessible in essence; any consistent theory that can correlate experimental observations and make predictions can be an acceptable theory. The result of any experiment to explore a quantum system will result in decoherence, which is the mechanism by which the quantum systems interact with their environments to exhibit probabilistically additive behavior – a feature of classical physics

– and give the *appearance* of wave function collapse. We have no means by which individual quantum world can be explored.

The idea posed above confirms neither realist nor Copenhagen interpretation of quantum mechanics. It does not confirm a realist interpretation because it holds that we cannot demonstrate the reality of the individual quantum process by any means. Therefore, there is no way intuitively or otherwise to understand what is happening in such processes. However, it does not imply that the individual quantum process is not real or there is no consistent realistic interpretation of quantum mechanics. Any interpretation which can consistently correlate the results of experimental observations, either realist or anti-realist, gets equivalent credit comparing to the others with the same descriptive and predictive power. This is why, as we shall see in more detail in the next chapter, Bohmian mechanics, which gives a non-local causal interpretation of quantum mechanics predicts the same results as the orthodox interpretation, which is inherently indeterministic and statistical.

This implies any description at quantum domain to be *our* description. Coming back to the notion of wave function, one can say that "the wave function is an abstract mathematical bridge in *our* description of a quantum system, linking the initial conditions that *we* set for a system *we* have defined with the responses that *we* obtain with *our* measurement instruments – and these instruments as well as ourselves participate as an *integral part* of the system." (Roederer, 2005, p. 195)

The same goes for the causal or deterministic interpretations of quantum mechanics, such as Bohmian interpretation. The notions of causality and determinism in quantum domain are inherently deferent from what we use in the classical domain. Deterministic or statistical descriptions in quantum realm are *our* descriptions. This is why we see that in deterministic interpretations of quantum mechanics, say in Bohmian mechanics, causal effects are nonlocal, that is, we have action-at-a-distance, while in the macroscopic world, according to special relativity as a law of nature, such an action-at-a-distance is prohibited.

Causality and determinism as representing laws of nature apply only to macroscopically measurable effects. Measurement in quantum systems leads to the collapse of the wave function, which in turn leads to a change in the individual quantum system. By doing measurement, we will acquire nothing more than some knowledge about the interaction between a quantum system and our macroscopic measuring device. Measurement does not lead to the increase of our knowledge about individual quantum systems *per se*.

What was stated above about the inaccessibility of individual quantum processes by implementing measurement also holds for consciousness. We, as conscious beings, have direct and intimate knowledge of the facts about our own consciousness. Such direct knowledge leads to a subjective, intrinsic, and qualitative feel, i.e., conscious experience. This is one of the differences between quantum processes and consciousness as we are not able to have any direct knowledge of quantum processes, in principle. However, as far as obtaining knowledge through implementing measurement is concerned, consciousness and quantum systems are similar. Consciousness is not measurable in principle. At best, we are able to measure neural correlates of consciousness (NCC).

On the other hand, although we have direct knowledge of our own experiences, trying to know facts about others' conscious states will simply fail. This is the epistemic asymmetry mentioned in section 1.2, page 11, the restatement of which seems to be useful: such an asymmetry, epistemic asymmetry, brings an important distinction between two systematic approaches about, namely, first-person perspective versus third-person perspective: "From the inside, consciousness seems all-pervasive, self-evident, and undeniable […] from the outside, firsthand exploration of the consciousness of others just seems to be out of the reach of ordinary scientific methods, others' experiences being neither directly nor non-inferentially verifiable." (Güzeldere, 1997, p. 25) So, as far as third person perspective is concerned, both consciousness and quantum domain are inaccessible.

In the above, I briefly explored similarities and dissimilarities between the realm of phenomenal consciousness and quantum domain. Quantum domain will be further explored in chapter 5 where I probe Bohm's ideas of the relationship of mind and matter. In the next chapter, I will also explore the notion of intentionality giving a new natural informational account of the notion. Chapter 6 binds the realms of consciousness, intentionality, and quantum mechanics to one another via a new thesis.

# Chapter 4
# Intentionality

## 4.1 Introduction

This essay projects a thesis on consciousness dealt with which informationally. However, as stated in section 1.2, consciousness and intentionality go hand in hand: if pragmatic information provides an appropriate context for projecting a theory of consciousness, it must also be able to account for intentionality. In this chapter, the latter notion is examined in virtue of pragmatic information. Biological systems and artifacts have been considered the only intrinsically intentional systems because the interaction between their parts are information-driven.

However, biological systems are *original* intentional systems, while the intentionality of artifacts is *derived*. Original intentional systems are “teleonomical” systems: we do not presuppose any kind of ultimate purpose, but their existence according to the laws of nature. However, artifacts are produced by humans or other intelligent living systems (purposeful agents) with regard to a particular application or performance, planned in advance. They will not happen naturally or would happen just by chance. In this chapter, we will see how one can use the thesis of pragmatic information in giving an account for intentionality and how she can use the latter thesis to distinguish between original in derived intentionality.

## 4.2 Naturalizing Intentionality

The usage of pragmatic information as an appropriate universal notion of information also leads to a new notion of intentionality, the notion that can be classified as *naturalized intentionality*. To depict the idea, we have to start with the initiator of the concept in its contemporary use, namely, Frantz Brentano.

In his book, *Psychology from an Empirical Standpoint*, the following paragraphs can be considered the material for launching discussions on the nature of intentionality, discussions that extend to the present day.

> "Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not do so in the same way. In presentation, something is presented, in judgment something is affirmed or denied, in love loved, in hate hated, in desire desired and so on.
>
> This intentional inexistence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it. We can, therefore, define mental phenomena by saying that they are those phenomena which contain an object intentionally within themselves." (Brentano, 1874, pp. 88-89)

As seen, Brentano's account of intentionality is not so clear. What is quite clear, however, is that he characterizes intentionality exclusively as the mark of mental. This has been a thesis for a research program. As Jacob (2003) emphasizes, two other theses can be outlined: 1. the thesis that mental states, which exhibit the phenomenon of intentionality are directed towards things different from themselves and 2. the thesis that intentional inexistence is the property of the objects towards which the mind is directed by virtue of intentionality. In other words, objects have a property in common, that is, the property of making intentional mind be directed towards them.[81] The two latter theses straightforwardly make the projection of the following question felicitous: are there intentional objects?

Chisholm's (1957) answer to the above question is "no." To reject the existence of intentional objects, Quine (1960, p. 220) also take advantages of Chisholm's argument according to which the intentional vocabulary cannot be reduced to some non-intentional vocabulary concluding that intentionality is the mark of mental. From this conclusion, "Quine (1960, 221) presented an influential dilemma with both epistemological and ontological implications. The first horn of the dilemma is to accept the "indispensability of intentional idioms and the importance of an autonomous science of intention" and to reject a physicalist ontology. The second horn of the same dilemma is to accept physicalism and renounce the "baselessness" of the intentional idioms and the "emptiness" of a science of intention." (Jacob, 2003)

The linguistic view of intentionality advocated by Chisholm (1957) and Quine (1960), which results in the confirmation of Brentano's thesis that intentionality is the mark of mental, is questionable. By virtue of having meaning, however, sentences of natural languages, which are non-mental commodities, can be directed towards things other than themselves, just like intentional states of mind.

The above objection has been influentially discussed and responded by John Searle in his numerous works (Searle, 1980, 1983, 1992, and 1995).[82] Searle, at the first step, hints of explaining the intentionality of mind as similar to intentionality of language.

> "Notice also that we cannot explain the intentionality of the mind by saying it is just the intentionality of language. In the case of language, the utterance "Caesar crossed the Rubicon" is about Caesar and says of him that he crossed the Rubicon. I cannot say that a mental representation derives its intentional capacity from language, because of course the same problem arises for language. How is it possible that a mere sentence, sounds that come out of my mouth or marks that I write on paper, can refer to, be about, or describe objects and states of affairs that are 2,000 years in the past and 10,000 miles away? The

[81] As Jacob himself has noticed, the two theses can hardly be divorced.
[82] There are also similar accounts in this respect, e.g., Haugeland (1981) and Fodor (1987).

intentionality of language has to be explained in terms of the intentionality of the mind and not conversely." (Searle, 2004, p. 161)

Searle also distinguishes between original or intrinsic intentionality[83] on the one hand and derived intentionality on the other. When we have in our head information about how to get to San Jose, this information and the set of true beliefs about the way to San Jose count as intrinsic intentionality. Such information can also be found in the map in front of us containing of symbols and expressions that *refer to* or are *about* or *represent* cities, however, in a different sense. The intentionality containing in the map is derived from the intentionality of the map makers and users. The intentionality the map contains is imposed on it by the intrinsic intentionality of humans.

As will be discussed soon, the distinction between intrinsic and derived intentionality, however in a different way from what Searle has posed, plays an important role in pragmatic information as depicted in this essay. For the time being, another important distinction, according to Searle, should be mentioned, namely, the distinction between those features of the world that are observer independent and those that are observer dependent or observer relative. Force, mass, gravitational attraction, the planetary system are examples of observer-independent things. They exist regardless of the thought and the act of human beings. On the other hand, money, property, government, football games are things whose existence depends on human attitudes. These things are what they are because that's what we think they are. "Observer-dependent facts are created by conscious agents, but mental states of the conscious agents that create observer-dependent facts are themselves observer-independent mental states."(ibid, p. 6)

From the latter distinction, Searle moves to a proposal on the notion of "function." "Functions are never intrinsic to the physics of any phenomenon but are assigned from outside by conscious observers and users. *Functions, in short, are never intrinsic but are always observer relative*.

We are blinded to this fact by the practice, especially in biology, of talking of functions as if they were intrinsic to nature. But except for those parts of nature that are conscious, nature knows nothing of functions."(Searle, 1995, p. 14)

[83] Searle uses original intentionality and intrinsic intentionality interchangeably. My proposal, however, distinguishes between the two.

According to his view, when we say, "the function of the heart is to pump blood," we have not discovered any natural facts beyond the causal facts. We have just added to the vocabulary of "causes" a set of values.

As far as we are in the realm of natural non-living systems, I totally agree with Searle in that we assign or impose functions to the causal relations between systems' elements. In other words, systems with physical interactions between their parts are all causal systems with no purpose, meaning, or means for information generation included (recall the example given in section 2.5.2, namely, the moon of a planet in orbit around it). Only when biological systems interact with such systems, the interaction would count as information-driven. This occurs, for example, when we model physical systems. Roughly, "a model is an approximation, a mental construct in which a correspondence is established between a limited number of parameters that control the model and the values of certain physical magnitudes (degrees of freedom) pertinent to the system of study." (Roederer, 2005, p. 24) Through the process of learning, brains also build models that are approximation of reality outside. These models are amended in a way that errors can be corrected in later experiences.

In summary, in pragmatic information vocabulary, systems with physical interactions between their parts are just causal systems running independent of us. However, as stated in the last section, any interaction with these systems by biological systems makes the interaction become information-driven. Even "*when it is us who choose the initial state,* this physical correspondence would also represent an information-driven interaction: we set the initial state with the *purpose* to achieve a wanted final state, or with the purpose of finding out what the final state will be at time *t*. We have *used* the physical system for a mapping purpose; we can repeat the same process in the laboratory, in our mind or on a computer as often as we want, and for the same initial conditions (the same input pattern) we will always obtain the same final state […]." (ibid, p.170)

In the realm of *biological* systems, however, things are different. In this realm, as stated in section 2.5.2, interactions are information-driven in which a pattern as the sender with *purpose* triggers a change in the recipient such that such a change would not happen if that certain pattern were not present. According to this view, as far as information-driven interactions are concerned, there is no difference between a heart, which pumps blood and a conscious mind, as a biological system, which assigns functions to a purely physical interaction; both fulfill the required conditions of information-driven interactions. Recall the example I posed in section 2.5.2 and appendix 1 about how the heart interacts with the brain based on information-driven interactions. To show the similarity between the latter and the mechanisms in which some mental states and processes take place according to information-driven interaction framework, I restate the Roederer's (2005) hypothetical account on "perception," "short-term memory," "long-term memory," "memory recall," "imagination," "thinking," and "reasoning."

Suppose that some environmental signals in sensory system S, say in the retina of the eye, are converted into homologous patterns of neural activation (Fig. 4.1). Suppose also that S is a pattern that represents information in a nervous system in a dynamic form, given by the specific *spatio-temporal distribution of electrical impulses* in the neural network (analog postsynaptic potentials and standardized action potential). This is one way through which information is represented or encoded in the nervous system – to be called Type I of information representation in the neural network hereafter. Another way of information encoding is as what happens in patterns at higher levels A, B, C, …, that is, changes in the synaptic architecture during their use – to be called Type II of information representation hereafter. When pattern $P_A$ is activated by the input pattern $P_S$, a chain of information-driven interactions happens by the activation of patterns $P_B$, $P_C$, etc.

The persistence of these patterns for an interval of time longer than the actual duration of the input pattern $P_S$ is short-term, activity-dependent or working memory images of $P_S$. As seen in Fig. 4.1, there can also be information-driven feedback relationships between the consecutive neural levels A, B, C. When a higher-level pattern, say $P_B$, is evoked in some independent way at a later time, the original pattern $P_A$ to which it corresponds is, at least partially, elicited. In this case, we say that information in $P_A$, and also in $P_S$, is stored in the *long term or structural memory*. Roederer (2005, p. 141) calls this feedback process, through which an internal pattern (without any need to an external input $P_S$) can elicit corresponding patterns at lower levels, *memory recall*. Some evidence confirms this top-down controlling model. For example, Miyashita (2004) states:

"The types of memory identified in humans are extended into animals as episodic-like (event) memory or semantic-like (fact) memory. The unique configurational association between environmental stimuli and behavioral context, which is likely the basis of episodiclike memory, depends on neural circuits in the medial temporal lobe, whereas memory traces representing repeated associations, which is likely the basis of semantic-like memory, are consolidated in the domain-specific regions in the temporal cortex. These regions are reactivated during remembering and contribute to the contents of a memory. Two types of retrieval signal reach the cortical representations. One runs from the frontal cortex for active (or effortful) retrieval (top-down signal), and the other spreads backward from the medial temporal lobe for automatic retrieval. By sending the top-down signal to the temporal cortex, frontal regions manipulate and organize to-be-remembered information, devise strategies for retrieval, and also monitor the outcome, with dissociated frontal regions making functionally separate contributions." (Miyashita, 2004, 435)

Another experiment, confirming said top-down controlling model, has been conducted by Dierks et al. (1999) with paranoid schizophrenic patients whose on- and offset of auditory hallucinations could be monitored within functional magnetic resonance imaging (fMRI) session. They demonstrate an increase of the blood oxygen level-dependent (BOLD) signal in Heschel's gyrus during the patients' hallucinations. Their results provide direct evidence of the involvement of primary auditory areas in auditory verbal hallucinations.[84]

Memory recall plays a crucial role in some mental phenomena: thinking, reasoning, and imagination, for example, are the act of information recall, alteration, and restorage without any external input. Anticipation also comes about when "certain features of a frequently occurring sensory input trigger neural activity in relevant primary sensory areas of the cortex even *before* an actually occurring feature can elicit the corresponding response, if the expected feature is missing, the corresponding neural pattern appears anyway." (Roederer, 2005, p. 203) Note that this storage would have to be of the Type II

[84] There are some studies emphasizing the overlap in the neural substrates supporting visual perception and visual imagery. See for example: Ishai and Sagi (1995); Ishai et al.(2000); Kosslyn et al.(1999); Kreiman et al.(2000); Lambert et al.(2004); Mechelli et al.(2004); O'Craven and Kanwisher( 2000).

of representing information in the neural network, that is, a change in the synaptic architecture.

According to Roederer, the most important difference between animals and humans can be that only human can recall the past at will. Moreover, only human's brain can "recall stored information as images or representations, manipulate them, and restore modified or amended versions thereof without any concurrent external sensory input."[85] (Roederer, 2005, p. 220)

Getting back to Searle's idea and concerning the distinction between the concept of function, which can only be assigned to the purely physical interactions by a conscious being on the one hand and the concept of purpose, which is intrinsic to the biological systems, not observer relative, on the other, we can change the vocabulary as what follows: when we are in the realm of biological systems and the interactions in this realm satisfy the conditions of information-driven interactions, "purpose" as an intrinsic concept will be used. In the realm of physical interactions, however, the concept of "function" applies where a conscious biological system interacts with a purely physical system such that he assigns the causal interactions between the different parts of the system to do something teleological or with a certain goal to achieve. Once again, I have to emphasize that such assignment makes the system under consideration become informational. Notice that it is also possible to assign a function to a biological system. For example, when we say, "the function of the heart is to pump blood," this function has been assigned by us as conscious beings. It is not intrinsic to the heart.

This is why we can speak of malfunctioning hearts or heart diseases, as Searle has noticed. A heart is a biological system with information-driven interactions. As "life" and "survival" are values for us, pumping blood by heart counts as a function while a heart disease as a malfunction. "If we thought the most important value in the world was to glorify God by making thumping noises, then the function of the heart would be to make a thumping noise and the noisier heart would be a better heart." (Searle, 1995, p. 15)

[85] Note that the account of the mentioned mental processes and states by Roederer in no way can give any explanation of the subjective, intrinsic, and qualitative feel, namely, conscious experience, which might be associated with them. This criticism will be posed in detail in chapter 6.

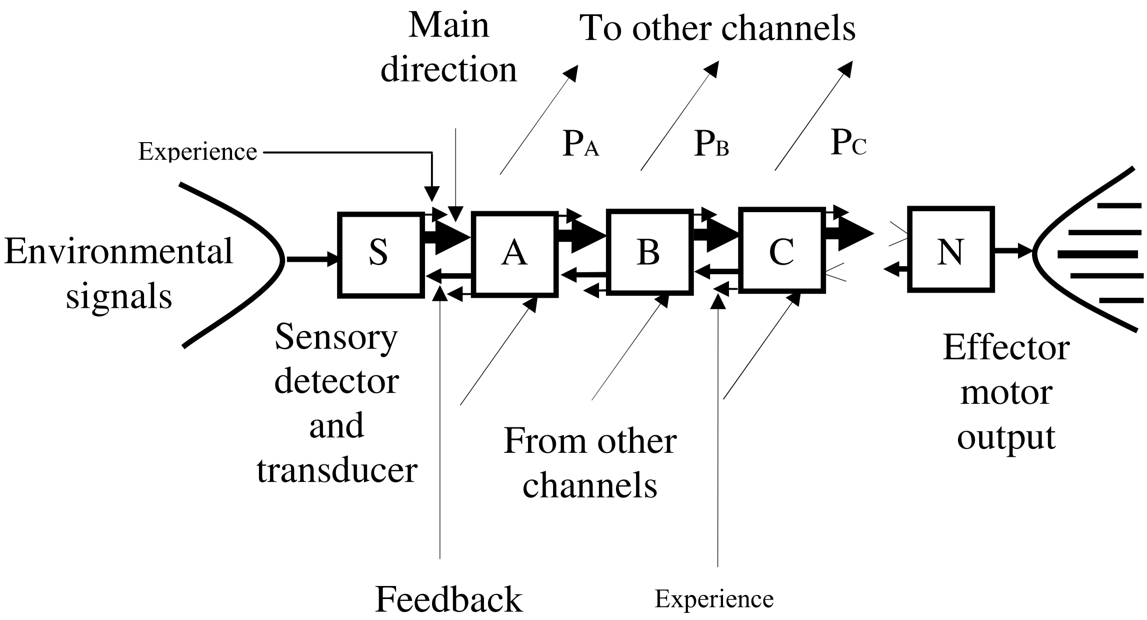


**Fig. 4.1.** Hypothetical set of sensory information-processing stages in a nervous system. The figure has a lot in common with what one finds in Roederer, 2005, p. 141.

Changing the common code between the sender and the receiver, for any reason, will result in changing the correspondence between them. If for any reason, the brain sends different pulses to the heart, different changes will occur and the heart will not function as it did with the first common code. Let me call the first common code, the right one. It leads to the right functioning of the heart, as we call it. Let me also call the second common code, the wrong one. It leads to a malfunctioning heart, as we humans call it. As far as the notions of "right" and "wrong" are concerned, it is us who attribute them to the heart. These terms have to do with the notions of functioning and malfunctioning hearts, respectively. However, both hearts are controlled by brain and in both, the interaction between the heart (and its parts) and brain (and its parts) is information-driven including all its elements, i.e., purpose, change, common code, correspondence, etc. As we saw, such elements in information-driven interactions have naturalistic objective definitions. Thus, speaking of, say, "mal-purpose" is senseless. Such an objective definition of the concept of purpose as depicted in pragmatic information does not include any value or teleological aspect. It differs from the notion of "purpose" as used in everyday life.

Where "purpose" comes to such an objective, non-teleological sense mentioned above, Küppers (1990) uses the term "purposiveness" to distinguish it from the notion of purpose, which in its ordinary use requires an intentional agent. According to him, "purposiveness" is the characteristic of all biological systems.

> "It would thus be absurd to deny the purposiveness of the natural organ while conceding the purpose of the artifact. On the contrary: this purposiveness is a characteristic of all organisms, and not just an isolated phenomenon based on a coincidence of specific structure and meaningful function."[86] (Küppers, 1990, p. 7)

In elucidating the above difference, Küppers refers to Colin Pittendrigh, who first introduced the term "teleonomy" in contrast to teleology. The fact of purposiveness of living systems in the former is denoted in a purely descriptive way without implying any explanation of the cause of the phenomenon. Teleology, however, interprets the purposiveness of the biological systems in terms of the existence of a final cause or purpose. "This conceptual clarification is important, as it allows a distinction between natural and artificial objects. The artifact is produced by humans with regard to a particular application or performance, planned in advance. Thus the final purpose determines the form of a man-made object. For natural objects, however, we do not presuppose any kind of ultimate purpose, at any rate not if we are prepared to accept what Monod called the "postulate of objectivity" as a basic postulate of scientific method: "Nature is *objective* and not *projective*.""[87] (ibid, p. 7)

Let us come back to the distinction between "function" and "purpose." In spite of what was mentioned above about the difference between biological systems and artifacts, as far as information-driven interactions are concerned, there is no difference between animate systems and artifacts as they fulfill the conditions of the latter interactions. Recall Bohm's example about the radio waves as broadcast from a station and received by our radios. The same distinction, drawn between the application of the notions "function" and "purpose" in living systems can be drawn in the radio. The desired function of a radio for us is to receive the desired program if it is tuned correctly with respect to the specific wavelength or frequency of the waves. If it is not so, the radio is said to be

[86] Küppers also states that "the phenomenon of purposiveness was dealt with by Kant in his *Kritik der teleogischen Urteilskraf* (1790). He was particularly occupied with the question of whether and, if so, to what extent it is possible to explain the obvious purposiveness of animate Nature by a final goal or cause." (Küppers, 1990, p. 178)
[87] Monod (1972, p. 15)

malfunctioning. However, there might still exist a correspondence between another certain pattern, as the sender, and a correspondent change elsewhere, at the recipient side.

The distinction between observer dependent and observer independent features of the world, here depicted with respect to the distinction between "function" and "purpose" respectively, leads us to another essential distinction. According to Searle, beliefs, fears, hopes, and desires, or in general genuine mental states, are intrinsically intentional. (Searle, 1983, p.27) Utterances borrow their derived intentionality, as Searle calls it, from the intrinsic intentionality of human beings (and some other animals) with minds.

To analyze such a distinction, first we have to notice that where we speak of "intentionality," we are speaking of a form of *representation*. Remember figure 4.1; there is an information-driven interaction between, for example, $P_A$ and $P_B$. Where $P_A$ is present, $P_B$ is triggered. $P_B$ represents the presence of $P_A$. Furthermore, the presence of $P_A$ has a meaning for $P_B$: $P_B$ will be triggered just if $P_A$ or a very similar pattern is present. All this is done *naturally* and *intrinsically*. The same goes for the radio waves and the radio with just one difference: the correspondence has been constructed artificially, not naturally. However, such a correspondence is still intrinsic. If the radio is "on" and it is working properly, it will receive the required program. Once we have determined the conditions according to which the radio is supposed to work, once we have built it with a certain purpose to achieve a certain task, the system will turn to an information-driven, which fulfills our purposes. As far as there exists a one-to-one correspondence between the pattern of the sender and a change in the receiver, the receiver is considered intrinsically intentional. In such an interpretation, the distinction between intrinsic and derived intentionality, as Searle projects, disappears. All information-driven interactions, both in artifacts and biological systems, contain intrinsic intentionality in the sense that the change in the receiver is elicited where a certain correspondent pattern is present as the sender.

So I will change Searle's vocabulary in a way consistent to the claim posed in this essay. To do so, I will first distinguish between the terms "original" and "intrinsic." *Original intentional systems* are systems occurring in the realm of biological systems with information-driven interactions between the parts. *Derived intentional systems* are artifacts the interaction between whose parts also fulfills the requirements of information-driven interactions. They are derived because an intelligent conscious being has designed them to achieve a certain goal. By providing the setup of the artifacts and building them with a predetermined plan, artifacts turn to informational systems whose intentionality shows the intentionality of the setup providers and builders. If the derived intentionality of artifacts were not put there by the designers, setup providers, or initial condition determiners, such systems could occur in nature in a repeatable way (not just by chance). As mentioned earlier, both original and derived intentionality, are intrinsic. Moreover, both are *relative*. By the term relative I mean that the receiver only represents the presence of a specific pattern as the sender.

This way construed, maps (or utterances) get their intentionality from the original intentionality of the map makers who created them with a certain purpose. However, the interaction with the user is intrinsic such that the pattern made of the ink stains on the sheets of cellulose fibers, as the sender, has a certain meaning for the user, as the recipient. There is a one-to-one correspondence between the mentioned pattern and the user.

It is now clear that my proposal on intentionality in this essay falls into the class of the proposals, which try to naturalize intentionality. The common strategy of such proposals is to show that Brentano was wrong in claiming that intentionality exclusively is the mark of mental. Among others, Dretske (1980, 1981, 1991, and 1994) is also a proponent of naturalized intentionality. In his book, *Explaining Behavior* (1991), Dretske introduces three types of conventional systems of representation. By a representational system (RS) he means “any system whose function it is to indicate how things stand with respect to some other object, condition, or magnitude.” (Dretske, 1991, p. 52)

Let us start with Type I of representational systems. Dretske considers the dime on his table as Oscar Robertson, the nickle (heads uppermost) as Kareem Abdul-Jabbar, and another nickle (tails uppermost) as the opposing center. Some pieces of popcorn are also considered other players and a glass as the basket. The stage is now ready to represent the positions and movements of the players by moving coins and popcorn around on the table. These objects can be used to describe a basketball play he has witnessed. (ibid, pp. 52-53)

In the above case, a temporary function has been assigned to the coins and popcorn. They have no intrinsic power to do what they have been assigned the function of doing. Their success in the performance of the job is derived from the interaction with the manipulator. The objects he has used to represent what happened in the basketball court are, considered by themselves, representationally lifeless. They are his representational instruments. (ibid, p. 53)

> "The elements of Type I systems have no *intrinsic* powers of representation – no power that is not derived from us, their creators and users. Both their function (what they, when suitably deployed, are *supposed* to indicate) and their power to perform that function (their success in indicating what it is their function to indicate) are derived from another source: human agents with communicative purposes. Many familiar RSs are like this: maps, diagrams, certain road signs (of the informational variety), prearranged signals, musical notation, gestures, codes, and (to some degree, at least) natural language." (ibid, p. 53)

It is obvious that the coins and popcorn have no intrinsic powers of representation of the basketball play. However, although the power of representation of the coins and popcorn is derived from human agents, the case is not similar to the other examples Dretske poses. Maps, diagrams, certain road signs, prearranged signals, musical notation, gestures, codes, and even natural language are already intentional, while the coins and popcorn are not (as far as representing the basketball play is concerned). The former examples are not representationally lifeless. Patterns of the former can immediately trigger the correspondent neural patterns of the user, but the coins and popcorn cannot do in advance the task Dretske defines for them.[88] So it seems that the notion of derived intentionality in Dretske's point of view is slightly different from Searle's. The case of the coins and popcorn is very similar to what was mentioned about modeling of a purely physical system discussed in page 130. They are not intentional at all until we model them with a

[88] Of course coins and popcorn will trigger the correspondent patterns in the brain as coins and popcorn. But concerning the task of representing a basketball play, they do not have such power in advance.

certain purpose. Maps, however, are intentional in advance though their intentionality is derived.

Note that, on the one hand, the way I use the term "meaning" in pragmatic information is very similar to Dretske's extension of Paul Grice's (1957) notion of *natural meaning* versus non-natural meaning in a language. He uses the verb "mean" as a synonym for "*indicate*." The English word "fire" non-naturally means fire, while smoke naturally means fire. According to Grice, if *P* is not the case, nothing can mean *P* in the natural sense. However, something (e.g., a statement) can non-naturally mean that *P* without *P*'s being the case. "The word "fire" can be tokened in the absence of any fire either for the purpose of expressing a thought about what to do if there was one or perhaps to mislead somebody else into falsely thinking that there is one. But there cannot be any smoke unless there is a fire." (Jacob, 2003)

On the other hand, there is a major difference between the naturalized notion of "meaning" in pragmatic information and such a notion in Dretske's point of view. "Meaning" in pragmatic information is meaning for the user (or informee or recipient). It is dependent on the informee. Dretske, however, claims that things have meaning, in the natural sense, independent of the users or makers.

> "Although a great deal of intelligent thought and purpose went into the design and manufacture of an ordinary bathroom scale, once the scale has been finished and placed into use there is nothing conventional, purposeful, or intelligent about its operation. This device indicates what it does without any cooperation or help from either its maker or its user. All you do is get *on* it. It then gives you the bad news. Somebody put the numbers on the dial, of course, and did so with certain intentions and purposes; but this is merely a convenience, something that (to use fashionable jargon) makes it user-friendly. It has nothing to do with what the instrument indicates. A clock doesn't stop keeping time if the numbers are removed from its face." (Dretske, 1991, p. 54)

Contrary to Dretske, I do not consider the bathroom scale as independent of the user and maker. Of course, like the example of the radio, once the scale has been finished and placed into use, it does not need the maker to give you the bad news. But the mechanism according to which the laws of physics in the scale give you the bad news is purposefully designed by the maker. The bathroom scale shows the intention of the maker in achieving a certain goal, that is, to give you the bad news. Furthermore, the circle is not complete unless a user gets on the scale and the pointer, as the sender, changes the cognitive state of her, as the recipient, with the bad news. If my dog gets on the scale, there is no news for it. In this interpretation, meaning, though naturalized, is meaning for us as users and recipients.

Before dealing with Type II of representational systems, some words should be said on the difference between symbols and signs according to Dretske. Recall the example of the coins and popcorn as the representational elements in representing the basketball play. Such elements are called *symbols* in Dretske's vocabulary. "Symbols are, either explicitly or implicitly, *assigned* indicators function, functions that they have no intrinsic power to perform. *We* give them their functions, and *we* (when it suits our purpose) see to it that they are *used* in accordance with this function. Such representational systems are, in this sense, *doubly* conventional: *we* give them a job to do, and then *we* do it for them." (ibid, p. 54) In contrast, natural signs are "events and conditions that derive their indicative powers, not (as in the case of symbols) from us, from our use of them to indicate, but from the way they are objectively related to the conditions they signify." (ibid, p. 54) In this sense, tracks in the snow, bird songs, and finger prints are natural signs. These signs do what they do independent of us.

In systems of Type II, natural signs are used partly to exploit their natural meaning and partly for conventional purposes. In such systems, we give a natural sign the job of doing what it can already do. The function of indication assigned to these systems depends on our interests, values, and purposes although they can typically indicate a great many things. The pointer in the bathroom scale also indicates how long the spring mounted in it is stretched. However, we give the bathroom scale the function of indicating our weights, which is mostly of our interest.

According to Dretske, the difference between Type I and Type II representational systems can be drawn as what follows: In systems of Type I, the function comes first. With systems of Type II, however, first comes the power of their elements to indicate something, and then comes their function.

"A natural system of representation is not only one in which the elements, like the elements of Type II systems, have a power to indicate that is independent of the interests, purposes, and capacities of any other system, but also one in which, in contrast with systems of Type II, the functions determining what these signs *represent* are also independent of such extrinsic factors. Natural systems of representation, systems of Type III, are ones which have their *own* intrinsic indicator functions, functions that derive from the way the indicators are developed and used *by the system of which they are a part*. In contrast with systems of Type I and II, these functions are not assigned. They do not depend of the way *others* may use or regard the indicator elements." (ibid, p. 62)

According to Dretske, the heart is already supposed to pump blood. Such a function is independent of us.

> "We are accustomed to hearing about biological functions for various bodily organs. The heart, the kidneys, and the pituitary gland, we are told, have functions – things they are, in this sense, *supposed to do*. The fact that these organs are supposed to do these things, the fact that they have these functions, is quite independent of what *we* think they are supposed to do. Biologists *discovered* these functions, they didn't invent or assign them. We cannot, by agreeing among ourselves, *change* the functions of these organs." (ibid, p. 63)

Once again, we notice that as far as the concept of function is concerned, Dretske's idea is in contrast with John Searle's. Representational systems of Type III are systems the interactions between whose parts are information-driven. Let me analyze one of the examples Dretske poses. Then it will become clear that what Dretske considers as "function," I consider as "purpose."

> "The firing of neural cells in the visual cortex, by indicating the presence and orientation of a certain energy gradient on the surface of the photoreceptors, indicates the whereabouts and the orientation of "edges" in the optical input and therefore indicates something about the surface in the environment from which light is being reflected. The activity of these cells, not to mention comparable activity by other cells in a wide variety of *sensory systems*, is *as much a natural sign or*

> *indicator as are the more* familiar events we commonly think of as signs – the autumnal change in maple leaves, growth rings in a tree, and tracks in the snow." (ibid, pp. 62-63)

Dretske poses such an example to support the claim that in biological realm, function is intrinsic, not observer dependent. To restate Dretske's example in pragmatic information vocabulary, one should say that firing of neural cells in the visual cortex takes place in the form of producing a specific spatio-temporal distribution of electrical impulses (Type I of information representation in the neural network, discussed in page 131. Such a pattern, as the recipient, represents a certain energy gradient on the surface of the photoreceptors, as the sender. There is a one-to-one correspondence between the pattern of the sender and a specific change in the recipient and the interaction is information-driven.[89] The sender has a certain purpose to elicit a change in the recipient. Such a purpose has a *meaning* for the recipient.

What Dretske considers as intrinsic function in the representational systems of Type III is what I have posed in the framework of pragmatic information in discussing Searle's notion of function as an observer-dependent notion. According to Dretske, representational systems of Type III, not just mental things, *are originally intentional*, the idea that has been advocated in this essay.

> "Only when we reach RSs of Type III – only when the functions defining what a system is supposed to indicate are intrinsic functions – do we find a source, not merely a reflection, of intentionality. Only here do we have systems sufficiently self-contained in their representational efforts to serve, in this one respect at least, as models of thought, belief, and judgment" (ibid, p. 67)

Any representational content must carry information about two things: the *reference* and the *sense*. By "reference" it is meant that any representation is representation *of* something. Representations are directed towards objects, persons, conditions, etc. By speaking of "sense," one is concerned with the way what is represented is represented, with what a representation says or *indicates*, or with what a representational content *means*. These two notions are also additional strands of intentionality.

[89] Note that there is a chain of intermediate information-driven interactions between mutually correspondent patterns in different stages. See figure 4.1.

According to Dretske, animals' (including humans) perception, recognition, knowledge, beliefs, and memory reflect the difference between a representation's reference and its sense. He believes that this "lends support to the idea that a cognitive system *is* a representational system of some kind, presumably a system of Type III." (ibid, p. 72)

Concerning the accepted and criticized parts of John Searle and Fred Dretsle's ideas on the notions of "function," "intrinsic intentionality," "original intentionality," "representational systems," and with respect to information-driven interactions in a way treated in this essay, we are now ready to sum up the claim of this essay on "intentionality" in virtue of pragmatic information. Every system with information-driven interactions between its parts is an intentional system. Whether the system is a mental system or not, whether the system is a biological system or an artifact, such intentionality is intrinsic to the system. For the change will occur in the recipient just where there is a specific pattern of the sender correspondent to it. The purpose of the sender has a unique meaning for the recipient in a natural way stated above. The change in the recipient, then, *represents* or *indicates* the existence of a certain pattern as the sender. Such a change is *about* or *directed towards* the pattern of the sender and will occur just where it is present.

On the other hand, every intentional system with information-driven interactions between its parts is either derived, in case of artifacts, or original, in case of biological systems. Artifacts borrow their intentionality from the intentionality of the makers, designers, setup providers, or initial condition determiners, who made them to achieve a certain goal, who determined the initial conditions such that the system can fulfill a specific purpose repeatedly.

# Chapter 5

# Bohmian Mechanics and Its Implications for Mind

# 5.1 Introduction

As Atmanspacher (2006) argues, as regards the relation between mind and matter, two general lines of ideas are discernable. The first considers such a relation in a direct way as illustrated below.

$$[\text{Mind}] \rightleftarrows [\text{Matter}]$$

**Figure 5.1.** Schematic illustration of the frameworks in which the relationship between mind and matter is considered in a direct way. (Atmanspacher, 2006)

Such a framework can yield both reductionist and non-reductionist pictures. Non-reductionist theses may lead to Cartesian dualism. But they may also lead to a sort of property dualism according to which "conscious experience involves properties of an individual that are not entailed by physical properties of that individual, although they may depend lawfully on those properties." (Chalmers, 1996, p. 125)[90] Property dualism is a framework in which the relation between mind and matter is conceived indirectly as shown below.

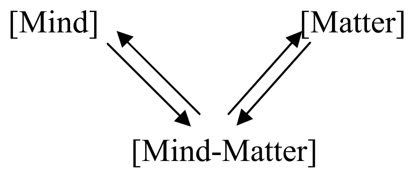


**Figure 5.2.** Schematic illustration of the frameworks in which the relationship between mind and matter is considered in an indirect way. (Atmanspacher, 2006)

In property dualism, there is no priority between mind and matter as regards the underlying unseparated "background reality" [Mind-Matter], as Atmanspacher (2006) calls it, though it is widely accepted that mind arises from material basis. Mind and matter are two aspects of such background reality that are correlated nomologically.

[90] Chalmers distinguishes between this sort of property dualism and a weaker one: "It is sometimes said that property dualism applies to any domain in which the properties are not themselves properties invoked by physics, or directly reducible to such properties. In this sense, even biological fitness is not a physical property. But this sort of "dualism" is a very weak variety. There is nothing *fundamentally* ontologically new about properties such as fitness, as they are still logically supervenient on microphysical properties. Property dualism of this variety is entirely compatible with materialism." (Chalmers, 1996, p. 125)

Among others, the thesis put forward by David Bohm (1980 and1990) counts as a sort of dual aspect theory on the relationship of mind and matter. According to this thesis, the whole universe is somehow enfolded in everything. On the other hand, each thing is enfolded in the whole. In other words, at the level of *implicate order*, each thing is internally related to everything else. But when it comes to the level of our ordinary experience, each thing *appears* to be relatively separate and extended, being related to other things only externally. The latter counts as *explicate order* according to Bohm. Explicate order involves classical (Newtonian) physics, as well as our experience. As far as the relationship between mind and matter is concerned, according to such an idea, implicate order unfolds into the explicate domains of mind and matter.

What makes the above thesis beneficial to be discussed in this essay is the use of the notion of "active information" as depicted by Bohm. The idea in this chapter is that the notion of active information at quantum level is very similar to the notion of pragmatic information as posed by Roederer as far as manmade experimental setup for preparing the needed states is concerned. It is also claimed that the extension of the thesis of active information to the natural inanimate world would count as a violation of the key constituent concepts of pragmatic information and as an ill generalization of Bohm's thesis itself.

Care should also be taken that even in an experimental arrangement, the notion of determinism at quantum level serves differently from what is conceived by the notion in classical world. The interactions in the former are nonlocal while spooky action at a distance in the latter is prohibited.

In what follows in this chapter, first I will briefly unfold some basic elements of Bohmian mechanics such as "field," "quantum potential," "nonlocality," "active information," etc. After examining the similarities and differences of the two notions of active information and pragmatic information, I shortly explain Bohm's ideas on the relationship of mind and matter according to the notions of "implicate" and "explicate" order.

## 5.2 Bohmian Mechanics

By means of Boltzmann's statistical mechanics, relations between quantities are identified without urging us to deal with the enormously complex, but nonetheless real, "hidden motion" of atoms and molecules. By analogy, one can think of quantum mechanics to provide us with statistical relations between local hidden variables associated with real individual, localized quantum particles moving in a real space-time frame. The term "local" in a local hidden variable theory means that distant events are assumed to have no *instantaneous* effect on local ones. A local hidden variable theory thus might solve problems associated with the collapse of the wave function, entanglement, and spooky action at a distance; problems that Copenhagen interpretation of quantum mechanics was faced with.

In 1964, John Bell wrote a paper entitled "On the Einstein-Podolsky-Rosen Paradox," in which he had presented an inequality, called Bell inequality. Such an inequality must be obeyed under any local hidden variable theory but can in certain circumstances be violated under conventional quantum mechanics. It placed restrictions on the statistical results of experiments on sets of entangled particles that have interacted and moved apart.

In the early 1980s, with collaborators in France, Alain Aspect performed the crucial "Bell test experiment" providing overwhelming support to the thesis that Bell's inequalities are violated. It showed that Albert Einstein, Boris Podolsky and Nathan Rosen's (EPR) *reductio ad absurdum* of orthodox quantum mechanics, namely that it implied "ghostly action at a distance," did in fact appear to be realized when two entangled particles were separated by an arbitrarily large distance. The Aspect experiment suffered from having some loopholes, which were resolved later.[91]

A completely deterministic, locally real version of quantum theory requires that the physical states of the particles be prescribed at the moment of their interaction and that

[91] There were two loopholes: "Locality loophole," which was resolved by Aspect, Dalibard, and Roger (1982) and Gregor Weihs et al. (1998) and "efficiency loophole," which was resolved by physicists from the national Institute of Standards and Technology in Boudler, Colorado, and the Department of Physics at the University of Michigan in (2001).

the particles separate as individually real entities in those physical states. These are the specifications of any local hidden variable theory. But as we saw, the experimental demonstration of quantum entanglement leaves out the idea of local hidden variables. So, either we have to consider the Copenhagen interpretation, as an anti-realist theory, to be complete[92] or in case of allegiance to realism, we have to construct a nonlocal hidden variable theory. In both cases, Einstein would be unhappy: The former contrasts with realism and the latter with special relativity.

Nonlocal hidden variable theories can be alternatives to the unfortunate local hidden variable ones. In such theories, the particles' motions might still be predetermined while there is spooky action at a distance. In 1926, Louis de Broglie suggested such an alternative. He considered the quantum entities, like electrons and photons, as independently real particles moving in a real field. A second field also exists which has the same statistical importance as Schrodinger's wave mechanics. So, equations of quantum mechanics admit double solution: a continuous wave field solution with statistical importance and a point-like solution which corresponds to a localized particle. The simplified version of de Broglie's interpretation of quantum mechanics is considered a construction involving a point particle guided by a continuous *pilot wave*. According to this view, the wave field guides the classical particle to follow a path in which the amplitude of the wave field is large. De Broglie explained his new interpretation of quantum mechanics at the 1927 Solvay Congress stating how this motion could account for quantum interference phenomena. However, he responded poorly to the objection posed by Wolfgang Pauli (Pauli, 1928) concerning inelastic scattering. Later, de Broglie's further discussions with members of the Copenhagen school led him to abandon his theory.

In the course of writing a book on the Bohr's notion of complementarity, entitled *Quantum Theory* (1951), the American physicist David Bohm presented a new version of the EPR thought experiment that eventually made him dissatisfied with the Copenhagen interpretation of quantum mechanics, the interpretation he had initially intended to advocate. Strongly encouraged by Einstein, Bohm initiated establishing a sort of hidden

[92] This means that, regarding orthodox interpretation of quantum mechanics, we have to close our eyes to the problems of interpretation of the meaning of quantum probability and the collapse of the wave function; problems that have been untouched in such a theory.

variable theory, which appeared through two papers in *Physical Review*. As Bohm's hidden variable theory owes a lot to de Broglie's theory, as far as the notion of pilot wave is concerned, his theory is often referred to as the de Broglie-Bohm theory.

Bohm's objective was to confront Copenhagen interpretation of quantum mechanics according to which quantum systems could not be described causally and objectively. He writes:

> "The usual interpretation of the quantum theory is self-consistent, but it involves an assumption that cannot be tested experimentally, *viz.*, that the most complete possible specification of an individual system is in terms of a wave function that determines only probable results of actual measurement processes. The only way of investigating the truth of this assumption is by trying to find some other interpretation of the quantum theory in terms of at present 'hidden' variables … the mere possibility of such an interpretation proves that it is not necessary for us to give up a precise, rational, and objective description of individual systems at a quantum level accuracy." (Bohm, 1952)

The above quotation shows that Bohm believes that other descriptions of quantum theory, especially causal ones as he wished, are possible *in principle*.

In Bohmian mechanics, particles, say electrons, are regarded as inseparable union of a particle and a field. So, his theory is regarded as a holistic one. The field he speaks of has some new properties that make it different from what is conceived by the term in classical mechanics. In physics, fields can generally be represented mathematically. Such certain mathematical expressions are called potentials. A potential in physics describes a field with the possibility or potentiality that is present at each point of space to give rise to act on a particle at that point.

However, one must distinguish between the use of the term in classical and quantum domain. "What is crucial in classical (-Newtonian) physics is then that the effect of this potential on a particle is always proportional to the intensity of the field. One can picture this by thinking of the effect of water waves on a bobbing cork, which gets weaker and weaker as the waves spread out. As with electric and magnetic fields, the quantum field can also be represented in terms of a potential which I call the quantum potential. But unlike what happens with electric and magnetic potentials, the quantum potential depends only on the form, and not on the intensity of the quantum field. Therefore, even a very weak quantum field can strongly affect the particle. It is as if we had a water wave that could cause a cork to bob up with full energy, even far from the source of the wave. Such a notion is clearly fundamentally different from the older Newtonian ideas. For it implies that even distant features of the environment can strongly affect the particle." (Bohm, 1990)

So, at quantum domain, this is the form, or pattern, of a field, not its energy, that guides the particle. Such a field is regarded as containing objective and active information. Such information is potentially active everywhere, but actually active only where the particle is. Active information also implies the possibility of a certain sort of wholeness of the particle with distant features of its environment. According to Bohm, when particles interact, it is as if they are all connected to each other by invisible links consisting a single whole. As regards the thesis of wholeness, Bohm sees a commonality with the Copenhagen interpretation. However, when it comes to the notion of realism, Bohm separates his way from Bohr:

> "This is in certain ways similar to Bohr's notion of wholeness, but it is different ill that it can be understood in terms of the concept of a particle whose motion is guided by active information. On the other hand, in Bohr's approach, there is no corresponding way to make such wholeness intelligible." (ibid)

Furthermore, Bohm's point of view on the notion of "measurement" has both commonalities with and differences from Copenhagen school. In Bohm's opinion, the particle's position and its trajectory are predetermined at all time during its motion. So, it is not necessary to resort to possibilities. But when we consider a large number of particles which can be described with the same wave function, because of our inaccessibility to, or our ignorance of, the particles' initial conditions, we calculate probabilities as a *practical* necessity.

This is in strong contrast with the notion of probability in orthodox interpretation of quantum mechanics. In the conventional quantum theory, the wave function is just a mathematical tool for the calculation of the relative frequencies of possible outcomes of repeated measurements on an ensemble of systems that have been prepared identically. The outcomes are determined only when a measurement is made.

Bohmian mechanics reverts Heisenberg's uncertainty principle to a "disturbance" picture. Measurement has significant impacts on the wave function (and hence on the quantum potential), leading to the disturbance (and hence the change) of the precisely predetermined position and momentum of the particle. So, as regards the central role of the measuring device, Bohm and Bohr share the ideas: in Bohm's thesis, changing the measuring device leads to the change of quantum potential which in turn leads to the new trajectories determined by the new experimental setup. In Bohr's thesis, the measuring device affects the system in a way that an outcome, out of possible outcomes, is realized. In both theses, the observed system and the observing apparatus are inseparable. Both adhere to a sort of wholeness. Notice that the observing apparatus above includes human beings, or perhaps other biological systems, as the ending part of the chain of measurement.

The notion of active information also implies another notion to come into the play, i.e., non-locality.

> "[T]he quantum potential depends only on the form of the wave, so that it can be strong even when the wave intensity is weak. Or to put it differently, what is basically new here is the feature that we have called non-locality, i.e. the ability for distant parts of the environment (such as the slit system) to affect the motion of the particle in a significant way (in this case through its effect on the quantum field)." (ibid)

Concerning said elements of the Bohm's nonlocal hidden variable theory, the following points are mentionable:

In the above theory, as the wave and the particle are not linked energywise, the interaction between the pilot wave and the particle can be a good candidate for a sort of information-driven interaction. In other words, what Bohm has in mind as "active information" is quite close to the notion of pragmatic information as Roederer projects. What counts in active information is the form or pattern of the quantum potential and any change elsewhere as the effect of the presence of this pattern is correspondent to such a pattern. Recall what was said about the way measurement affects quantum potential. Alteration of the form of the quantum potential changes the motion of the particle in a way that a correspondence can be considered between the form of the quantum potential and the motion of the particle.

What about other elements of pragmatic information in Bohm's account of active information, such as purpose? Remember Bohm's examples quoted at the beginning of section 2.5.2. He proposes the examples as analogies to the notion of active information in our classical world. The first two examples show active information in artifacts and the third in biological systems. It is not so hard to find the elements of pragmatic information in the two examples other than the one I have explained in that section. Then, Bohm extends the notion of active information to matter at quantum level:

> "Our proposal is then to extend this notion of active information to matter at the quantum level. The information in the quantum level is potentially active everywhere, but actually active only where the particle is (as, for example, the radio wave is active where the receiver is). Such a notion suggests, however, that the electron may be much more complex than we thought (having a structure of a complexity that is perhaps comparable, for example, to that of a simple guidance mechanism such as an automatic pilot). This suggestion goes against the whole tradition of physics over the past few centuries which is committed to the assumption that as we analyze matter into smaller and smaller parts, their behaviour grows simpler and simpler. Yet, assumptions of this kind need not always be correct. Thus, for example, large crowds of human beings can often exhibit a much simpler behaviour than that of the individuals who make it up."(ibid)

To analyze the above quotation, I draw the reader's attention to the three ways of establishing a one-to-one correspondence between the sender and the recipient, leading to the same change at the recipient side each time the sender is present, posed in section 2.5.2, pages 85 and 86.

The first was to design purposefully a resetting device to guarantee such a correspondence. At quantum level, it is done by providing appropriate experimental setup by a biological system or a program written by a programmer with the purpose of fulfilling a certain task. It is clear that the established correspondence between any two components in a quantum system, say between the wave and the particle, results from the intention of such a biological system and the way she arranges the experiment. When we make measurement, again we affect the system with a certain purpose.

> "And as has been pointed out, a different slit system would produce a different quantum potential which would affect the motion of the particles in a different way. Therefore the motion of the particles cannot be discussed in abstraction from the total experimental arrangement." (Bohm, 1993, p.38)

As seen above, also in Bohm's opinion, the correspondence between the quantum potential and particle's motion is just a function of the way we organize experimental setup.

The second way was the direct intervention of a purposeful, intelligent biological system in establishing the correspondence between the sender and the receiver. As we do not have direct access to the individual quantum systems, such a goal must be reached through the first way discussed above.

The third way was by chance in any natural system. As I have argued, in this case any purpose will disappear. Here, Bohm's ideas on active information do not match pragmatic information as depicted in this essay. Extending the notion of active information to all interactions at quantum level, regardless of whether they are done by chance or by the intervention of a living system in the laboratory, is a sort of ill generalization.

To sum up, when quantum particles are prepared by a purposeful living system in the laboratory to fulfill a certain goal, Bohm's description of active information at quantum level can count as a good instance of the existence of information-driven interactions in such a domain. This can be seen when he describes the two-slit interference experiment according to the notion of active information (see Bohm 1990 & 1993). But the extension of the idea to the natural inanimate world would count as a violation of the key constituent concepts of pragmatic information such as purpose.

Furthermore, even in the laboratory, the way the particle interacts with the quantum potential according to the notion of active information is different from similar interactions in classical world, such as the information-driven interaction between radio waves and as the sender and radio as the receiver. At quantum level, "the wave in de Broglie-Bohm theory can exert a strong influence on the particle through the form of the quantum potential, but there is no reciprocal reaction of the particle on the wave, seemingly at odds with classical mechanics in the form of Newton's third law of motion." (Baggott, 2004, p. 217) So, it seems that while the interaction between the wave and the particle, say an electron, is predetermined and information-driven, according to Bohmian mechanics, their interaction does not exactly match the picture we have of the interactions in the macroscopic world. The electron moves under its own energy. Although the wave containing active information interacts with the particle fulfilling the conditions of information-driven interactions, there is no energy and matter transfer between interacting components. In other words, they interact *non-mechanically*.[93]

> "We therefore emphasize that the quantum field is not pushing or pulling the particle mechanically, […]." (Bohm, 1993, p. 37)

Nonlocality is another notion that makes deterministic interactions at quantum level, as Bohm depicts them, differ from similar interactions in classical domain. In the macroscopic world, action-at-a-distance is prohibited, but as quantum potential is not so dependent on the intensity of the quantum field, but on the form, it can instantaneously affect distant features of the Universe. As emphasized in section 3.3, individual quantum world is inaccessible to our classical senses. So, any description at this level would be *our* description. Nonlocal interactions play a crucial role in our description of individual quantum phenomena, regardless of whether we are faithful to the anti-realist orthodox quantum theory or Bohm's causal hidden variable theory. Inaccessibility and nonlocality at quantum domain make the notion of determinism differ from what is conceived by the

[93] As stated in chapter 1, non-mechanical interactions can also be characterized by not being reducible to any of the four fundamental interactions known in physics.

notion at classical domain. Once again I have to notice the reader that inaccessibility at quantum level does not mean that there is nothing real out there.

## 5.3 Implicate and Explicate Order: Implications to Mind

As said in the introduction section in this chapter, as regards the interdependence of things to each other, Bohm distinguishes two levels of implicate and explicate order. In his opinion, everything in some way implicates or enfolds everything. This enfoldment relationship between things is active and is not superficial at all. Things appear to be independent of each other just under typical conditions of our ordinary experience. At this stage, we are in the realm of explicate order. The claim thus is that the relative independence of things is just a sort of appearance. “The explicate order, which dominates ordinary experience as well as classical (Newtonian) physics, thus appears to stand by itself. But actually, it cannot be understood properly apart from its ground in the primary reality of the implicate order.” (Bohm, 1990)

The general implicate process of ordering in Bohm’s thesis can apply to both mind and matter. This suggests that mind and matter are at least closely analogous, if not two aspects of one thing. They are not so different as they appear at the level of explicate order. Further development of this idea leads to the consideration of the notion of implicate order to serve as a means of expressing the actual relationship between mind and matter. We already saw how wave and particle are interrelated and considered a single whole in Bohm’s quantum theory. The role of active information in unifying the wave-particle duality was also clarified. “Indeed, going further, the whole notion of active information suggests a rudimentary mind-like behaviour of matter, for an essential quality of mind is just the activity of form, rather than of substance. Thus, for example, when we read a printed page, we do not assimilate the substance of the paper, but only the forms of the letters, and it is these forms which give rise to an information content in the reader which is manifested actively in his or her subsequent activities. A similar mind-like quality of matter reveals itself strongly at the quantum level, in the sense that the form of the wave function manifests itself in the movements of the particles. This quality does not, however, appear to a significant extent at the level at which classical physics is a valid approximation.” (ibid)

In the above view, this is active information that is simultaneously physical and mental. In other words, active information serves as a sort of bridge between mind and matter as two sides of reality, two inseparable sides that are aspects of a single whole. This means that what is felt to be the information contained in thought as the mental side of active information is at the same time the information contained in the physical side. The latter is the information related which gives rise to neurophysiological, chemical, and physical activity. “This means that that which we experience as mind, in its movement through various levels of subtlety, will, in a natural way ultimately move the body by reaching the level of the quantum potential and of the 'dance' of the particles. There is no unbridgeable gap of barrier between any of these levels. Rather, at each stage some kind of information is the bridge. This implies, that the quantum potential acting on atomic particles, for example, represents only one stage in the process.”(ibid)

In the next chapter, which will contain the core idea of this essay, a very similar idea to Bohm’s claim on the relationship of mind and matter will be posed. However, my proposal on the mentioned relationship will differ from Bohm’s as far as the generalization of the idea to the natural nonliving systems is concerned.

# Chapter 6
# Double Aspect Theory, Pragmatic Information: A Seemingly Legitimate Combination

## 6.1 Introduction

As stated in chapter 1, consciousness is a property that is ontologically independent of physical properties, but arises from a physical substrate in virtue of certain contingent laws of nature. Such an idea, which considers consciousness as a feature ontologically over and above physical features of the world, requires the construction of a new fundamental psychophysical law specifying the dependence of phenomenal to physical properties.

Chalmers (1996) suggests a few connections between consciousness and physical processes that deserve to be called psychophysical laws. One is a set consisting of two principles of coherence. The first coherence principle states that for each experience, there exists a corresponding cognitive state. This principle connects consciousness to awareness, leading to the idea that where there is consciousness, there is awareness. Note that as stated in chapter 1, Chalmers distinguishes between consciousness and awareness. In his opinion, "awareness is the psychological correlate of consciousness, roughly explicable as a state wherein some information is directly accessible and available for the deliberate control of behavior and for verbal report,"(Chalmers, 1996, p. 220) while phenomenal consciousness is a different concept, having to do with the immediate and direct knowledge of an experience. It should also be mentioned that one can also say that when there is awareness, there is consciousness, if we use the narrower sense of the term awareness, that is, if we take non-occurrent thoughts as not being qualified as part of the contents of awareness.

> "[N]on-occurrent thoughts do not qualify as part of the contents of awareness, but occurrent thoughts do. Correspondingly, we should expect that occurrent thoughts will be associated with experiences, even if non-occurrent thoughts are not. This is just what we find. My non-occurrent thought that Clinton is president has no impact on my phenomenology, but an occurrent thought to that effect will be associated with an experience. To see this, note that there is *something* it is like to think to oneself that Clinton is president; if I had not been thinking that thought just now, it would have been like something subtly different to be me." (Chalmers, 1996, p. 222)

According to this principle, for example, "what is experienced in audition is represented in our auditory system, in such a way that later processes have access to it in the control of behavior; in particular, the contents are available for verbal report." (ibid, p. 221) If there were no cognitive states corresponding to the experiences, then the content of the experience could not be reflected in behavior at all. The principle does not claim that whenever we have a conscious experience, we are aware of that experience. The idea is that when we have an experience, we are aware of the contents of that experience.

The second coherence principle is called "the principle of structural coherence." According to this principle, "the structure of consciousness is mirrored by the structure of awareness, and the structure of awareness is mirrored by the structure of consciousness." (ibid, p. 225)

The second principle that can count as a psychophysical law is called the "principle of organizational invariance" according to which "any two systems with the same fine-grained *functional organization* will have qualitatively identical experiences. If the causal patterns of neural organization were duplicated in silicon, for example, with a silicon chip for every neuron and the same patterns of interaction, then the same experiences would arise. According to this principle, what matters for the emergence of experience is not the specific physical makeup of a system, but the abstract pattern of causal interaction between its components." (Chalmers, 1995) The principle of organizational invariance will be discussed in more detail later in this chapter.

Chalmers does not see the above principles to be fundamental laws in a theory of consciousness. At best, they can be components of a final theory and can constrain the form of a final theory of consciousness. "The concept of awareness (or global availability) is a high-level concept, for example, and its boundaries are somewhat vague; it is very unlikely that this concept would be involved in a fundamental law. The principle of organizational invariance may be less vague, but it still expresses a regularity at a level that is far from fundamental. Another problem: these principles grossly underdetermine the nature of the psychophysical connection. All sorts of questions about connection remain unanswered." (Chalmers, 1996, p. 226)

So, Chalmers seeks for precise fundamental laws analogous to fundamental laws in physics, which when combined with the physical facts about the system, can enable us to predict the phenomenal facts about that system. In search of finding such a fundamental law, he suggests Double-Aspect Theory of Information according to which "whenever we find an information space realized phenomenally, we find the same information space realized physically. And when an experience realizes an information state, the same information state is realized in experience's physical substrate."(ibid, p. 284).

In his double-aspect theory, the concept of information Chalmers adopts has much in common with what Shannon (1949) has in mind. Such a concept was dealt with in chapter 2 where Shannon's conception of information was criticized as not having to do with information at all but with data transmission.

In this chapter, first I explain Chalmers' double-aspect theory of information in more detail while criticizing the syntactic conception of information as not being an appropriate candidate for double-aspect theory. The claim thus will be that the concept of information that best fits double-aspect is pragmatic information as put forward by Roederer.

However, as Roederer's theory of information is totally physicalist, it will not be able to give a complete account of the phenomenal consciousness associated with the mental phenomena, because it neglects the subjective character of consciousness. To change the theory in a way that it comprises the subjective qualities of conscious states, we will consider information as clarified by Roederer to have another aspect, namely, subjective or phenomenal or non-mechanical aspect (in the Bohmian sense). In this way, when information is processed physically (or mechanically), another aspect may be realized, i.e., phenomenal or non-mechanical. By such consideration, we have indeed combined double-aspect theory of information with pragmatic information. Such a combination modifies both theses and leads to a stronger explanatory power for phenomenal consciousness, while escaping from the subsequent ambiguities associated with the theses of pragmatic information and double-aspect theory. Non-mechanical informational interactions at quantum level as described by David Bohm are considered the evidence of the existence of phenomenally realized information.

## 6.2 Chalmers' Double-Aspect Theory of Information

In the last section, double-aspect theory was shortly explained. According to this view, a phenomenally realized information space is also realized physically. The notion of information used here by Chalmers is a syntactic or formal notion of it as depicted by Shannon (1949) in which, according to Chalmers, the key is the concept of a "state" selected from an ensemble of possibilities. Chalmers calls such possible states as *information states*, which comprise an abstract space called an *information space*. There is a basic structure of *difference relations* between such states. Consider our pinball machine with two equiprobable states; |0> and |1> are two states which comprise a two-state information space. According to Chalmers, the nature of these two states is exhausted by the fact that they are different from each other.

"In the most general case, an information space will have two sorts of structure: each complex state might have an internal structure, and each element in this state will belong to a subspace with a topological difference structure of its own. We might call these first of these the *combinatorial* structure of the space, and the second of these the *relational* structure of the subspaces. Much of the time, each subspace will have the same relational structure, so we can just speak of the relational structure of the space itself. The *overall* structure of the space is given by these combinatorial and relational structures together." (Chalmers, 1996, p. 279)

As said in chapter 2, Shannon's mathematical theory of information has to do with amounts of information. But Chalmers is not concerned with such amounts. Rather, he is concerned with the states themselves. The relation between the states and the amount of information in Chalmers' point of view is analogous to the relation between matter and mass.

I think such an analogy is wrong. Matter is out there in the real world. It is the substance, of which physical objects are composed. It constitutes much of the observable universe. The notion of matter per se has some explanatory power compatible with the laws of physics, capable of standing independent of the notion of mass.

Perhaps the above analogy stems from the conventional definition of matter as anything that occupies space and has mass. Compatible with such a definition, a more recent definition of matter that at least some physicists use is that matter is everything that is constituted of elementary fermions.

However, things may have mass without being matter: W and Z bosons have mass, but are not elementary fermions. Any two photons, which are not moving parallel to each other, taken as a system, have an invariant mass. Glueballs have mass due to their binding energy, but contain no particle with mass, nor any elementary fermions.

There may also be matter without having mass: most of the mass of protons and neutrons comes from the binding energy between the quarks, not the masses of the quarks themselves. One of the three types of neutrinos may be massless.

What the mathematical theory of communication (MTC) is concerned with is how probabilities are distributed among the possible states, not what possible states are and how they are related or structured. So, Chalmers, who "seems to think that the information theoretic notion of information is a matter of what possible states there are, and how they are related or structured […] rather than of how probabilities are distributed among them (Harms, 1998, p. 480)," has been confused. States in MTC are our abstractions to know the probability distribution among them in order to *calculate* the average informativeness in data transmission. In MTC, using the notion of state without having to do with amounts of information is senseless. The relation between states and amounts of information in MTC is not like the relation between matter and mass.

In short, Chalmers is either faithful to MTC or wants to use the notion of "state" in a new theory of information other than MTC. In case that the former holds, he should use such a notion just for quantitative purposes as the theory serves just as a mathematical tool for calculation of average informativeness in data transmission and optimization of such transmission. If one leaves out the computational power of MTC, there remains nothing of the theory. MTC has no explanatory power as far as the notion of information is concerned.

In case that the latter holds, Chalmers should introduce a new notion of information. What is Chalmers' notion of information in which what counts are the possible states themselves and the way they are related to each other and structured? How does he deal with the elements of MTC without going through quantification? In what follows, I will explore the possible answers that Chalmers may give to the above questions.

Explaining physically realized information, Chalmers states:

> "The natural way to make the connection between physical systems and information states is to see physically realized information in terms of a slogan due to Bateson (1972): information is a *difference that makes a difference*. While my light-switch can take on an infinite number of positions in a continuous range, most of this variation makes no difference at all to my light. Whether the switch is all the way up, or one-quarter of the way down, the light will be on. When it is in positions more than about one-third of the way down, on the other hand, the light will be off. As far as the light is concerned, there are only two relevant states of the switch, which we can call "up" and "down." The difference between these two states is the only difference that makes a difference to the light. So we can see the switch as realizing a two-state information space, with some physical states of the switch corresponding to one information state and with some corresponding to the other." (Chalmers, 1996, p. 281)

An information space correspondent to a physical object is always defined with respect to a causal pathway and also to an information space consisting of possible effects at the end of the pathway. In the above example, the causal pathway is what connects light-switch to light and the information space of possible effects consists of on/off state of the light.

The same goes for continuous information spaces. Suppose that the light in the above example has a dimmer switch. Rotating the knob to different positions in this case produces different intensities of light in a continuous range.

Concerning Chalmers' account of physically realized information, some objections may raise which are explored in what follows.

As stated in chapter 2 and earlier in this chapter, Shannon's theory (1949) mainly focuses on communications, control systems, computers, and the like, defining a mathematical measure of the *amount* of information contained in a given message, regardless of its meaning, purpose and the means used to generate it. Shannon declares:

> "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design." (Shannon, 1948, P. 379)

As seen in the above quote, the theory is mainly constructed to solve the engineering problems, which have to do with computational issues in artifacts mentioned above. At best, MTC can have computational applications in the study of biological systems, as for example MacKay (2005) uses the theory:

> "In the 1960s, a single field, cybernetics, was populated by information theorists, computer scientists, and neuroscientists, all studying common problems. Information theory and machine learning still belong together. Brains are the ultimate compression and communication systems. And the state-of-the-art algorithms for both data compression and error-correcting codes use the same tools as machine learning." (MacKay, 2003, p. v)

Now, suppose four stones on different parts of a hill as illustrated in figure 6.1. An earthquake causes stone 1 to hit stone 4. There were some other possibilities: stone 1 could have hit stone 3 or stone 2 could have moved hitting stone 3 or 4. Can one say that when stone 1 hits stone 4, we have physically realized information? In case that the answer is "yes," does the informational account in the form of "the difference that makes a difference" principle give any further explanation than causal explanation? What would Shannon have said if he had wanted to analyze such a case?

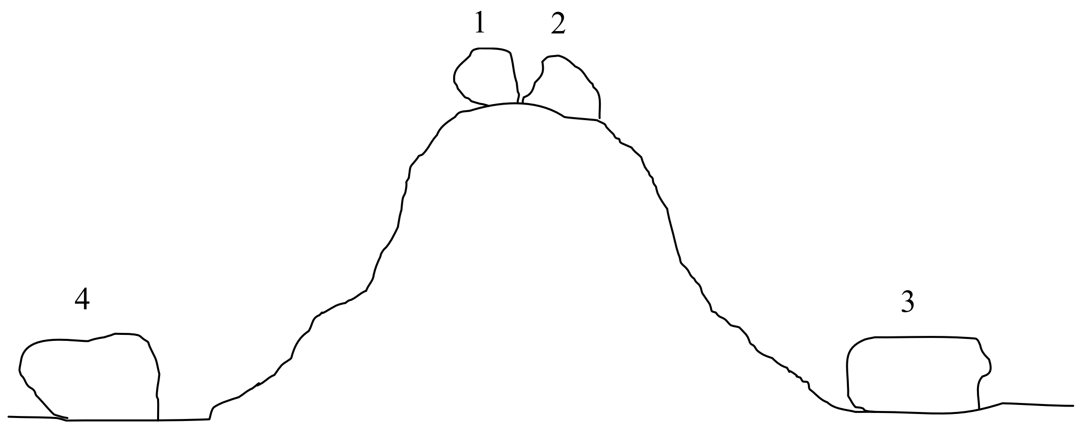


**Figure 6.1.** Four stones on different parts of a hill. The informational explanation of this case will add nothing more than causal explanation. See text for more detail.

It seems that Chalmers does not see any difference between the switch-light example and stone-hill one. Both function according to "the difference that makes a difference" principle. But Shannon would distinguish between the two examples. In the switch-light example, Shannon could apply MTC for mathematical calculations, because aside from its mathematical power, MTC has nothing to tell us. For Shannon, all matters is the mathematical treatment of the actual message selected from a set of possible messages. The theory does not have any explanatory power. Causal explanation of MTC adds nothing to our knowledge. The way Chalmers deals with the example will not attract Shannon's attention. Thus, if we are not willing to do any calculations in the stone-hill example, we will have nothing to say according to MTC or Bateson's slogan. At best, we can give some causal explanation of why stone 1 hit stone 4. It seems that Chalmers himself is aware of such a lack in his accounts:

> "Another sort of underspecificity stems from the looseness of the definition of physically realized information: for a fully specific psychophysical theory, we will need to know precisely what it is for an information space to be physically realized. But all this is part of the process of developing a theory." (Chalmers, 1996, p. 286)

To summarize, the claim here is that MTC in the form of "the difference that makes a difference" is a theory for *quantitative* treatments in the realms of artifacts such as computers and control systems as well as biological systems. Any use of such a theory in explaining any phenomenon will not go further than the causal explanation.

The way Chalmers treats with phenomenally realized information is similar with what was said about physically realized information. According to Chalmers, "states of experience fall directly into information spaces in a natural way. There are natural patterns of similarity and difference between phenomenal states, and these patterns yield the difference structure of an information space. Thus, we can see phenomenal states as realizing information states within those spaces.

For example, the space of simple color experiences has a three-dimensional relational structure that we have already discussed. Abstracting the patterns of similarity and difference among these experiences, we obtain an abstract information space with a three dimensional relational structure, which the phenomenal space realizes. Any given simple color experience corresponds to a specific location within this space. A specific red experience is one phenomenally realized information state; a specific green experience is another." (ibid, p. 284)

However, information spaces realized phenomenally do not function based on causal "difference that makes a difference" principle used in physically realized information spaces. In the former, Chalmers relies on the intrinsic qualities of experiences and structures among them. "Any experience will bear natural relations of similarity and difference with other experiences, so we will always be able to find information spaces into which experiences fall." (ibid)

Let us now deal with the double-aspect principle according to which information has two aspects, that is, a physical and a phenomenal aspect. "Wherever there is a phenomenal state, it realizes an information state, an information state that is also realized in the cognitive system of the brain. Conversely, for at least some physically realized information spaces, whenever an information state in that space is realized physically, it is also realized phenomenally." (ibid, p. 286)

How does double-aspect principle work? Chalmers explains that "information spaces required by physics are themselves grounded in phenomenal or protophenomenal properties. Each instantiation of such an information space is in fact a *phenomenal* (or protophenomenal) realization. Every time a feature such as mass and charge is realized, there is an intrinsic property behind it: a phenomenal or protophenomenal property, or a *microphenomenal* property for short. We will have a set of basic microphenomenal spaces, one for each fundamental physical property, and it is these spaces that will ground the information spaces that physics requires. The ultimate differences are these microphenomenal differences." (ibid, p. 305)

Now, after reviewing Chalmers' double-aspect, mostly through his own words, the following questions may seem reasonable: let us change double-aspect theory of information to "double-aspect theory of causality" according to which cause-and-effect relationships have two aspects, a physical and a phenomenal aspect. Do we lose anything in such a change? Will there be any lack of explanation compared with the original double-aspect? If the account of physically realized information does not go further than the causal account of "the difference that makes the difference," why should we use the notion of information? The latter is legitimate, I think, when the notion of information has had more explanatory power than the notion of causality. Altering the title does not add any new knowledge to what we have, concerning causal relations.

Chalmers' double-aspect principle has some consequences. As stated above, according to him, we have information wherever we have causation. Causation is found everywhere, so information is also everywhere. In this respect, Chalmers encounters the issue dually: on the one hand, he suggests looking for constraints, not letting any physically realized information space be associated with experience. On the other hand, he bites the bullet and accepts that experience is ubiquitous; even a rock has experiences. The former shows that Chalmers sees the ubiquity of information counterintuitive, if not a problem associated with his double-aspect principle. But according to the latter, there can be something it is like to be, say, a thermostat: the only difference between a thermostat's experiences and, say, a mouse's experiences lies in the degree of complexity. Chalmers seeks for ways to make such an idea less crazy.

So, to make the idea of thermostats, as information-processing systems and entities with phenomenal experience, less crazy, Chalmers gradually moves down the scale of complexity. He begins with humans as complex information-processing systems and moves to less complex systems such as dogs, mice, lizards, and fish to slugs. He sees no reason to consider consciousness winking out altogether as he moves down the scale, although the experiences would be less complex. Then he proceeds with the case of thermostats as the extension to the simpler and simpler systems and asks:

> "Someone who finds it "crazy" to suppose that a thermostat might have experiences at least owes us an account of just *why* it is crazy."(ibid, p. 295)

Such an extension from living systems to nonliving ones, leading to panpsychism, stems from the notion of information Chalmers adopts for his double-aspect principle. I may not consider panpsychism as a problem associated with the theory, but it is at least very counterintuitive; it is hard to consider a rock having experiences.

My claim is that double-aspect theory of information will be more legitimate if the notion of physically realized information used by the theory can fulfill the following conditions:

1. If the notion can go further than a cause-and-effect explanation. In other words, If the notion can legitimately tell us why a rock cannot have experiences, [94] but, for example, a mouse can.

2. If the notion can legitimately distinguish between living and nonliving systems or in general, if it can legitimately give more clear-cut distinctions.

3. If the notion can successfully explain the physical correlates of mental states and processes in a way that the explanation fits recent experimental findings.

[94] Regarding panpsychism, there is an aspect that has been neglected, namely, the moral aspect of the thesis. Consider a robot, which is organized like me and has silicon chips where I have neurons. I think if we threw the robot away, we are not punished because of doing something immoral. If it is so, we have implicitly presumed that the robot is not conscious.

4. If it can answer more questions and provide more predictions of the system under consideration. In a nutshell, a new theory of information providing us more detailed knowledge of the system under consideration will be more legitimate.

So, if there exists such a notion, we can reliably substitute it with "the difference that makes a difference" principle. My candidate for replacing Chalmers' notion of information with is pragmatic information. In the next section, I will show how Roederer's reading of pragmatic information[95] fulfills all above-mentioned conditions, being a legitimate notion for double-aspect principle.

Care should be taken that in considering Chalmers' thesis as a whole, two points are considerable: First, Chalmers is a naturalistic dualist as mentioned in chapter 1. Second, although he suggests some strategies for avoiding epiphenomenalism, his thesis is known as being epiphenomenal. The thesis I propose in this essay does not go further in these respects. I am faithful to naturalistic dualism in my thesis and have no problem with epiphenomenalism.

[95] Again I have to emphasize that when I speak of pragmatic information in this chapter, I mean Roederer's interpretation of the notion.

## 6.3 Pragmatic Information: a Legitimate Candidate

To see whether pragmatic information is a legitimate notion to be replaced with “the difference that makes a difference” principle in double-aspect theory, we have to see whether the new notion can fulfill the conditions stated above. So I will deal with the conditions one by one.

The first condition explores whether pragmatic information can go further than a simple cause-and-effect explanation. Recall what was said about the distinction between physical interactions and information-driven interactions. In the former, the two interacting bodies were coupled energywise and the final state of the system was totally dependent on the initial conditions. It was mentioned that in macroscopic domain, the observable and measurable interactions usually appear as chains of irreversible cause-and-effect relationships. They are irreversible because to unfold *exactly* in reverse order without external intervention, too many conditions concerning all intervening interactions must be fulfilled *exactly*, making natural occurrence of the reverse process extremely unlikely. In such interactions, which occur in the natural nonliving domain, there need not be any purpose, meaning and means for information generation.

But informational systems acquire more properties as soon as the notion of “purpose” comes into the play. In pragmatic information, when there is purpose, there is “meaning” in a way that a one-to-one correspondence between the sender and the receiver is constructed through a common code. The idea is Darwinian evolution friendly. In any natural living system with information-driven interaction between its parts, the interaction mechanism could only emerge through biological evolution, and in the neural system, through a learning process. So it is clear that the first condition is fulfilled.

The distinction between physical interactions and information-driven ones as depicted in this essay brings the distinction between living and natural nonliving systems about. The claim of this essay is that the conditions of information-driven interactions are only fulfilled in the living systems together with artifacts. Later in the chapter, I will differentiate between information-driven interactions in natural living systems and artifacts, extending the ideas put forward in chapter 4 about intentionality.

Concerning conditions 3 and 4, the thesis of pragmatic information is able to explain several phenomena, some hypothetically and some confirmed (at least partly) experimentally, the phenomena about which Chalmers is either silent or leaves them as open questions. In chapter 4, we saw how “perception,” “short-term memory,” “long-term memory,” “memory recall,” “imagination,” “thinking,” and “reasoning were hypothetically explained by Roederer according to pragmatic information. The idea predicted a sort of top-down control, which was confirmed by experimental findings as stated in the mentioned section. Roederer successfully explains the way information is represented in the nervous system, both microscopically and macroscopically, and the way DNA molecules function and interact based on information-driven interactions. Associative memory as a very fundamental neural processing operation is also explained (see Roederer, 2005, chapter 4).

Finally, it is worth trying to come closer to the very crucial problem associated with the science of consciousness, i.e., the unity of consciousness. Chalmers leaves the issue as an open question:

> “How, within this framework, can one account for the *unity* of consciousness? That is, what makes my visual experiences, auditory experiences, and so on, all experiences of the same subject? I suspect that the answer involves the way that the relevant information is processed, so that the unity of consciousness corresponds to the fact that the relevant information is available to be integrated in a certain way. But just how to cash this out is unclear.” (Chalmers, 1996, p. 309)

According to Roederer, hypothetically, binding is the transformation of radically different patterns into single patterns that, for example in seeing an object, are in correspondence with the topological properties of the form of the object and with the spatial position of the object in the environment. (Roederer, 2005, p. 214)

However, despite the explanatory power of the thesis of pragmatic information, Roederer's account of consciousness fails to capture the subjective character of consciousness. The failure stems from his physicalist point of view. He considers consciousness as arisen from the coherent interplay between cortical and limbic functions:

> "Cortical activity is consistently monitored by the limbic structures and so is the information on the state of the organism […]; the resulting information-based interactions in turn are mapped onto certain cortical regions, especially in the prefrontal areas. In this interplay a balance *must* be achieved, or the organism would succumb to conflicting behavioral instructions. Somehow out of this balance or compromise, based on real-time input, memory of experienced events and instinct, emerges one primary goal for action at a time – this is the essence of animal *core consciousness*." (ibid, p. 205)

Roederer's has adopted the above idea from Damasio (1999). I think such an account of consciousness at best can be regarded as a candidate for being a neural correlate of consciousness (NCC). The definition of a NCC given in the program of the ASSC (Association for the Scientific Study of Consciousness) concerns it as what follows: a neural correlate of consciousness is a specific system in the brain whose activity correlates directly with states of conscious experience.[96]

So, the proposal in this essay is that accommodating pragmatic information in double-aspect principle modifies both theses: On the one hand, it contributes to pragmatic information, as an originally physicalist thesis, another aspect, *viz.*, a phenomenal one. On the other hand, it substitutes the explanatorily powerful notion of pragmatic information for syntactic information.

In the next section, I will use the above combination to approach the hard problem of consciousness.

[96] See Metzinger (2000) for more detail on NCC.

## 6.4 Coming Closer to the Hard Problem of Consciousness

As mentioned above, although pragmatic information seems a legitimate concept for explaining what happens in the realm of living systems, as Roederer is a physicalist and his third-person view leaves out the first-person ontology of consciousness, the theory cannot capture the whole truth. So, because of the arguments, which a dualistic naturalist such as Chalmers may pose, there is something missing in the physical explanation of information-driven interactions, that is, first-person ontology or the realization of subjective quality of consciousness. Combining pragmatic information with double-aspect theory resolves the problem.

An example of such a combination is seen in figure 4.1 in which hypothetical set of sensory information-processing stages in a nervous system was illustrated. There are information-driven interactions between the consecutive neural patterns A, B, C, etc. In such interactions, the two aspects of information arisen from each two interacting patterns have been shown by two arrows drawn from left to right. The same goes for information-driven feedback relationships between the consecutive neural levels A, B, C, etc. The aspects of information in the latter have been shown by two arrows from right to left.

Here a question may arise: If every physically realized information is accompanied by experience, why do not we have multiple experiences each of which associated with each correspondent physically realized information? There are some strategies of constraining double-aspect theory in order to explain unconscious information processing. Recall Roederer's account of binding according to which radically different patterns were transformed into single patterns. So, presumably the answer to the above question is that only the interactions between transformed patterns at higher levels give rise to *my* experience. In other words, a phenomenally realized information counts as unconscious if it is not *my* experience.

Another way of constraining double-aspect is the way Chalmers suggests:

> "The most natural strategy may be to constrain the *way* that the information is processed. After all, I have already said that the

information in my system that corresponds most directly to my experience is the information that is directly available for global control. As it stands, this "criterion" is most unlikely to play a role in a fundamental law, as it is too vague and high-level a notion; indeed, we can use the principle only if we have already individuated a high-level system such as a person or a brain. But perhaps there is a more precise, simpler criterion that could do the work.

One possibility is that *amplification* of information is crucial. Physically realized information is also realized in experience only if the information is *amplified* in certain ways, becoming available to make a large difference along certain causal pathways." (Chalmers, 1996, pp. 300 & 301)

Now we are equipped to summarize the whole thesis proposed in this essay through figure 6.2.

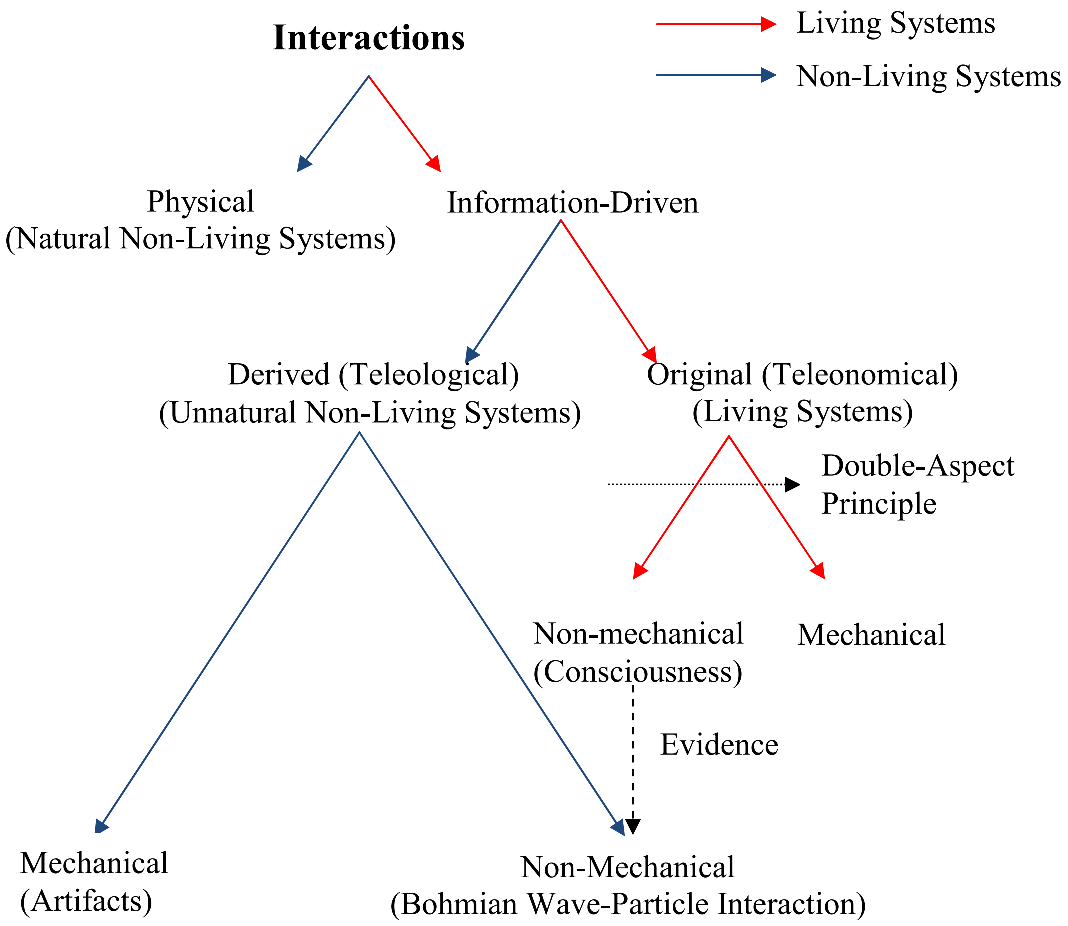


**Figure 6.2.** Two fundamentally different classes of interactions and the subclasses.

The figure shows two fundamentally different classes of interactions the characteristics of which were clarified in chapter 2. We saw that in the natural abiotic world, information plays no role. Physical interactions are driven by energy exchange between the interacting parts without any need to any operations of information processing. It is only by interaction of a living thing with non-living physical world that information comes into play: when a scientist extracts information through observation and measurement or sets up initial conditions of a system, making it an artifact.

However, for living organisms, information, as the essence of their existence, plays a crucial role from the very beginning: "to maintain a long-term state of unstable thermodynamic equilibrium with its surroundings, consistently increase its organization and reproduce, an organism has to rely on information-based interactions in which form or pattern, not energy, is the controlling factor. This latter class comprises *biomolecular* information processes controlling the metabolism, growth, multiplication and differentiation of cells, and *neural* information processes controlling animal behavior and intelligence. The only way new information can appear is through the process of biological evolution and, in the short term, through sensory acquisition and the manipulation of images in the nervous system." (Roederer, 2003, pp. 3-4)

Information-driven interactions include two different subclasses: original and derived. Derived information-driven interactions, which themselves are divided into mechanical and non-mechanical, happen in unnatural non-living systems planned by human beings or by other living systems. They are purely physical mechanisms "but with a purpose to achieve something that would not happen naturally, or just happen by chance." (Roederer, 2005, p.115) Derived information-driven interactions are categorized as informational because *we* humans deliberately set a *purpose* in them by setting the initial conditions or determining measurement setup. Without such a purpose, interactions will be purely physical. Derived information-driven interactions between components of a system are just like mirrors reflecting the information purposefully imposed on them by intelligent beings through setting the initial conditions and measurement setup. They would not happen naturally, or just happen by chance.

A computer or a thermostat is an example of systems realizing mechanical derived information-driven interactions. Moreover, we can prepare a setup such that it shows non-mechanical information-driven interactions. In the experimental setup provided for double-slit experiment, Bohm's description of the wave-particle interaction refers to such an interaction.

To have given definitions, information-driven interactions in a system count as *mechanical* where there is energy and matter transfer between the interacting parts of the system. The latter interactions are in principle reducible to one of the four fundamental interactions known in physics. Where there is no energy and matter transfer between the interacting parts of a system, while fulfilling all conditions of information-driven interactions, the interactions count as *non-mechanical* information-driven interactions. In double-slit experiment, wave and particle as comprising the whole system interact non-mechanically according to Bohm. If Bohm's description of the interaction between wave and particle corresponds to *reality* at the level of hidden variables, or to put it another way, if such an interaction really happens, it can be good evidence for the existence of non-mechanical information-driven interactions. It shows that such interactions are possible to happen. As said earlier, as individual quantum domain is inaccessible, there is no means by which we can explore whether Bohm's description is true or false.

An important point about non-mechanical interactions is that they require nonlocality. So, wherever we have non-mechanical interactions, we have nonlocality. Another important point as regards Bohm's hidden variable theory is that although at individual quantum level we are faced with non-locality, action-at-a-distance is prohibited in the classical world. The same goes for transmission of information.

> "In the simplest case of one Qbit we represent it by two real numbers […] – two "full" real numbers, not truncated by some experimental error, each one of which may carry a nondenumerable *infinite amount of information.* But to conform to experimental results, this quantum information must be pictured as a nonlocal concept, traveling with *infinite speed* from one point to another, even backwards in time […]. This is, indeed, a double whammy for our classical intuition! But not to worry: We cannot access quantum information directly to verify its strange behavior – we can only access its effects on the classical

> macroscopic world, and those behave very tamely, indeed! There is no infinite amount of information and no infinite speed of propagation." (Roederer, 2005, p. 74)

It is time now to focus on original information-driven interactions, which occur in the realm of living systems. The core idea here is that in the domain of living systems, when information is processed non-mechanically, it is also processed mechanically and, at least in some cases, vice versa. It is indeed in this realm that double-aspect principle applies. It seems reasonable to think that it is non-mechanical information processing that gives rise to phenomenal consciousness and at least can be considered a prototype of experience.

Note that the claim in this essay, as opposed to Bohm and Hameroff & Penrose (1996) for example (1996), is not that consciousness arises out of quantum effects. We already got familiar with Bohm's ideas on the implications of his quantum theory for mind. The idea that Hameroff and Penrose put forward suggests that "aspects of quantum theory (e.g. quantum coherence) and of a newly proposed physical phenomenon of quantum wave function "self-collapse" (objective reduction: **OR** -Penrose, 1994) are essential for consciousness, and occur in cytoskeletal microtubules and other structures within each of the brain's neurons. The particular characteristics of microtubules suitable for quantum effects include their crystal-like lattice structure, hollow inner core, organization of cell function and capacity for information processing." (Hameroff & Penrose, 1996) The claim of this essay is that the existence of non-mechanical information-driven interactions at the level of hidden variables shows the possibility of the existence of such interactions. In other words, if consciousness arises out of a sort of information-driven interaction, such interaction must be non-mechanical. Bohm's description of non-mechanical wave-particle interaction provides room for thinking of the possibility of the existence of such interactions. Unfortunately, Bohm did not deal with the phenomenon entanglement applying the notion of active information. But it is reasonable to think that he could have analyzed the interaction between two entangled particles to be a nonlocal, non-mechanical, information-driven interaction. It would be information-driven as we prepare the setup to make the system an entangled one. The phenomenon entanglement as occurring in nature just happens by chance; it may not have the characteristics of the entangled particles provided in the laboratory, unless we have modeled it in our mind.

The thesis stated above rejects Chalmers' claim according to which there is no reason to consider consciousness winking out altogether as we move down the scale of complexity, though as things get less complex, the associated experiences would also be less complex. In Chalmers' opinion, the less complex system, the simpler experience, no matter the system is a living one or a nonliving. So, as experience regards, we can start with human beings who seem to be the most complex information processing systems moving down the scale of complexity to, for example, dogs, mice, lizards, fish and slugs, continuing with thermostats and rocks. As said, he does not see anything crazy with this way of reasoning.

I would not claim that the above thesis is crazy. I claim that the combination of pragmatic information with double-aspect principle as I project it in this essay prevents us from dealing with living information processing systems and non-living ones (artifacts) in the same manner. Moreover, in this essay, I distinguish between artifacts and natural non-living systems. The former are information-processing systems while the latter are not. Such a difference between my proposal and Chalmers' lies in the notion of information I have adopted to apply to double-aspect principle.

Again, the notion of pragmatic information does not consider the principle of organizational invariance as a legitimate one. Nobody knows whether the same experience would arise if the causal patterns of neural organization were duplicated in silicon with a silicon chip for every neuron and the same patterns of interaction. But according to pragmatic information, there is an essential difference between my brain and its silicon duplicate. The former is a teleonomical system with original information-driven interactions between its parts, while the latter is a teleological system, which would not happen in nature or would happen just by chance. The information-driven interactions between the parts composing the latter system reflect the purpose of the maker just as when we read a sad story in a book, the words reflect the purpose of the author to impress the reader. Without such a purpose, the words are nothing but ink stains on the sheets of cellulose fibers. Indeed, if the words are put in book in the same order, but by chance, there is no information associated with them.[97]

[97] This holds true even in MTC. Recall figure 2.10 in which the connection between the sender (pinball machine) and the receiver (measuring device) was cut off. Although the measuring device aimed at the

However, the coherence principles as proposed by Chalmers nicely match my proposal in this essay. As I do not have anything to add to Chalmers' proposal in this respect, I refer the interested reader to the detailed discussion of Chalmers (1996) on coherence principles.

### 6.4.1 Open Questions

The account presented in this essay leaves a large number of questions as open. There might be some plausible answers to such questions, but an explanatory powerful theory does not deal with plausible answers, but exact answers. Here comes some of these questions:

1. What is the mechanism according to which phenomenal aspect of information is realized? To what degree does the complexity of a system play role in the realization of phenomenal information? It is reasonable to think that when all conditions of information-driven interactions are fulfilled, the system becomes potentially ready to realize phenomenal information. As mentioned in section 2.5.2, to have an information-driven interaction, the system must reach a certain degree of complexity, there must exist a one-to-one correspondence between the sender and the receiver while being decoupled energywise, etc. Such conditions to be fulfilled make the system more complex. So, it seems to me that it is more plausible to have phenomenally realized information out of such system compared to a simple cause-and-effect system like a rock. In other words, perhaps one can say that consciousness can be seen as an irreducible systemic property that arises out of information-driven interactions, which just occur in complex systems. However, the exact mechanism is not known.

2. What are the criteria according to which phenomenally realized information corresponds to *my* experiences? Some suggestions were posed earlier in this section, but they are just plausible ones.

---

correct state realized in the pinball machine, as the process was done by chance, the average informativeness was zero.

3. Regarding the idea that experience arises from its physical base, how is the correspondence between experience and its physical base established? In other words, what mechanisms make the phenomenally realized information change in correspondence with a certain change in its physical base? There are two ways of dealing with such questions. Recall figure 5.2 in which arrows are bidirectional. This means that mind and matter not only arise out of a sort of background reality, but also impose some backward causation to it. If the physical property changes, it changes its substrate, which in turn changes the phenomenal property. The alternative idea would be as the following: figure 6.3 below shows a modified version of figure 2.17 compatible with double-aspect theory. The psychophysical law which can answer the above question then would be as the following:

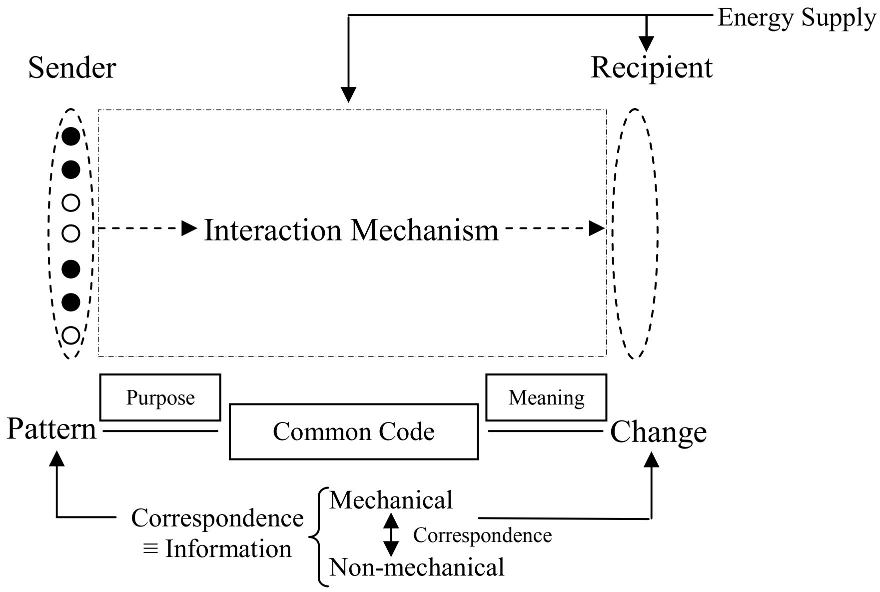


**Figure 6.3.** Modified version of double-aspect principle in virtue of pragmatic information. There are two sorts of correspondence. See below for more detail.

When information is processed phenomenally, two sorts of correspondence are established: 1. A correspondence between the spatial or temporal (or both) feature of the pattern of the sender and a specific change in the recipient. At this stage, we are talking about mechanically realized information; and 2. a correspondence between the relevant phenomenal property (arisen from non-mechanically realized information) and the mechanically realized information. As seen, modifying our double-aspect principle in terms of pragmatic information would cost of adding another principle to the theory, namely, the principle that there is also a correspondence between the mechanically realized information and the non-mechanically realized information. Again, it is not known which of the two alternatives are true to answer the above question.

Finally, it should be mentioned that the idea I put forward in this essay is still in its infancy. As stated earlier, from third-person perspective, consciousness is an inaccessible domain. This is why there are many theses approaching the hard problem of consciousness from different angles all claiming that their thesis captures the truth about consciousness. Like individual quantum domain, in consciousness studies there is no strict criterion to distinguish a thesis as the final one. We can only choose theses with more explanatory power as better ones. What was proposed in this essay was an attempt to locate the thesis among the latter theses.

# Appendix 1

### The frontal cortical control over cardiovascular functioning according to Skinner and Greetje (1997)

As seen in figure 2.18, when senses (eyes, ears, etc.) are elicited by a stressor, the information is delivered to the thalamus to the primary cortex and then to the frontal cortex. A cellular process occurring in the frontal cortex may initiate activity in the frontocortical-brainstem (FC-BS) pathway. The FC-BS pathway reaches the brain stem through the hypothalamus representing direct effect of frontal cortical activity on the viscera (Skinner, 1985a). Whether the frontal cortical process initiate frontocortical-brainstem activity is highly dependent on whether the stressor is perceived as stressful or not. The frontal cortical process is accompanied by a number of electrochemical responses including event-related slow potential, a slow membrane potential shift, and the release of noradrenaline (= norepinephrine). (Greetje, 1997, p. 7) In the brain stem, cardioregulatory centers become active through one of the following pathways: the FC-BS pathway or the indirect pathway via the amygdale. Such activation leads to dual autonomic outflow, namely both sympathetic and parasympathetic activation. The dual autonomic outflow hypothetically increases cardiac vulnerability.

The nucleus ambiguous (NA), the rostral ventrolateral medulla (RVLM), the nucleus of the solitary tract (NTS), and the dorsal motor nucleus of the vagus nerve are the centers in the brain stem which receive the frontal cortical activity. The heart and the circulatory system send signals to NTS:  Baroreceptors send signals to the dorsolateral, medial, and commissural parts of NTS, playing a crucial role in blood pressure regulation. Baroreceptors include those in the auricles of the heart and vena cavae, but the most sensitive baroreceptors are in the carotid sinuses and aortic arch. The carotid sinus baroreceptors are innervated by the glossopharyngeal nerve (CN IX); the aortic arch baroreceptors are innervated by the vagus nerve (CN X). NA and RVLM are activated as the result of NTS projection. NA is triggered by the central nucleus of the amygdala, various hypothalamic areas, the mesencephalic reticular formation (MRF), and NTS. This is the medial part of NTS, which mainly triggers cardioinhibitory part of the NA, the part that has cholinergic (vagal) projections to the heart. There is also a reverse projection from NA back to NTS. The main sympathetic input to the heart and circulatory system is provided by RVLM, which is located near NA. The result of projection from NTS to RVLM is the elicitation of the  intermediolateral cell column (IML), which in turn projects to the sympathetic preganglionic neurons that leads to noradrenergic effects on the heart and circulation.

Finally, the central nucleus of the amygdala, various parts of the hypothalamus, NTS, and the mesencephalic reticular formation (MRF) send afferent projections to the dorsal motor nucleus of the vagus nerve (DMV). The elicitation of DMV leads to the projection to the vagal nerve. “Most of the vagal activity flows to the gastrointestinal system, but neurons located in the lateral part of the DMV project to the heart.” (Martin, 1989; Loewy, 1990; Holstege, 1991)

## Bibliography


Abelli, B. (2007) *On Stage! Playwriting, Directing and Enacting the Informing Processes.* (doctoral dissertation)
Available at: <http://www.diva-portal.org/diva/getDocument?urn_nbn_se_mdh_diva-252-2__fulltext.pdf>

Ackoff, R.L. (1989) "From Data to Wisdom." In *Journal of Applied Systems Analysis* 16, pp. 3-9.

Allen, C. (2005) "Animal Consciousness." In *Online Stanford Encyclopedia of Philosophy*. Available at: <http://plato.stanford.edu/entries/consciousness-animal/>

Armstrong, D. (1968) *A Materialist Theory of Mind*. London: Routledge and Kegan Paul.

Armstrong, D. (1997) "What Is Consciousness?" In *The Nature of Consciousness*. Block, N., Flanagan, O., and Güzeldere, G. (eds). MA: MIT Press.

Aspect, A. Grangier, P. and Roger, G. (1982) "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities." In *Physical Review Letters* 49 (2), pp. 91-94.

Aspect, A., Dalibard, J., and Roger, G. (1982) "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers." In *Physical Review Letters* 49(25), pp. 1804-1807.

Atmanspacher, H. (2006) "Quantum Approaches to Consciousness." In *Online Stanford Encyclopedia of Philosophy*. Available at <http://plato.stanford.edu/entries/qt-consciousness/>

Awad, E. M., Ghaziri, H. M. (2004). *Knowledge Management*. Upper Saddle River, NJ: Pearson Education International.

von Baeyer, H. C. (2003) *Information, the New Language of Science*, Cambridge, Massachusetts: Harvard University Press.

Baggott, J. (2004) *Beyond Measure: Modern Physics, Philosophy, and the Meaning of Quantum Theory*. New York: Oxford University Press.

Bak, P. (1996) *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.

Bar-Hillel, Y., Carnap, R. (1953) "An Outline of a Theory of Semantic Information." Reprinted in: Bar-Hillel, Y.(1964) *Language and Information: Selected Essays on Their Theory and Application.* Reading, Mass; London: Addison-Wesley, pp. 221-74.

Bar-Hillel, Y., (1964) *Language and Information: Selected Essays on Their Theory and Application.* Reading, Mass; London: Addison-Wesley.

Bar-Hillel Y. (1969) "Wesen und Bedeutung der Informationstheorie." In *Informationen über Information*. von Ditfurth, H. (ed.). Hamburg: Hoffmann und Campe, pp. 13–42.

Barwise, J., Seligman, J. (1997) *Information Flow: The Logic of Distributed Systems*. Cambridge: Cambridge University Press.

Bates, M. J. (2005) "Information and Knowledge: an Evolutionary Framework for Information Science." In *Information Research* 10(4).

Available at: <http://informationr.net/ir/10-4/paper239.html>

Bell, J. S. (1987) *Speakable and Unspeakable in Quantum Mechanics.* Cambridge: Cambridge University Press.

Bernecker, S., Dretske, F. (eds) (2000) *Knowledge: Readings in Contemporary Epistemology*. Oxford: Oxford University Press.

Bertuglia C. S., Vaio F. (2005) *Nonlinearity, Chaos and Complexity: The Dynamics of Natural and Social Systems.* Oxford: Oxford University Press.

Block, N. J. (1980) "Are Absent Qualia Impossible?" In *Philosophical Review* 89(2), pp. 257-274.

Block, N. J. (1995) "On a Confusion about a Function of Consciousness." In *Behavioral and Brain Sciences* 18(2), pp. 227-287.

Block, N. J. (1997) "Biology versus Computation in the Study of Consciousness." In *Behavioral and Brain Sciences*, 20(1), pp. 159-165.

Blum, L. (1989) "Lectures on a Theory of Computation and Complexity over the Reals (or an Arbitrary Ring)." In Jen E. (ed.) (1989) *Lectures in Complex Systems*. Santa Fe Institute Studies in the Sciences of Complexity, Lectures Volume 2, Redwood City, California: Addison-Wesley, pp. 1-47.

Bohm, D. (1951) *Quantum Theory*. New York: Prentice Hall. Reprint (1989), New York: Dover.

Bohm, D. (1952) "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables, I and II." In *Physical Review* 85, pp. 166-193.

Bohm, D. (1953) "Proof that Probability Density Approaches $|\psi|^2$ in Causal Interpretation of Quantum Theory." In *Physical Review* 89, pp. 458-466.

Bohm, D. (1957) *Causality & Chance in Modern Physics.* Pennsylvania: University of Pennsylvania Press.

Bohm, D. (1980) *Wholeness and the Implicate Order*. London: Routledge & Kegan Paul.
Bohm, D. (1990) "A new theory of the relationship of mind and matter." In *Philosophical Psychology* 3, pp. 271-286.
Available at: <http://members.aol.com/Mszlazak/BOHM.html>
Bohm, D., Hiley, B. J. (1993) *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge.
Bohr, N. (1935) "Can Quantum Mechanical Description of Physical Reality be Considered Complete?" In *Physical Review* 48, pp. 696-702.
Boyer, C. B. (1968) *A History of Mathematics*. New York: Wiley.
Brentano, F. (1874/1911/1973) *Psychology from an Empirical Standpoint*. London: Routledge and Kegan Paul.
Bricmont, j. (1995) "Science of Chaos or Chaos in Science?" In *Physicalia Magazine* 17 (3-4), pp.159-208. Available at: <http://dogma.free.fr/txt/JB-Chaos.htm>
Brillouin, L. (1962) *Science and Information Theory*. London: Academic Press.
Burry, M., Coulson, J., Preston, J., and Rutherford, E. (2001) "Computer-Aided Design Decision Support: Interfacing Knowledge and Information." In *Automation in Construction* 10, pp. 203-215.
Casti, J. L. (1986) "On System Complexity: Identification, Measurement, and Management." In Casti, J. L., Karlqvist, A. (eds.) (1986) *Complexity, Language, and Life: mathematical approaches*. Berlin: Springer-Verlag, pp. 146-173.
Chaffey, D., Wood, S. (2005) *Business Information Management: Improving Performance Using Information Systems*. Harlow: Prentice Hall.
Chaisson, E. J. (2001) *Cosmic Evolution*. Cambridge, MA: Harvard University Press.
Chaitin, G. J. (1987) *Algorithmic Information Theory*. Cambridge: Cambridge University Press.
Chalmers, D. J. (1995) "Facing Up to the Problem of Consciousness." In *Journal of Consciousness Studies* 2 (3), pp. 200-219.
Chalmers, D. J. (1996) *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.
Chalmers, D. J. (1997) "Availability: The Cognitive Basis of Experience?" In Block, N., Flanagan, O. and Güzeldere, G. (eds) (1997) *The Nature of Consciousness*. MA: MIT Press.

Chalmers, D. J. (1999) “Materialism and the Metaphysics of Modality.” In *Philosophy and Phenomenological Research* 59, pp. 475-496.

Chalmers, D. J. (2002) “Does Conceivability Entail Possibility?” In Gendler, T., Hawthorne, J. (eds) (2002) *Conceivability and Possibility*. Oxford: Oxford University Press, pp.145-200.

Chalmers, D. J. (2003) “Consciousness and its Place in Nature.” In Stich S., Warfield, F. (eds) (2003) The *Blackwell Guide to the Philosophy of Mind.* Oxford: Blackwell. Available at: <http://consc.net/papers/nature.html>

Chalmers, D. J. (2006) “Strong and Weak Emergence.” In Clayton, P. and Davies, P. (eds) (2006) *The Re-emergence of Emergence*. Oxford: Oxford University Press. Available at: <http://consc.net/papers/emergence.pdf>

Chisholm, R. M. (1956) *Perceiving: a Philosophical Study*. In Rosenthal, D. (ed.) (1990) *The Nature of Mind*. Oxford: Oxford University Press.

Churchland, P. M. (1985) “Reduction, Qualia, and Direct Introspection of Brain States.” In *Journal of Philosophy* 82, pp. 8-28.

Clark, A. (1993) *Sensory Qualities.* Oxford: Oxford University Press.

Cleveland, H. (1982) “Information as Resource.” In *The Futurist*. December 1982, pp. 34-39.

Colburn, T. R. (2000) *Philosophy and Computer Science*. Armonk, N.Y.: M.E. Sharpe.

Cohen, I. B. (1993) *The Natural Sciences and the Social Sciences: Some Historical and Critical Perspectives*. Boston, Massachusetts: Boston Studies in the Philosophy of Science.

Damasio, A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* San Diego, New York, Boston: Harcourt, Inc.

Davidson, D. (1983) “A Coherence Theory of Truth and Knowledge.” In LePore, E. (ed.) (1983) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Basil: Blackwell.

Davidson, D. (1986) “Empirical Content.” In E. LePore, (ed.) (1986) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Basil: Blackwell.

Dennett, D. C. (1990) “Quining Qualia.” In Lycan, W. (ed.) (1990) *Mind and Cognition*. Oxford : Blackwell.

Dennett, D. C. (1991) *Consciousness Explained*. Boston: Little, Brown and Company.

Dierks, T., Linden D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., and Singer, W. (1999), “Activation of Heschl’s Gyrus during Auditory Hallucinations.” In *Neuron* 22, pp. 615–621.

Dodig-Crnkovic, G. (2005) “System Modeling and Information Semantics.” In Bubenko, J., Eriksson, O., Fernlund, H., and Lind, M. (eds) (2005) *Proceedings of the Fifth Promote IT Conference*. Borlänge, Sweden, Lund: Studentlitteratur.

Dretske, F. (1980) “The intentionality of cognitive states.” In Rosenthal, D. (ed.) (1990) *The Nature of Mind*. Oxford: Oxford University Press.

Dretske, F. (1981) *Knowledge and the Flow of Information*. Cambridge, Mass.: MIT Press.

Dretske, F. (1991) *Explaining Behavior*. Cambridge, Mass.: MIT Press.

Dretske, F. (1995) *Naturalizing the Mind*. Cambridge, Mass.: MIT Press.

Edmonds, B. (1999) “What is Complexity? – The Philosophy of Complexity per se with Application to some Examples in Evolution.” In Heylighen, F., Bollen J., and Riegler A. (eds.) (1999) *The Evolution of Complexity, the Violet Book of ‘Einstein Meets Magritte.’* Dordrecht: Kluwer Academic Publishers.

Available at: <http://bruce.edmonds.name/evolcomp>

Einstein, A., Podolsky, B., and Rosen, N. (1935) “Can Quantum Mechanical Description of Physical Reality Be Considered Complete?” In *Phys. Rev.* 47, pp. 777-780.

Fetzer, J. H. (2004) “Information, Misinformation, and Disinformation.” In *Minds and Machines* 14(2), pp. 223-229.

Flake, G. W. (1999) *The Computational Beauty of Nature*. Cambridge, Mass.: MIT Press.

Floridi, L. (1999) *Philosophy and Computing: An Introduction*. London, New York: Routledge.

Floridi, L. (2004) “Outline of a Theory of Strongly Semantic Information.” In *Minds and Machines* 14(2), pp. 197-222.

Floridi, L. (2005a) “Semantic Conceptions of Information.” In *Online Stanford Encyclopedia of Philosophy*.

Available at: http://plato.stanford.edu/entries/information-semantic

Floridi, L. (2005b) “Is Information Meaningful Data?” In *Philosophy and Phenomenological Research* 70(2), pp. 351–370.

Fodor, J. A. (1987) *Psychosemantics*. Cambridge, Mass.: MIT Press.

Fodor, J. (1991) “Too Hard for Our Kind of Mind?” In *London Review of Books* 13(12).

Foley, R. (1987), “Dretske’s “Information-Theoretic” Account of Knowledge.” In *Synthese* 70, pp. 159-184.
Fox, C. J. (1983) *Information and Misinformation: An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming*. Westport, CT: Greenwood Press.
Frieden, B. R. (2004) *Science from Fisher Information: A Unification* (2nd ed.). Cambridge: Cambridge University Press.
Gäng, P. (1967) “Pragmatische Informationen.” In *Grundlagenstudien aus Kybernetik und Geisteswissenschaft* 8, pp. 77–90.
Gell-Mann, M. (1994) *The Quark and the Jaguar*. London: Little, Brown and Company.
Gernert, D. (2006) “Pragmatic Infoemtion: Historical exposition and General Overview.” In *Mind & Matter* 4(2), pp. 141-167.
Graham, G. (1999) *The Internet: A Philosophical Inquiry*. London: Routledge.
Greenberg W.M. (1998) “On Chalmers’ principle of organizational invariance and his ‘dancing qualia’and ‘fading qualia’thought experiments.” In *Journal of Consciousness Studies* 5(1), pp. 53-58.
Greetje, K. (1997) “*Brain Control of Heart Regulation* (doctoral dissertation).
Available at: http://dissertations.ub.rug.nl/faculties/ppsw/1997/g.koers/
Grice, H. P. (1989) *Studies in the Way of Words.* Cambridge, Mass.: Harvard University Press.
Grice, P. (1957) “Meaning.” In Grice, P. (1989) *Studies in the Way of Words*. Cambridge, Mass.: Harvard Univesity.
Güzeldere, G. (1997) “The Many Faces of Consciousness: A Field Guide.” In *The Nature of Consciousness*. Block, N., Flanagan, O. and Güzeldere, G. (eds.) (1997) Mass.: MIT Press.
Hameroff, S. and Penrose, R. (1996) “Orchestrated Reduction of Quantum Coherence in Brain Microtubules: A Model for Consciousness.” In *Mathematics and Computers in Simulation* 40(3), pp. 453-480.
Available at: http://www.quantumconsciousness.org/penrose-hameroff/orchOR.html
Harms, W. F. (1998) “The Use of Information Theory in Epistemology.” In *Philosophy of Science* 65(3), 472-501.
Hartshorne, C. and Weiss, P. (eds.) (1932) *Collected Papers of Charles Sanders Peirce II*. Cambridge, MA: Harvard University Press.

Haugeland, J. (1981) “Semantic Engines: an Introduction to Mind Design.” In Haugeland, J. (ed.) (1981) *Mind Design, philosophy, Psychology, Artificial Intelligence*. Cambridge, Mass.: MIT Press.

Heylighen, F. (2008) “Complexity and Self-Organization.” In Bates, M. J., Maack, N. (eds) (2008) *Encyclopedia of Library and Information Sciences*. Taylor & Francis. Available at: http://pespmc1.vub.ac.be/Papers/ELIS-Complexity.pdf.

Hintikka, J. and Suppes, P. (eds) (1970) *Information and Inference*. Dordrecht: Reidel.

Holstege, G. (1991) “Descending motor pathways and the spinal motor system: Limbic and non-limbic components.” In *Progress in Brain Research* 87, pp. 307-421.

Horgan, P. (1995) “From Complexity to Perplexity.” In *Scientific American*, June, pp. 104-109.

Ishai, A. and Sagi, D. (1995) “Common Mechanisms of Visual Imagery and Perception.” In *Science* 268, pp. 1772–1774.

Ishai, A., Ungerleider, L.G., and Haxby, J.V. (2000) “Distributed Neural Systems for the Generation of Visual Images.” In *Neuron* 28, pp. 979–990.

Jackson, F. (1982) “Epiphenomenal Qualia.” In *Philosophical Quarterly* 32, pp. 127-136.

Jacob, P. (2003) “Intentionality.” In *Online Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/entries/intentionality/

Kant, I. (1790) *Kritik der teleogischen Urteilskraf*. Berlin and Libau

Kornwachs, K. and Lucadou, W. von (1982) “Pragmatic information and nonclassical systems.” In Trappl, R. (ed) *Cybernetics and Systems Research*. North-Holland, Amsterdam, pp. 191–197.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J.P., Thompson, W. L., Ganis, G., Sukel, K. E., and Alpert, N. M. (1999) “The Role of Area 17 in Visual Imagery: Convergent Evidence from PET and rTMS.” In *Science* 284, pp. 167–170.

Kreiman, G., Koch, C., and Fried, I. (2000) “Imagery Neurons in the Human Brain.” In *Nature* 408, pp. 357–361.

Kripke, S. (1972) “Naming and Necessity.” In Davidson D., Harman, G. (eds) (1972) *Semantics of Natural Language.* Holland: Reidel, pp. 253-355.

Küppers, B. O. (1990) *Information and the Origin of Life.* Cambridge, Mass.: The MIT Press.

Lambert, S., Sampaio, E., Mauss, Y., and Scheiber, C. (2004) "Blindness and Brain Plasticity: Contribution of Mental Imagery? An fMRI study." In *Brain Res. Cogn.* 20, pp. 1–11.

Laplace, P. S. (1825) A Philosophical Essay on Probabilities, Transl. by F. W. Truscott and F. L. Emory, Dover Pub., New York, 1951. Essai philosophique sur les probabilit´es, r´e´ed. C. Bourgeois, Paris, 1986 (texte de la 5`eme ´ed., 1825).

Levine, J. (1983) "Materialism and Qualia: The Explanatory Gap." In *Pacific Philosophical Quarterly* 64, pp. 354-361.

Levine, J. (2001) *Purple Haze: The Puzzle of Consciousness*. Oxford, New York: Oxford University Press.

Lioyd S. (1989) "Physical Measures of Complexity." In Jen E. (ed.) (*1989*) *Lectures in Complex Systems*, *Santa Fe Institute Studies in the Sciences of Complexity*. Lectures Volume 2, Redwood City, California: Addison-Wesley, pp. 67-73.

Loewy, A. D. (1990) "Central Autonomic Pathways." In A. D. Loewy & K. M. Spyer (eds.) (1990) *Central Regulation of Autonomic Functions.* New York: Oxford University Press, pp. 88-103.

Loewy, A. D. and Spyer, K. M. (1990) "Vagal Preganglionic Neurons." In Loewy, A. D. and Spyer, K. M. (eds) (1990) *Central Regulation of Autonomic Functions.* New York: Oxford University Press, pp. 68-87.

Lombardi, O. (2004) "What Is Information?" In *Foundation of Science* 9, pp. 105-134.

Lormand, E. (1998) "Consciousness." In *Routledge Encyclopedia of Philosophy*. New York: Routledge.

Lucas, C. (2007) "Complexity Philosophy as a Computing Paradigm." In Roetzheim, W. (2007) *Why Things Are - How Complexity Theory Answers Life's Toughest Questions*. Level 4 Press, Inc.

Available at http://www.calresco.org/lucas/compute.htm

Lycan, W. (1996) *Consciousness and Experience*. Cambridge, MA: the MIT Press.

Mach, E. (1883) *Die Mechanik in ihrer Entwickelung, historisch-kritisch dargestellt* (2nd ed.). Leipzig: Brockhaus. English translation (6th ed., 1960): *The Science of Mechanics*. La Salle, Illinois: Open Court Press.

Mackay, D. (1980) *Brain, Machines and Persons.* St. James's Place, London: Collins.

MacKay, D. J. C. (2003) *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Martin, J. H. (1989) *Neuroanatomy: Text and Atlas*. Norwalk, Conn: Appleton & Lange.

McLaughlin, B. P. (1992) “The Rise and Fall of British Emergentism.” In Beckermann, A. Flohr, H. and Kim, J. (eds.) (1992) *Emergence or Reduction?: Prospects for NonreductivePhysicalism.* Berlin: De Gruyter, pp. 49–93.

Mechelli, A., Price, C.J., Friston, K.J., and Ishai, A. (2004) “Where Bottom- Up Meets Top-Down: Neuronal Interactions During Perception and Imagery.” In *Cereb. Cortex* 14, pp. 1256–1265.

Meixner, U. (2004) *The Two Sides of Being*. Paderborn: Mentis.

Metzinger, T. (ed.) (2000) *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. The MIT Press.

Miyashita, Y. (2004) “Cognitive Memory: Cellular and Network Machineries and Their Top-Down Control.” In *Science* 306, pp. 435-440.

Mingers, J. (1997) “The Nature of Information and Its Relationship to Meaning.” In Winder, R. L. *et al.* (eds) (1997) *Philosophical Aspects of Information Systems*. London: Taylor and Francis, pp. 73-84.

Monod, J. (1972) *Chance and Necessity.* London: Collins. (Originally published in 1970: *Le Hasard et la Nécessité* Paris).

Morris C. W. (1938) “Foundations of the Theory of Signs.” In Neurath, O. *et al.* (eds) *International Encyclopedia of Unified Science I* (2nd ed.) Chicago: University of Chicago Press, pp. 78–137.

Nagel, T. (1974) “What is it Like to Be a Bat?” In *Philosophical Review* 83, pp. 435-450. (Reprinted in his (1979) *Mortal Questions.* Cambridge, England: Cambridge University Press.)

Newton, I. (1687) *Philosophiae naturalis principia mathematica.* London: Royal Society.

O’Connor, T. & Yu Wong, H. (2006) “Emergent Properties.” In *Online Stanford Encyclopedia of Philosophy*.

Available at: http://plato.stanford.edu/entries/properties-emergent/

O’Craven, K. M., and Kanwisher, N. (2000) “Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions.” In *Cogn. Neurosci.* 12, pp. 1013–1023.

Pauli, W. (1928) in *Solvay* 1928, pp. 280-282.

Peirce, C.S. (1902), “Logic, Considered as Semeiotic.” Manuscript L75.

Available at: http://www.cspeirce.com/menu/library/bycsp/l75/l75.htm, and "On the Definition of Logic." (Memoir 12)
Available at: http://www.cspeirce.com/menu/library/bycsp/l75/ver1/l75v1-05.htm#m12
Penrose, R. (1994) *Shadows of the Mind*. London: Oxford University Press.
Pittendrigh, C. S. (1958) "Adaptation, Natural Selection and Behavior." In Roe, A. and Simpson, G. G. (1958) *Behavior and Evolution*. New Haven: Yale University Press, pp. 390-416
Plato, *Timaeus*. Translated by Jowett translation.
Plato, *The Republic*. (New CUP translation by Tom Griffith and G.R.F. Ferrari into English)
Prigogine, I. (1980) *From Being to Becoming*. New York: Freeman.
Prigogine, I. (1983) Un si`ecle d'espoir, in Temps et Devenir, Colloque de Cerisy, `a partir de l'oeuvre d'Ilya Prigogine; ´eds J. P. Brans, I. Stengers, P. Vincke; Patino.
Prigogine, I. and Stengers, I. (1984) *Order out of Chaos.* London: Heinemann.
Prigogine, I. (1994) *Les Lois du Chaos*. Paris: Flammarion.
Prigogine, I. (1995) "Why Irreversibility? The Formulation of Classical and Quantum Mechanics for Nonintegrable Systems." In *International Journal of Quantum Chemistry* 53, pp. 105-118.
Prigogine, I. (1996), La fin des certitudes. Temps, chaos et les lois de la nature, Odile Jacob, Paris.
Prigogine, I. and Stengers, I. (1979) La Nouvelle Alliance. M´etamorphoses de la science, Gallimard, Paris, (coll. Folio, Paris, 1986).
Prigogine, I. & Stengers, I. (1988), Entre le temps et l'´eternit´e, Fayard, Paris, (coll. Champs, Flammarion, Paris, 1992).
Putnam, H. (1975) "The Meaning of 'Meaning'." In *Mind, Language, and Reality: Philosophical Papers II*. Cambridge University Press.
Quine, W. V. O. (1960) *Word and Object*. Cambridge, Mass.: The MIT Press.
Revonsuo, A. and Kamppinen, M. (1994) "General Introduction: The Riddle of Consciousness." In Revonsuo, A. and Kamppinen, M. (eds) (1994) *Consciousness in Philosophy and Cognitive Neuroscience.* Hillsdale, NJ: Erlbaum.
Roederer, Juan G. (2003) "On the Concept of Information and Its Role in Nature." In *Entropy* 5, pp. 3-33

Roederer, J.G. (2005) *Information and Its Role in Nature.* Berlin Heidelberg: Springer-Verlag.

Rosenthal, D. (1991) "Two Concepts of Consciousness." In Rosenthal, D. (ed.) (1991) *The Nature of Mind*. Oxford: Oxford University Press.

Rosenthal, D. (1997) "A Theory of Consciousness." In Block, N., Flanagan, O. and Güzeldere, G. (eds.) (1997) *The Nature of Consciousness*. MA: MIT Press.

Rowley, J. (2006) "Where Is the Wisdom that We Have Lost in Knowledge?" In *Journal of Documentation* 62(2), pp. 251-270.

Rowley, J. (2007) "The wisdom hierarchy: representations of the DIKW hierarchy." In *Journal of Information Science* 33(2), pp. 163-180.

Ryle, G. (1949) *The Concept of Mind*. Chicago: University of Chicago Press.

Searle, J. (1980) "Minds, Brains and Programs."In *The Behavioral and Brain Sciences* 3, pp. 417-424.

Searle, J. (1983) *Intentionality*. Cambridge: Cambridge University Press.

Searle, J. (1992) *The Rediscovery of the Mind*. Cambridge, Mass.: MIT Press.

Searle, J. (1995) *The Construction of Social Reality*. Allen Lane: The Penguin Press.

Searle, J. (2004) *Mind, A Brief Introduction*. New York, Oxford University Press.

Sellars, W. (1956) *Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Sen A. and Smith T. E. (1995) *Gravity Models of Spatial Interaction Behavior*. Berlin: Springer Verlag.

Shannon, C. E. (1948) "A Mathematical Theory of Communication." In *Bell System Technical Journal* 27, pp. 379–423, 623–656.

Shannon C. E. (1956) "The Bandwagon." In *IRE Transactions on Information Theory* IT-2(1), March.

Shannon, C. E., and Weaver, W. (1949) *The Mathematical Theory of Communication* (twelfth printing 1971). Urbana, IL: University of Illinois Press.

Sharma, N. (2005) "The Origin of the "Data Information Knowledge Wisdom" Hierarchy"
Available at: http://www-personal.si.umich.edu/~nsharma/dikw_origin.htm

Shedroff, N. (2001) "An overview of understanding." In Richard Saul Wurman, R. S. (ed.) (2001) *Information Anxiety 2*. Indianapolis: Que.

Siewert, C. (2003) "Consciousness and Intentionality." In *Online Stanford Encyclopedia of Philosophy*.
Available at: http://plato.stanford.edu/entries/consciousness-intentionality/

Skinner, J. E. (1985a) "Psychosocial Stress and Sudden Cardiac Death: Brain Mechanisms." In R. E. Beamish, P. K. Singal & N. S. Dhalla (eds) (1985) *Stress and Heart Disease*. Boston: Martinus Nijhoff Publishing, pp. 44-59.

Stoljar, D. (2001) "The Conceivability Argument and Two Conceptions of the Physical." In Tomberlin, J. E. (ed.) (2001) *Philosophical Perspectives* 15. Boston MA and Oxford: Blackwell, pp. 393-413.

Stephan, A. (1992) "Emergence - A Systematic View on Its Historical Aspects." In Beckermann, et. al. (eds.) (1992) *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism.* Berlin: Walter de Gruyter, pp.25-47.

Stephan, A. (1997) "Armchair Arguments against Emergentism." In *Erkenntnis* 46, pp. 305-314.

Stephan, A. (1998) **"**Varieties of Emergence in Artificial and Natural Systems." In *Zeitschrift für Naturforschung* 53c, pp. 639–656.

Stephan, A. (1999a) *Emergenz, Von der Unvorhersagbarkeit zur Selbstorganisation.* Dresden, München: Dresden University Press.

Stephan, A. (1999b) "Varieties of Emergentism." In *Evolution and Cognition* 5(1), pp. 49-59.

Stephan, A. (2004) "Phenomenal Emergence."In *Networks* (3-4), pp. 91-102.

Taylor, E. W. et al. (1999) "Central Control of the Cardiovascular and Respiratory Systems and Their Interactions in Vertebrates." In *Physiological Reviews* 70(3). Available at: <http://physrev.physiology.org/cgi/reprint/79/3/855>

Thompson, D. W. (1961) *On Growth and Form* (abridged). Cambridge University Press.

Tye, M. (1995) *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.

Tye, M. (2000) *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Tye, M. (2007) "Philosophical Problems of Consciousness." In Velmans, M. and Schneider, S. (2007) *The Blackwell Companion to Consciousness*. Blackwell Publishing Ltd.

Van Gulick, R., (1993) "Understanding the Phenomenal Mind: Are We All just Armadillos?" In Davies, M. and G. Humphreys (eds.) (1993) *Consciousness*. Oxford: Basil Blackwell.

Van Gulick, R. (2004) “Consciousness.” In *Online Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/entries/consciousness

Weizsäcker C.F. von (1985) *Aufbau der Physik*. München: Deutscher Taschenbuch Verlag.

Weizsäcker E. von and Weizsäcker C. von (1972) “Wiederaufnahme der begrifflichen Frage: Was ist Information?” In *Nova Acta Leopoldina* 206, Band 37/1, J.A. Barth, Leipzig: pp. 535–555.

Wittgenstein, L. (1953) *Philosophical Investigations*. Trans. By G.E.M. Anscombe. Macmillan.

Wolfram, S. (1994) *Cellular Automata and Complexity: Collected Papers*. Addison-Wesley.

Young, P. (1987) *The Nature of Information*. Praeger.

Zeleny, M. (1987) “Management Support Systems: towards Integrated Knowledge Management.” In *Human Systems Management* 7(1), pp. 59-70.

Zurek, W. H. (1989) “Algorithmic Information Content, Church-Turing Thesis, Physical Entropy, and Maxwell’s Demon.” In Jen E. (ed.) (1989) *Lectures in Complex Systems*, Santa Fe Institute Studies in the Sciences of Complexity, Lectures Volume 2, Redwood City, California: Addison-Wesley, pp. 49-65.