Cognitive Biology

Dealing with Information from Bacteria to Minds

Gennaro Auletta

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2011

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978–0–19–960848–5

1 3 5 7 9 10 8 6 4 2

Alla memoria di mia madre e mio padre, Chiara e Ferdinando.

Foreword

I imagine you reading the first page of this book with a sense of anticipation and, perhaps, a degree of trepidation. This is a big book that addresses big questions. It will take you on a long journey, probably through unfamiliar terrains. This book addresses one of the most important challenges facing science and philosophy; namely a deep and easy understanding of ourselves. This is clearly not a trivial challenge. Indeed, despite remarkable progress over the past century, no single discipline has provided a satisfactory answer. The path taken in this book captures and incites a paradigm shift in neuroscientific thinking; it transcends any single discipline to provide an inclusive and principled approach to questions about why we are here and how we work. Its agenda is to establish a *cognitive biology* that situates cognitive neuroscience in its biological substrates. This approach allows the author to call upon powerful constraints and constructs, which are used to dissolve fundamental problems in cognitive science and related fields. The result is a picture of the brain that is illuminated from several perspectives; a picture that could describe any adaptive, self-organizing biological system.

The inclusive and eclectic treatment offered by this book rests on building bridges between different disciplines and formalisms. For example: bridging between quantum physics and information theory; information theory and perception; perception and semiotics; semiotics and behavior. The author's remarkable capacity to bridge these distinct fields reflects his role as Professor at the Pontifical Gregorian University. The breadth of knowledge and scholarship that underwrites this grand synthesis is truly remarkable. I was genuinely surprised to find that one person could write so fluently and didactically about so many diverse scientific takes on the same problem. Furthermore, in the areas that I am familiar with, I found his treatment very contemporary; as if the author had magically attended all the key conferences and workshops relevant to this work in the past couple of years. His panoramic approach works very well and enables each part of the synthesis to be presented in a concise and contextualized fashion.

This book has three parts, all framed around the notion of information. The first deals with the fundamental nature of information, its acquisition and measurement. The second considers information in the exchange of agents with their environment and considers the general principles of self-organization and dynamics in biological systems. In the third part of the book, we turn to the interpretation of information, touching on high-level issues that attend language acquisition and awareness. Information is a crucial theme throughout; it ties together the most elementary formulations of our physical world, in terms of quantum physics, all the way through to symbolic representations and pragmatics. It speaks to the underlying premise that the brain is an active organ constructing, selecting and exchanging information with its environment. This builds upon the early work of physicists like Helmholtz, who viewed the brain as making inferences about its sensorium, generating predictions and testing hypotheses through interaction with the world. This view has dominated modern treatments of brain function (and indeed machine learning) in recent years. Information is inherently probabilistic and is only defined in relation to some model of the system generating information. Again, this probabilistic view is very current and is at the heart of things like the Bayesian brain hypothesis. However, this book takes us beyond the Bayesian brain and considers why the selection and control of sensory information is mandated in all complex self-organizing systems. It is this contextualization of modern treatments of the brain that lends the author's synthesis simplicity and substance.

If you read this book you will learn a lot. Furthermore, it may help leverage your own insights in a larger context. Having said this, there are lots of ideas and material that need to be assimilated. This assimilation is made relatively easy by the book's primer-like style when covering technical material. I confess I was a bit surprised that the publishers allowed so much mathematics to adorn the text. Although it looks a little daunting at first glance, the mathematics are used in a largely iconic and useful way and, with a little perseverance from the reader, it works very well.

We are invited to start this book with some quantum physics. Although this is probably not in the comfort zone of most neuroscientists, let me reassure you and encourage you to start reading. The quantum mechanical treatment is simply presented to underscore the fact that information is inherently probabilistic and is the product of an explicit measurement or action. In one sense, we are invited to engage in our own information acquisition at this level, and realize the book's latent information, in terms of our understanding of what we are. I guarantee that when you reach the end of the book you will look at the brain, and perhaps yourself, in a new light.

Karl J. Friston*
University College London

* Professor Karl J. Friston MB, BS, MA, MRCPsych, FMed.Sci, FRS Scientific Director: Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, k.friston@fil.ion.ucl.ac.uk.

Author's Preface

This book is the result of a lifetime of study and especially of ten years of intensive research. The original idea was to search for the biological bases of cognition. However, during this preparatory study I understood that we cannot deal with this issue without resorting to an engagement with both physics and information theory. These two aspects are deeply connected in quantum mechanics and, as the reader will see, in many respects this is very helpful for reading the book.

The turning point in my research was when I became acquainted with the work of those neuroscientists who, against a traditional cognitivist understanding of cognition, have stressed the active role of the brain. I recall here Marc Jeannerod in particular. This framework was particularly suitable to me since I was (and still am) very much influenced by C. Peirce's pragmatism. The insights provided by this neurological approach allowed me to consider the role of cognition *for* biology for the first time.

A second step was represented by the contributions coming from biology, especially from the field of studies known as epigenetics and its impact on traditional evolution theory (the so-called Evo-Devo). Here, I was very much influenced in particular by the work of Scott Gilbert. This was a new, dynamic and interactive way of understanding the organism. Such a point of view was not only in perfect agreement with the previous one but opened the door to unexpected possibilities of conceiving the interconnections between cognition (understood here in the largest sense of the word) and biology. In such a context I was also supported by the important insights developed by Francisco Ayala on the distinction between teleonomy and teleology. I wish also to recall here that some interactions with George Ellis (Capetown University) were very fruitful for the definition of complexity and the issue of top-down causation.

A third important turn was when I became aware of Karl Friston's research. His work on Bayesian probabilities and on the connection between lowering the informational surprisal and lowering the statistical equivalent of the free energy provided me with strong ties between cognitive and biological processes, at least at an ontogenetic level. Moreover, it explained how the specific information-controlling activity of organisms could have arisen at all. Previously, I had looked on my own for a connection between inverse Bayesian probabilities and information in order to establish an informational framework in which the specific activity of organisms in controlling their environment could have arisen. Unfortunately, I could not arrive at a straight result like that provided by Friston, so that when I had the opportunity to read his papers it was a sort of revelation.

This book is addressed to any scholar (from undergraduate students to senior researchers) interested in learning a little more about biology and cognition. It is not necessary to read the whole book but is also possible to simply read specific arguments or select a personal reading path thanks to the summary, indexes, definitions, and cross-references. Moreover, interested scholars do not need to be in the field of biology and cognitive or neurological sciences: Anybody interested tangentially in these issues may benefit from reading the book or some of its parts. I expect that philosophers, ethologists, primatologists, physicists, and chemists in particular will be interested, at least for methodological reasons. There are no specific preliminary requirements for reading this book: Basic information for each field or problem is provided. It is true that this book uses some mathematics, but in general a high-school level is sufficient. More sophisticated mathematical topics have been kept separated in the appendices.

Finally, allow me to thank all the people who have helped me in this enterprise. I would like very much to thank my coworkers Ivan Colagé and Paolo D'Ambrosio for their constant and very engaging interaction with me, and the numerous suggestions and improvements that they proposed. I also warmly thank Marc Jeannerod (French Institute for Cognitive Sciences) who not only provided me with important insights but encouraged and helped me to make this work possible. A special thanks to Karl Friston (University College, London) for having supported me with open-mindedness and friendship as well as for having kindly accepted to write a Foreword. I also warmly thank Francisco Ayala (University of California, Irvine) for the many discussions and his authoritative sympathy for my work. Moreover, I wish to thank Jean-Pierre Changeux (Institut Pasteur), Simon Conway-Morris (University of Cambridge), Geoffrey Cook (University of Cambridge), Miguel R. Fuentes (Pontifical Gregorian University), Arthur Gibson (University of Cambridge), Scott Gilbert (Swarthmore College), Sir Brian Heap (University of Cambridge), Stuart Kauffman (University of Calgary), Giorgio Parisi (University La Sapienza), Angelo Tartabini (University of Parma), and Robert E. Ulanowicz (University of Maryland) for their very helpful and constructive contributions. I cannot omit to thank my student Carly Andrews for having helped me to purify my English. Finally, I would also like to thank Sonke Adlung, senior editor at OUP, very much for his kind assistance and help.

Nature is not economical of structures: only of principles of fundamental applicability.
[SALAM 1979]

Contents

List of abbreviations

ADP	Adenosine diphosphate
AI	Artiticial Intelligence
AL	Artiticial Life
AM	Amplitude modulation
AMPA	Alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANN	Attractor neural network
AP	Action potential
ARAS	Ascending reticular activating system
ASL	American Sign Language
ATP	Adenosine triphosphate
BRAC	Basic rest–activity cycle
BS	Beam splitter
BZ	Belousov–Zhabotinsky (reaction)
cAMP	Cyclic adenosine monophosphate
Ch.	Chapter
CNS	Central nervous system
CR	Conditioned response
CRP	cAMP receptor protein
CS	Conditional stimulus
Def.	Definition
DNA	Deoxyribonucleic acid
EEG	Electroencephalogram
EQ	Encephalization quotient
Fig.	Figure
FMRI	Functional Magnetic Resonance Imaging
GABA	Gamma-aminobutyric acid
HIS	Hereditary immune system
iff	If and only if
IPL	Inferior parietal lobule
iRNA	Interference RNA
IT	Inferior temporal cortex
JND	Just noticeable difference
LASER	Light Amplification by Stimulated Emission of Radiation
LF	Logical form

LGN	Lateral geniculate nucleus
lhs	Left-hand side
LIP	Lateral intraparietal region
LTM	Long-term memory
LTP	Long-term potentiation
M	Magno cell
M1	Primary motor cortex or neurons
mRNA	Messenger RNA
MSE	Mean Standard Error
MTL	Medial temporal lobe
NMDA	N-methyl-D-aspartic acid
P	Parvo cell
PDP	Parallel distributed processing
PF	Phonetic form
PFC	Prefrontal cortex
PET	Positron emission tomography
PM	Premotor cortex
pp.	Pages
Pr.	Principle
PSP	Postsynaptic Potential
PVN	Paraventricular nucleus
REM	Rapid eye movement
rhs	Right-hand side
RNA	Ribonucleic acid
RNAP	RNA polymerase
rRNA	Ribosomal RNA
SD	Structural description
Sec.	Section
S1	Somatosensory cortex
S–R	Stimulus–reaction
S-structure	Surface structure
STM	Short-term memory
Subsec.	Subsection
SWS	Slow-wave sleep
Tab.	Table
TE	Anterior inferotemporal cortex
TNGS	Theory of neural group selection
Th.	Theorem
tRNA	Transfer RNA
UG	Universal grammar
UR	Unconditional response
US	Unconditional stimulus
V1	Primary visual cortex
V2	Secondary visual cortex

Introduction

Biological and cognitive sciences are at a conceptual turning point at the beginning of this century. Let us consider first *biology* and then the cognitive sciences.

- There have been amazing steps taken at the level of both molecular and evolutionary biology (the extent of their progress will be shown in the following chapters). However, we also have the growing feeling that traditional reductionist methodologies are largely insufficient for dealing with the complexity of systems and problems that are increasingly at the core of our current scientific inquiry.[1] The traditional language of biology has its roots in the classical language of physics and its explanatory machinery. The latter consists in mechanical causation and its proper terminology is that of mass, energy, and force, which at a molecular and chemical level involves molecular forces, concentration of certain chemicals, rate of a reaction, and so on, which is a formal language that does not seem completely satisfactory today when dealing with higher levels of biological organization. It is much more suitable here to deal with parameters like formal constraints, dissipative events, differential timing, degenerate states and processes. As we shall see all these aspects are crucial for epigenetic processes. Moreover, from the following inquiry it will become evident that top-down constraints must be added to the list of explanatory mechanisms. All these processes and features are rooted in the language and the mechanisms of *complexity theory*.
- On the other hand, *cognitive science* (cognitive psychology, cognitive neuropsychology, neural network theory, and so on) has also taken huge steps in conceptualization and experiments. However, the general trend, at least in the last decades, has been a sort of functionalism through which the mind or the brain is conceived as an entity separated from its biological background. The functionalist language of cognitive science is rooted in cybernetics. The aim of cybernetics was the utilization of the framework of classical information and communication, as it was firmly established by Shannon,[2] for the foundations of cognition.[3] For reasons that will be clear in the following chapters, the final result of this scientific enterprise was relatively unsatisfactory. Artificial intelligence (AI), the most qualified scientific vector of cognitive sciences, did not lead to the expected successes.[4] This methodology does not seem to catch the fundamental complexity of the brain and the mind as it is rooted in biology. As a matter of fact, at the beginning of this scientific enterprise, it was assumed that the brain represented the external world in a passive

[1] [*GORDON* 1999]. [2] [SHANNON 1948].

[3] [*WIENER* 1948][*ASHBY* 1956]. Wiener stressed that cybernetics can be seen as an attempt at transforming communication engineering in a branch of statistical mechanics, and this could not be resolved without a classical theory of information [*WIENER* 1948, p. 10].

[4] [*HAWKINS/BLAKESLEE* 2004, pp. 13–18].

manner, mirroring external objects and their properties, and that the role of thought consisted essentially of classically (logically) processing the information acquired in this way. In the latter 20 years there has been a growing body of evidence that this is not the right story. The brain has been increasingly understood not only as an active and interactive entity (as biological systems already are) but as a system whose nature lies rather in perception–anticipation and in strategic pursuit of aims.[5] These results have completely transformed the character, scope, and methodologies of the cognitive sciences. For these reasons, a new language is required, centered on concepts like information sharing and information selection out of an initial variety. The roots of this language can be found in the basic theory of our physical universe, i.e. in quantum mechanics, and especially in the growing field known as *quantum information*.

- Finally, I stress that the tendential split between a too reductionist understanding of biology and a functionalist approach to cognition may have controversial implications for future development in natural sciences. As a matter of fact, today we understand more and more that cognition is rooted in biology. This means that we must find in biology the elements or at least the presuppositions that lead to cognition. However, this also implies that cognitive processes (even at a very elementary level) are important for the organism's survival. To this extent, cognition is not irrelevant to biology just as biology is not irrelevant to cognition.[6] Thus, the object of this book should be considered as an attempt at building a new discipline, that of *cognitive biology*, which endeavors to bridge these two domains. From what has been said before, a reference to physics seems unavoidable here, especially when foundational issues are at play.[7] This was the path chosen by thinkers like W. Wundt or H. von Helmholtz in the 19$^{\text{th}}$ century.[8] E. Mach opened his book on sensation with a presentation of classical mechanics.[9] Only this approach can assure a sufficient *solid* framework for dealing with the foundations of biology and cognition. The specific problem that we are concerned with in cognitive biology is to find an opportune mediation between the theory of complexity (helpful for biology) and the quantum-mechanical treatment of information (helpful for cognition). As we shall see this connection can be found in the notion of *information control* that bridges metabolic processes and informational aspects. It is understood that I am not claiming that quantum mechanics or the theory of complex systems offers a direct solution to biological and cognitive issues. However, those physical disciplines provide us with a general conceptual and methodological framework that is useful for dealing with these problems: My principal aim in this book is to show that the notion of information control implies a shift from explanations that are centered on single trajectories, properties, behaviors and systems, to explanations focused on *classes* of behaviors and systems.[10]

Such an aim cannot be accomplished without simplifying strategies[11]; however, these must avoid opposing and risky oversimplifications:

(1) The elimination of the specificity of any level of complexity involved in the investigation, which can be the consequence of assuming that the same basic laws apply to, and the same structures form all domains of reality. Therefore, the specific form that my methodology takes is to recognize that any system or aspect of reality under consideration is operationally complete,[12] that is, (a) its behavior is determined by laws that are specifically formulated for that level of reality and (b) lower levels of reality are integrated into those laws and structures at the

[5] [*JEANNEROD* 2006].
[6] It is likely that the first scholar to have understood this connection is Herbert Spencer [*REZNIKOVA* 2007, p. 4].
[7] [*WOODGER* 1929, p. 84]. [8] [*WUNDT* 1873–4] [*VON HELMHOLTZ* 1867]. [9] [*MACH* 1875, pp. 6–22].
[10] [PARISI 1999]. [11] [*ROBERT* 2004, p. 3]. [12] [*NEWELL* 1990, 49].

higher level, assuring the ability to work in the proper way (for instance, no biology without chemistry).

(2) When investigating a specific domain, the risk is to completely isolate it from its contingencies connected with the rest of reality. Indeed, in any work like this the different stages considered are neatly separated and abstractly distinguished. In fact, things are much more entangled and the evolution of species on the Earth shows many mixed cases and exceptions. This fact, however, does not diminish the importance of a systematic treatment of the matter at the specific level. A tension between the latter approach and a strict empirical or historical one is natural and even salutary.

Therefore, much of what I shall say is hypothetical. I will explicitly stress this point many times. However, the book tries to gather some of the best hypotheses that we can advance given the current data at our disposal. A significant number of concepts introduced here have an empirical basis and will be presented in this way. On the other hand, it is very difficult to proceed in science in a fruitful and long-ranging perspective without formulating hypotheses.[13] If this investigation is of any interest, perhaps further research will fill some gaps, reformulate and even disprove some assumptions. This is what I expect and hope. A research program will be judged not by the labels one attaches to it but by the practical effects that it will eventually produce. In the words of Francis Bacon:

> Qui tracteverunt scientia aut Empirici aut Dogmatici fuerunt. Empirici, formicæ more, congerunt tantum et utuntur; Rationales, aranearum more, telas ex se conficiunt: apis vero ratio media est, quæ materiam ex floribus horti et agri elicit, sed tamen eam propria facultate vertit et digerit. Neque absimile philosophiæ verum opificium est; quod nec mentis viribus tantum aut precipue nititur, neque ex historia naturali et mechanicis experimentis præbitam materiam, in memoria integram, sed in intellectu mutatam et subactam, reponit.[14]

Some of the main results of many disciplines have been used in order to provide the necessary theoretical framework and empirical evidence for the necessity of new foundations of life and cognitive sciences. My work could therefore be defined as a *cross-disciplinary* enterprise, i.e. as an investigation connecting various and even distant fields with the aim of finding specific results and indicating a general research path. My own work strongly relies on recent results that have been developed in these different fields. What I have done is to establish connections between different approaches and problems as well as to infer some theoretical conclusions in order to make the scientific community aware that we have already entered into a new stage of research and even set the presuppositions for a new research strategy. It should be clearly understood that it has not been my aim to present or to develop these research areas on their own (which is probably an impossible task). Indeed, the specialist may find that the empirical material adduced in his or her field is not sufficient or has been selected by criteria that he or she does not share. In the words of Martha Farah,[15] the story of this book "can be told from start to finish only if one is willing to accept educated guesses at some junctures and frank declarations of ignorance at others."

[13] [PEIRCE 1898] [*BERNARD* 1865, pp. 35–47].

[14] Those who have handled sciences have been either men of experiment or men of dogmas. The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes a middle course: it gathers its material from the flowers of the garden and of the field, but transforms and digests it by a power of its own. Not unlike this is the true business of philosophy; for it neither relies solely or chiefly on the powers of the mind, nor does it take the matter which it gathers from natural history and mechanical experiments and lay it up in the memory whole, as it finds it, but lays it up in the understanding altered and digested [*BACON* 1620, I, 95].

[15] [*FARAH* 2000a, p. 1].

In connecting so many different fields, one of the main tasks of this book is to help in providing a fundamental conceptual clarification about many terms that are currently used in cognitive sciences (this is less true for biology where the terminology is much more fixed). Indeed, many terms are used in different ways in adjacent sciences (that somewhat intersect), and even in the context of a single science, they are mostly used in a slippery way. For instance, words like "semantics," "reference," "representation," "image," "concept," "category,"[16] and so on, even in specialized sciences like the neurological ones and by famous scientists, are used in a questionable form, sometimes covering aspects and issues that are not only different but often unrelated. For these reasons, I have also changed the meaning of many words as they are currently used in some disciplines, and I have introduced new terms. There may indeed be objections to many terms I have coined; I will not claim that my terminology is the definitive one, but I would ask that the reader appreciate my effort at univocity. In this effort of clarification, I have also sometimes changed the terminology that I have used in my own previous papers and books. The glossary is useful whenever difficulties arise and in such cases I suggest that the reader refer to it.

Finally, for the sake of clarity I shall give a brief overview of the structure of the book. The order of the topics follows pragmatic principles, so that a deep and easy understanding may be acquired. The first part is centered on the concept of information acquisition. Chs. 1–2 are rather abstract and introductory but provide the general framework in terms of information and system theory, presenting the basic notions which we shall deal with later. Chs. 3–5 provide some fundamental elements of information-acquiring especially regarding the human brain, considered here solely in its function as an information-acquiring device. These elements will turn out to be relevant for many issues treated in the second part of the book, in particular in Chs. 12–17.

The second part of the book is about information control and related semiotic processes. Ch. 6 is again an abstract and fundamental treatment of basic physical notions dealing with self-organization and complexity. Ch. 7 gives a basic notion of the cell, while Ch. 8 introduces the notion of the organism as a cybernetic system. Chs. 9–11 are an analysis from an informational and systemic point of view of the three basic processes of life: Phylogeny, ontogeny, and epigeny. Ch. 12 represents a new start, dealing extensively with the representational processes of life. Chs. 13–14 come back to an analysis of the brain, but this time treated as an information-control system rather than as an information-acquiring device, as it was still considered in the first part of the book. Chs. 15–17 deal with the three fundamental functions of behavior, learning, and memory.

The third part of the book is about information interpretation and symbolic activity. Ch. 18 is devoted to conceptualization, the problem of choice and empathy. In Ch. 19 a basic treatment of symbolic activity is provided. Chs. 20–22 are devoted to the issues of intentionality, consciousness, development, and culture. Ch. 23 deals with language, while a short examination of the mind–body problem follows in Ch. 24. Finally a concluding chapter will draw some main philosophical lessons.

[16] [THOMPSON/ODEN 2000].

Part I

Acquiring Information

One of the prime characteristics of science is . . . its success in linking together diverse strands of nature that had previously seemed unrelated.

[*STAPP* 1993, p. 120]

It is not possible to do the work of science without using a language that is filled with metaphors. Virtually the entire body of modern science is an attempt at explaining phenomena that cannot be experienced directly by human beings . . . Such explanations, if they are to be not merely formal propositions, . . . must necessarily involve the use of a metaphorical language.

[*LEWONTIN* 2000, p. 3]

Facts are always examined in the light of some theory and therefore cannot be disentangled from philosophy.

[*VYGOTSKY* 1986, p. 15]

1

Quantum Mechanics as a General Framework

When dealing with the relations between biology and physics, three paths are theoretically possible: (1) To rely on classical physics and the old reductionist methodology. This has led to very important results and still will, but (as I hope this book will show) such an approach appears to be increasingly insufficient and not in accordance with the developments of physics. (2) To reject any connection with physics in order to guarantee an autonomous foundation of biology. This is scientifically very questionable and leads finally to a form of vitalism. (3) The third possibility is to show that the most revolutionary physical theory, i.e. quantum mechanics, allows a connection with physics that is both much more interesting for biology and sufficiently soft to allow an autonomous foundation *without* any violation of physical laws. I have chosen this third perspective. This justifies, I hope, the present and the next chapter. In the next part of the book I shall also consider another important contribution coming from physics, namely the theory of complex systems. Although crucial for the understanding of the topic of this book, the latter theory is less general than quantum theory and therefore less indicated as a start.

In this chapter we shall deal with quantum physics at a conceptual level. After having explained the main reasons for starting with a summary of quantum mechanics in a book about biological and cognitive processes, I shall try to summarize what are the main characters of quantum theory and which are the main general lessons. We shall discover that a quantum system shows not only well-localized properties but also nonlocal features.

1.1 Why Quantum Mechanics?

In some papers in biological sciences and in many in cognitive sciences one refers to physics as the science to which biological and neurological or psychological sciences should be reduced. Any attempt at developing an autonomous theory of biological and cognitive systems is often stigmatized as folk physics or folk psychology. As an extreme but interesting case of this form of ontological reductionism,[1] let us consider the example of G. E. Allen who affirmed that "Spemann's Nobel Prize should be revoked because organization did not seem to be a specific phenomenon describable in physical and chemical terms".[2] So, the problem we have still today is that of an autonomous (i.e. not reducible to physics and chemistry) foundation of biological sciences,[3] an issue to which I hope this book will make a contribution.

[1]The interested reader may have a look at [*MURPHY/BROWN* 2007, pp. 47–8] or [VAN GULICK 2007].
[2]Quoted in [*GORDON* 1999, p. 14]. See also [ANDERSON 1972].

Unfortunately, all those authors who speak of a reduction of biology to physics (or chemistry) have in mind classical physics (or chemistry), as if it were the ultimate paradigm of science. Actually, classical physics (as well as classical chemistry) has been overcome.[4] I will not deny that it still has many interesting applications and is very useful for solving a lot of practical problems, just as Ptolemaic astronomy is still used in navigation. However, classical mechanics is no longer the physical explanation of our world. It has been reduced to a sort of approximation. Quantum mechanics has replaced it as the ground physical theory.[5] According to Stapp,[6] there are two main revisions in physics: Heisenberg's uncertainty relations and the introduction of random events in physics. As we shall see, uncertainty relations are a consequence of quantum correlations. Anyway, both of these two fundamental revisions are due to quantum mechanics. These are the profound reasons why we need to start with such a theory.

It is true that, at a physical level, there are still a lot of problems in unifying quantum mechanics and relativity. However, any further progress in sciences (string theories, great unification, or further possibilities) will never bypass what has already been established on the firmest grounds by the most tested theory in history having the widest range of applications, i.e. by quantum mechanics. Moreover, it is already foreseeable that in the next 50 years quantum mechanics will progressively replace classical mechanics also in those macroscopic domains where a more detailed description becomes increasingly necessary.

I recall here that the first scientists to have suggested the necessity to find a connection between biology and quantum mechanics in order to build a new type of physical explanation more suitable for biology were some leading personalities in quantum mechanics like Bohr, Jordan, Delbrück, and especially Schrödinger.[7] Moreover, much later Roger Penrose remarked that a specific biological system such as the brain does not work like a classical computer.[8] Indeed, a brain can perform calculations even if the problem is not well-defined, in the sense that some of the data are lacking that are necessary for finding the solution through classical computational methods. He suggested that quantum mechanics could be the solution for this anomalous information treatment.[9] Even if it is not necessary to assume that brains directly follow quantum-mechanical laws (due to the very high complexity of their organization), the main idea was really revolutionary.

When we consider quantum physics, we are led to the conclusion that in this domain physical laws are very different from what one could expect in a classical framework.[10] Quantum laws do not rule single properties of physical systems but probability amplitudes and hence probabilities that these properties occur. By *property* I understand here the value of a physical quantity contributing to the description of a physical system's state (physical quantities are energy, time, speed, and so on). By *system*, in all its generality, I understand here a complex of interrelated elements that show some regularity in space or time. By *physical system*, I understand a system that can be (directly or indirectly) an object of experience through our senses. This means that quantum-mechanical laws rule our world on a global level (at the level of the possible general evolutions of the system that are described by those probabilities), but they do not determine singular and local events.[11]

[3][*GORDON* 1999, pp. 12–6]. [4][*SMOLIN* 1997, p. 25]. [5][AULETTA 2006c]. [6][STAPP 1992].

[7][*SCHRÖDINGER* 1944]. See [*SLOAN/FOGEL* 2009]. [8][*PENROSE* 1989, *PENROSE* 1994].

[9]We do not need to follow Penrose in all details as some aspects of his proposal may not be confirmed. They have actually been criticized by Tegmark [SEIFE 2000].

[10]Obviously, any formulation of laws in the framework of a theory cannot be understood as immediately representing the order of nature. They are the result of incomplete inferences and therefore at most fallible approximations. Nevertheless, to deny that they refer, although in an imperfect way, to such an order would be even worse. In the following I shall give some grounds for assuming that our inferences are not devoid of a certain objectivity.

[11][AULETTA 2006b].

This is completely different relative to classical physics where laws were thought of as ruling single events. This is a very interesting standpoint, as one of the first sciences to have provided a place for chance events in nature was biology, with the Darwinian theory of evolution.

Now, the fact that quantum mechanics does not rule single events has a very important consequence: There is room for the emergence of new types of physical systems. If, on the contrary, the world were strictly ruled by classical-mechanical laws, this would be impossible.[12] This state of affairs explains why, from the perspective of classical mechanics, living beings were always seen as exotic absurdities when not reducible to mechanical engines.[13] This is no longer the case when approaching the problem starting with quantum mechanics. As we shall see, it indeed provides some necessary conditions for a world in which life is possible.

1.2 What Quantum Mechanics Is

As we have seen, quantum systems do not behave in the same way as ordinary macroscopic bodies, neither they do follow the laws of classical physics. In particular, the whole of classical physics is based on two fundamental assumptions[14]: That any physical process and relevant parameter are continuous (continuity principle) and that all the properties of a physical system are determined (classically called *omnimoda determinatio*[15] or principle of perfect determination). Both these fundamental assumptions are violated in quantum mechanics: The principle of continuity by the quantization principle and the assumption of the perfect determination of the state by the superposition principle. Let us consider now the first violation.

1.2.1 Some Historical Remarks

The first time that an explanation violating the continuity of the classical world was proposed was in 1900, when Planck[16] introduced the idea that the energy of the resonators (here, the elements of matters emitting electromagnetic radiation) of an internal surface of a so-called black body (a hollow sphere that only absorbs but does not emit radiation) could be discontinuous.[17] Classical physics is not able to correctly predict the spectral properties of the black body. In fact, it predicts an infinite total intensity of the emitted radiation. This situation is called *ultraviolet catastrophe*, and contradicts experimental evidence. In order to solve this problem, Planck assumed that the energy of the resonators and of the electromagnetic radiation emitted was quantized (discrete). Planck's solution was the result of an *abduction*, which is the inferential process leading to a solution of a conflict between accepted laws and experimental evidences by pointing out new properties of the system under consideration,[18] which, although representing a novelty relative to any behavior previously observed under those laws, do not contradict the latter. In this case, the new property is the *quantization* of the emitted radiation. Planck's assumption may be formulated in a general

[12]Obviously, I am speaking of a classical world in which there is no complexity or chaos. It is true that these aspects have been successively englobed in the classical framework of physics. However, as a matter of fact the strong dependence of complex and chaotic systems on the initial conditions does not agree very well with the classical-mechanical assumption of the complete reducibility of these initial conditions to classical-mechanics laws.

[13][*SMOLIN* 1997, p. 25].

[14][AULETTA 2004]. For a whole examination of this section the interested reader may deepen their knowledge by making use of the textbook [*AULETTA et al.* 2009]. That book actually represents the foundations of all that I shall say in this chapter.

[15][*BAUMGARTEN* 1739, Par. 148]

[16][PLANCK 1900a, PLANCK 1900b].

[17][*KUHN* 1978] [*MEHRA/RECHENBERG* 1982–2001].

[18][PEIRCE 1878b, PEIRCE 1903d][AULETTA 2006a].

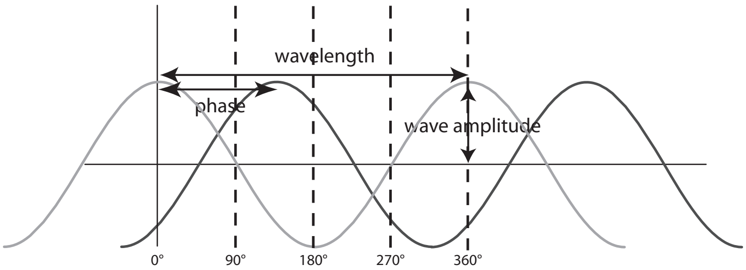

Fig. 1.1 Phase of a wave (on the right) relative to a reference wave (on the left) whose peak is individuated by the vertical line at the origin (0°). Since the behavior of a wave is oscillatory, it has a cycle (from 0° to 360° (or 2π), when it comes back to the initial situation). Its amplitude is the distance between a peak and the baseline. Its wavelength λ is the distance between two subsequent peaks and is inversely proportional to the frequency ν, that is, $\lambda = v/\nu$, where v is the velocity of the wave (which in the case of light can be taken to be a constant when going though void space).

form as follows: The energy E of the electromagnetic radiation in certain contexts can only assume discrete values for each of its frequencies ν, according to the formula

$$E = nh\nu, \tag{1.1}$$

where $n = 1, 2, \ldots$, and

$$h = 6.626 \times 10^{-34} \mathrm{J} \cdot \mathrm{s} \tag{1.2}$$

is the Planck constant that is expressed in terms of Joules times seconds. The frequency is inversely proportional to the wavelength—the distance between two peaks of a wave [Fig. 1.1].

This assumption was later used by Albert Einstein, who introduced the idea that the light could consist not of waves but of small, particle-like quanta (later called photons), to explain the photoelectric effect, i.e. the emission of electrons by a metal surface illuminated by light.[19] In fact, the classical picture of light in terms of waves was unable to account for the photoelectric effect: For a wave, a certain amount of time (of the order of several seconds) would be needed in order to deliver the energy required for the atoms to emit the electron. However, it is experimentally known that the effect is almost instantaneous. This is understandable only if one admits that light is made up of well-localized energy packets. It turns out then that the energy of photons occurs in quantized amounts, as proposed by Planck in 1900 as a solution to the black-body radiation problem.

This assumption was also, later on, applied by Niels Bohr to the atomic model.[20] In fact, in a continuous framework (like that proposed by Rutherford), negative-charged particles (electrons) revolving around a positive-charged nucleus (made of protons) would fall in a short time into a spiral trajectory into the latter. However, this would contradict the experienced stability of matter, which is, on the contrary, accounted for by discrete (quantized) energy levels.

I wish to stress, in conclusion, that the quantization assumption ran in contradiction with the whole understanding of physics at that time. Indeed, at the threshold between the 19$^{\mathrm{th}}$ and 20$^{\mathrm{th}}$ centuries a continuist view was dominant. All phenomena were essentially understood as local

[19] [EINSTEIN 1905]. [20] [BOHR 1913].

978–0–19–960848–5 01-Auletta-01-drv Auletta (Typeset by SPi) 11 of 869 June 13, 2011 14:45

OUP UNCORRECTED PROOF – REVISES, 13/6/2011, SPi

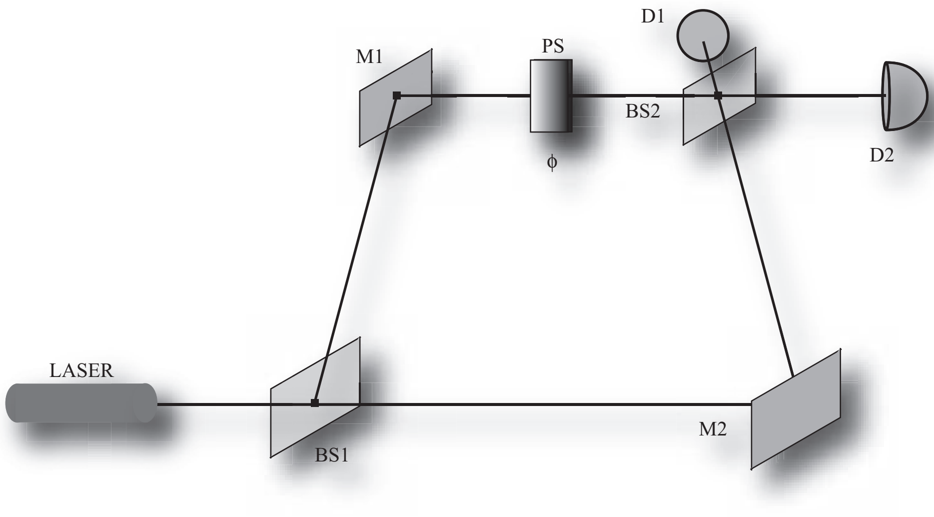

Fig. 1.2 Schematic setup of the Mach–Zehnder interferometer (top-down lateral view). The source beam coming from the laser is split at the first beam splitter BS1. After reflections at the two mirrors M1 and M2, the upper and lower paths are recombined at the second beam splitter BS2 and then detected at the photodetectors D1 and D2. PS denotes a phase shifter which causes a relative phase shift ϕ of the upper beam.

propagation phenomena.[21] Light, in analogy with acoustic phenomena, was understood as the local vibration of elements of a medium called aether, so that nothing was really in motion.[22]

1.2.2 The Superposition Principle

Let me stress that, being the quantum of light, a photon can be absorbed and emitted by single atoms. As a consequence, photons can be detected in well-defined positions, as happens for ordinary particles, by certain apparata (called *photodetectors*). It is worth mentioning that in optimal conditions a single photoreceptor (rod) of a human eye is able to detect a single photon[23] and therefore to function as a photodetector (even though with a small efficiency).

To see that and to infer another basic aspect of quantum mechanics let me describe an ideal experiment. The setup is shown in Fig. 1.2 and is known as a Mach–Zehnder *interferometer*. It essentially consists of two beam splitters, i.e. two half-silvered mirrors which partly reflect and partly transmit an input light beam, two mirrors, and two photodetectors. All the devices present in this setup are *linear*, i.e. such that the output is proportional to the input. Linearity is a mathematical property given by the combination of *additivity*

$$f(x + y) = f(x) + f(y), \tag{1.3a}$$

[21] Today general relativity supports this understanding of gravitational force.
[22] [*VON HELMHOLTZ* 1883, pp. 593–4]. [23] [*HUBEL* 1988].

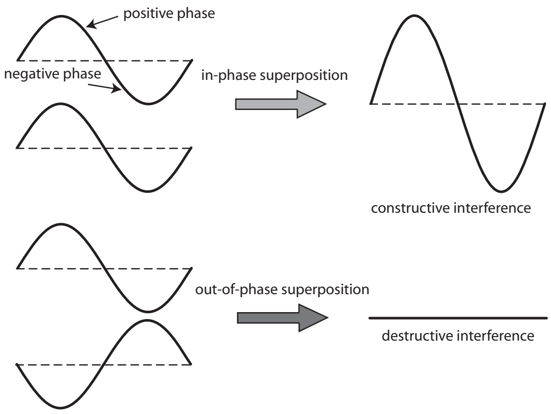

Fig. 1.3 Constructive and destructive interference depending on the phases of the photons. In the former case they are in phase, in the latter case they are completely out of phase (intermediate possibilities result in intermediate interferences).

for some function f and variables x and y, and of *homogeneity*

$$f(n \cdot x) = n \cdot f(x), \tag{1.3b}$$

where n can be any scalar number (multiplier).

A light beam coming from the source is split by the first beam splitter (BS1) into the upper and lower paths. These are reflected at the mirrors M1 and M2 and recombined at the second beam splitter BS2, placed before the photodetectors D1 and D2, which we assume to be ideal, i.e. with 100% efficiency. In the upper path a phase shifter (PS) is inserted in order to create a relative phase difference ϕ between the two component light beams. A phase shift which is a multiple of 2π (360°) brings the situation back to the original one, while a phase shift $\phi = \pi$ corresponds to the completely out-of-phase situation [Fig. 1.1]. At BS2 the two beams interfere and such interference may be destructive ($\phi = \pi$) or constructive ($\phi = 0$) [Fig. 1.3]. For example, destructive interference at D2 means that the observed intensity at the photodetector is equal to zero (*dark output*). This in turn means that D1 will certainly click (here there is constructive interference). The transmission and reflection coefficients T and R of the beam splitters can both vary between 0 and 1, with $\mathrm{R}^2 + \mathrm{T}^2 = 1$. When $\mathrm{T} = \mathrm{R} = 1/\sqrt{2}$, we have a $50\% - 50\%$ beam splitter.

We then see that, to a certain extent, photons still behave as classical light, that is, have wave-like properties. Since we have also seen that light can be considered as composed of corpuscular entities, this suggests a picture according to which light may display both *wave-like* and *corpuscular* aspects. We face here a new and surprising situation that appears paradoxical from a classical viewpoint. In the next subsections I shall try to explain this state of affairs and draw the necessary consequences. Let us first imagine what happens when we send a single photon at a time through the Mach–Zehnder interferometer. At each experimental run the photon will be detected either at D1 or at D2. This is because experimental evidence shows that the photon cannot be divided, which in turn means that to our knowledge the photon is an elementary (and discontinuous) entity. However, after many runs, which are required in order to obtain good statistics, we will experimentally

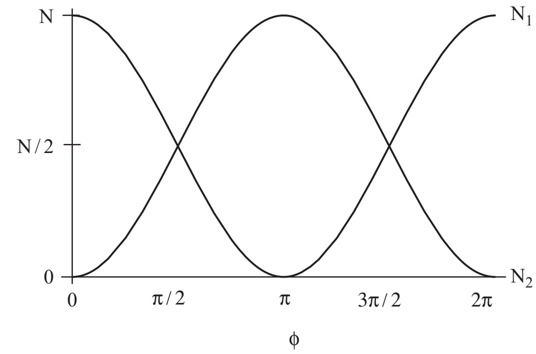

Fig. 1.4 The two curves show the statistical results of photon counting at detectors D1 and D2. N_1 and N_2 denote the number of photons counted at detectors D1 and D2, respectively. It should be noted that, for each value of ϕ, obviously $N_1(\phi) + N_2(\phi) = N$.

observe that the detector D1 will click $N(1 + \cos\phi)/2$ times and detector D2 $N(1 - \cos\phi)/2$ times, where N is the total number of experimental runs. Repeating the same experiment for a large number of times with different values of ϕ, we would obtain the plots shown in Fig. 1.4. This behavior represents a typical maximal interference. Since at most one photon at a time is present within the apparatus, one can speak of *self-interference* of the photon. This is something very surprising, needing some further consideration.

Because the photon cannot be split, self-interference forces us to admit that the photon is *not localized* in either of the two arms. For the sake of clarity, let us suppose that we remove BS1. Then the photon will certainly travel along the lower path (it is fully transmitted). We can label the state of the photon in such a case by the symbol $|\, l\rangle$, where l denotes the lower path. In the following, I will use the symbol $|\cdot\rangle$ to indicate the state of a physical system. On the other hand, if BS1 is replaced by a 100% reflecting mirror, the photon will take with certainty the upper path and its state may then be denoted in this case by the symbol $|\, u\rangle$, where u denotes the upper path. As a consequence, when the half-silvered mirror BS1 is put in its place, we are led to the conclusion that the state of the photon should be a combination (a *superposition*) of both the states $|\, l\rangle$ and $|\, u\rangle$ associated with the two arms of the interferometer. Therefore, we state in general terms the first theoretical consequence of our ideal experiment:

If a quantum system $\mathcal{S}$ can be in either of two states, then it can also be in any linear combination (superposition) of them.

Due to linearity [Eqs. (1.3)], inputs are added and multiplied in a way that holds also the output to be proportionally added and multiplied. In the example above, the total state of the photon after BS1 that we may conventionally denote with $|\, \psi\rangle$ can be expressed as the superposition

$$|\, \psi\rangle = c_u \,|\, u\rangle + c_l \,|\, l\rangle\,, \tag{1.4}$$

where c_u and c_l are some coefficients whose meaning will be explained below. Eq. (1.4) represents the above conclusion: It is not possible to assign a well-defined path to the photon, but this takes a combination of the two paths and is delocalized. We should emphasize that this state of affairs is a clear violation of the classical principle of perfect determination according to which the photon should be *either* in the upper path *or* in the lower path. This means that Eq. (1.4)—describing a

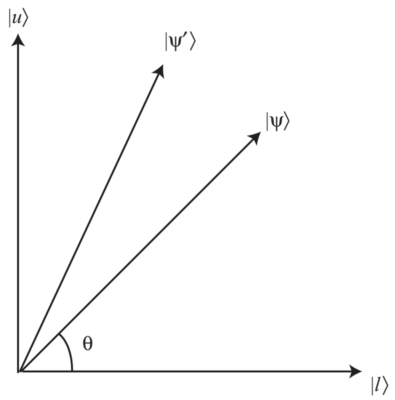

Fig. 1.5 The quantum state $|\psi\rangle$ is a vectorial sum of $|u\rangle$ and $|l\rangle$. The contribution of each basic state vector is a function of the angle θ.

superposition *of states*—cannot be interpreted as a superposition of classical waves, since, in the latter case, the components of the superposition would be two *different* spatial waves, whereas in the case of Eq. (1.4) the components $|l\rangle$ and $|u\rangle$ represent *possible states* of the *same photon*. Therefore, the wave-like properties of the photon discussed above cannot be taken in any classical sense. For that reason, the superposition principle is not just a consequence of our ignorance of the path *actually* taken by the photon.

Summing up, quantum mechanics corrects classical mechanics on two crucial points[24]: (a) There is not a continuity of motion and of the relevant physical parameters when systems interact, and (b) material elementary (local) entities are not identifiable.

1.2.3 Vectorial Nature of Quantum States

Quantum states can geometrically be represented as *state vectors* in a given vectorial space (called a Hilbert space). For instance, the superposition (1.4) may be considered as a vectorial sum of two basic vectors, $|u\rangle$ and $|l\rangle$ [Fig. 1.5]. These vectors can be described as

$$|l\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |u\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.5}$$

which corresponds to the x and y vectors of the Cartesian plane.

On the other hand, the coefficients c_u and c_l will represent the weights with which the two component vectors participate in the resulting state vector $|\psi\rangle$. In other words, they can be understood as functions of the angle θ that separates the vector $|\psi\rangle$ from $|u\rangle$ or from $|l\rangle$, as shown again in Fig. 1.5. Therefore, allowing us to write any arbitrary state vector as their sum, these two vectors are called a *basis*, and this is written as $\{|l\rangle, |u\rangle\}$. In the bidimensional case (with two basis vectors), we may then represent the coefficients as sine and cosine as follows [Fig. 1.6]:

$$|\psi\rangle = \sin\theta\,|u\rangle + \cos\theta\,|l\rangle . \tag{1.6}$$

[24][*MARGENAU* 1961, p. 131].

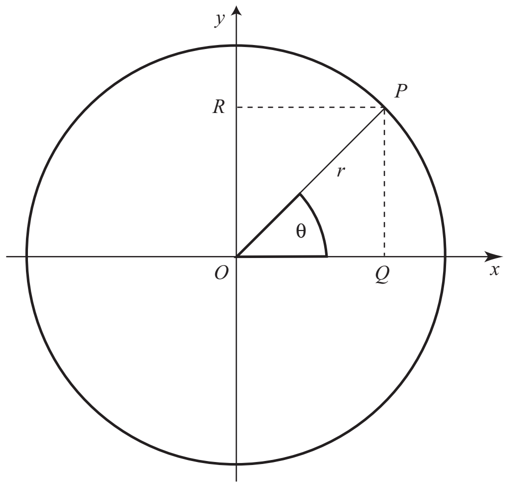

Fig. 1.6 Representation of sine and cosine. The sine of the angle θ is given by PQ/r, where r is the ray of the circumference: It is 0 for $\theta = 0°$ ($PQ = 0$), 1 for $\theta = 90°$ ($PQ = r$), 0 for $\theta = 180°$, and -1 for $\theta = 270°$ (because we are considering here the negative part of the y axis). The cosine is given by PR/r (or OQ/r). It is equal to 1 for $\theta = 0°$ (since $OQ = r$), 0 for $\theta = 90°$ ($OQ = 0$), -1 for $\theta = 180°$, and 0 for $\theta = 270°$. If we plot sine and cosine according to those values, we shall obtain two curves like those shown in Fig. 1.4, where N1 would represent the cosine and N2 the sine.

In Fig. 1.5, $|\psi\rangle$ is depicted as a symmetric vector, i.e. $\theta = 45°$ and therefore, in this case, we have (since both the sine and cosine of $\theta = 45°$ is $1/\sqrt{2}$)

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|u\rangle + |l\rangle). \tag{1.7}$$

However, the contribution of $|u\rangle$ and $|l\rangle$ may also be asymmetric, as happens for the vector $|\psi'\rangle$, also shown in Fig. 1.5. Obviously, the dimensions of a system may be greater than 2. These dimensions can be understood as the possible outcomes that we obtain when we measure a system. In other words, they must be understood as the alternative properties that a quantum system can instantiate; for example, to be localized in the upper path ($|u\rangle$) or to be localized in the lower path ($|l\rangle$) of an interferometer.

Summarizing, the vector $|\psi\rangle$ can be written

$$|\psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{or in the general form} \quad |\psi\rangle = \begin{pmatrix} c_l \\ c_u \end{pmatrix}, \tag{1.8}$$

which allows us to write the coefficients c_l and c_u as

$$c_l = \langle l \,|\, \psi\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} c_l \\ c_u \end{pmatrix} \quad \text{and} \quad c_u = \langle u \,|\, \psi\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} c_l \\ c_u \end{pmatrix}, \tag{1.9}$$

where $\langle l \,|\, \psi\rangle$, e.g., is the *scalar product* between the vectors $|l\rangle$ and $|\psi\rangle$. The scalar product can be computed in this way: The row vector on the left multiplies the column vector on the right (the first element in the row with the first element in the column and the second element in the row with the second element in the column), and then these results are summed, so as to obtain: $1 \cdot c_l + 0 \cdot c_u = c_l$. Similarly for the second scalar product: $\langle u \mid \psi \rangle = 0 \cdot c_l + 1 \cdot c_u = c_u$. The scalar product indicates the amount of proximity of these two vectors (indeed, when the two vectors coincide it is equal to 1, and when the two vectors are orthogonal it is equal to 0), explaining, in this way, at a general level, what is expressed in Fig. 1.5 with the angle θ.

1.2.4 Quantization Principle

We have seen that the energy of a quantum system can be quantized in certain circumstances.[25] Indeed, Planck assumed that otherwise continuous, radiation, when it was emitted by the black body's surface, was quantized [Subsec. 1.2.1]. Also Einstein's photoelectric effect concerns a certain interaction with light and matter. This shows that light, when interacting with matter, displays discontinuous aspects, and this is a key to our understanding of quantum mechanics. Indeed, such a situation can be generalized in a quantization principle:

When several quantum systems interact, they manifest discontinuous properties.

I wish to stress that not only energy can behave in this way, but also other physical quantities describing a quantum system (like speed), apart from those (like charge) that are intrinsically discontinuous. A consequence of this situation is that it is impossible to represent physical quantities in quantum mechanics with mathematical variables. Variables are in fact continuous whereas the behavior of quantum physical quantities can be both continuous or discontinuous. Mathematical entities able to represent this situation are *operators*, whose spectrum (the set of possible values of the corresponding physical quantity that one would obtain by measuring it) can indeed be continuous, discontinuous, or a combination of both.

An operator can be understood as the mathematical representation of an operation, e.g. a rotation. It is well known that the specific sequence in which several operations are executed does matter. For instance, if one walks and first turns left and then right, one would arrive at a different point than if one first turned right then left. In other words, the result of a sequence of operations will be different if the *order of the operations is changed*. In the general case, a sequence of operations is mathematically represented by a product of the operators representing the single operations. In our case, we have

$$\hat{L}\hat{R} \neq \hat{R}\hat{L}, \tag{1.10}$$

where $\hat{R}$ and $\hat{L}$ represent the operation of turn-to-right and turn-to-left, respectively, and the hat denotes here and in the following any operator. Eq. (1.10) represents a very interesting mathematical property, called *non-commutativity*.[26] In fact, variables and functions of variables (that represent classically physical quantities) do commute, that is, the product between any two variables or functions of variables is indifferent to the order of the factors, which means that, for any arbitrary couple of numbers a and b, we always have

$$ab = ba. \tag{1.11}$$

On the contrary, given two arbitrary quantum observables $\hat{O}$ and $\hat{O}'$, we may have

$$\hat{O}\hat{O}' - \hat{O}'\hat{O} \neq 0. \tag{1.12}$$

[25] [SCHRÖDINGER 1926a, SCHRÖDINGER 1926b, SCHRÖDINGER 1926c].
[26] A property first discovered by Heisenberg [HEISENBERG 1925].

The quantum physical quantities will be called in the following "observables" in order to distinguish them from the classical variables.

Let us come back to the formal aspects. Any (finite-dimensional) operator can be written as a matrix; for instance, in the bidimensional case, we will have

$$\hat{O} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{1.13}$$

which represents a generalization of the concept of vector (we have now not only a column but also rows). All the elements of a matrix (a, b, c, d, here) are numbers. The sum and product of bidimensional (with 2 columns × 2 rows) matrices $\hat{O}$ and $\hat{O}'$ is given by

$$\hat{O} + \hat{O}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a + a' & b + b' \\ c + c' & d + d' \end{bmatrix} \tag{1.14}$$

and

$$\hat{O}\hat{O}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}, \tag{1.15}$$

respectively. The result of the product is given by four elements, which respectively represent the first row times the first column, the first row times the second column, the second row times the first column, and the second row times the second column. A matrix can have complex elements. However, a matrix representing an observable must have real values, which we impose therefore as a requirement.

An example of operator is represented by projectors, for instance $\hat{P}_u = |\, u\rangle \langle u \,|$ and $\hat{P}_l = |\, l\rangle \langle l \,|$; taking into account expressions (1.5), we have

$$\hat{P}_l = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{1.16}$$

$$\hat{P}_u = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.17}$$

where we have the first row times the first column and times the second column, and the second row times the first column and times the second column. Note that the order of row and column matrices is inverted relative to the scalar product (1.9). The action of these projectors on a state vector (a ket, in short)

$$|\, \psi\rangle = c_u \,|\, u\rangle + c_l \,|\, l\rangle \tag{1.18}$$

is shown in Fig. 1.7 and can be mathematically described as

$$\begin{aligned} \hat{P}_u \,|\, \psi\rangle &= |\, u\rangle \langle u \,| \left(c_u \,|\, u\rangle + c_l \,|\, l\rangle \right) \\ &= c_u \,|\, u\rangle \langle u \,|\, u\rangle + c_l \,|\, u\rangle \langle u \,|\, l\rangle = c_u \,|\, u\rangle \cdot 1 + c_l \,|\, l\rangle \cdot 0 \\ &= c_u \,|\, u\rangle , \end{aligned}$$

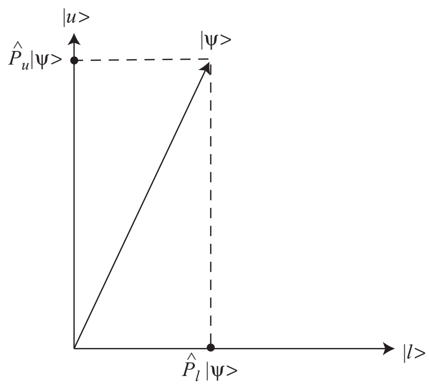

Fig. 1.7 Action of projectors. We have indeed $\hat{P}_u\,|\,\psi\rangle = c_u\,|\,u\rangle$ and $\hat{P}_l\,|\,\psi\rangle = c_l\,|\,l\rangle$, which nicely shows that the projection on $|\,u\rangle$ and $|\,l\rangle$ is shorter than $|\,u\rangle$ and $|\,l\rangle$ themselves, since (for reasons that will be explained at the beginning of Subsec. 1.2.7) both c_u and c_l are ≤ 1 (when c_u or c_l is equal to 1, the other coefficient is zero: In that case we are projecting a vector onto itself, e.g. $\hat{P}_u\,|\,u\rangle = |\,u\rangle$).

$$\begin{aligned}\hat{P}_l\,|\,\psi\rangle &= |\,l\rangle\,\langle l\,|\,(c_u\,|\,u\rangle + c_l\,|\,l\rangle\,) \\ &= c_u\,|\,l\rangle\,\langle l\mid u\rangle + c_l\,|\,l\rangle\,\langle l\mid l\rangle = c_u\,|\,l\rangle\cdot 0 + c_l\,|\,l\rangle\cdot 1 \\ &= c_l\,|\,l\rangle\,, \end{aligned} \tag{1.19}$$

where I have made use of the properties of the scalar product $\langle u\mid l\rangle = \langle l\mid u\rangle = 0$ and $\langle u\mid u\rangle = \langle l\mid l\rangle = 1$.

The fact that operators representing quantum physical quantities (energy, momentum, i.e. speed times mass, and so on) may not commute has some interesting consequences. The most important are known as the uncertainty relations, derived and stated for the first time by Heisenberg,[27] which define the minimum value of the product of the uncertainties of two conjugate observables. Both in classical physics and in quantum mechanics, conjugate variables or observables are the following couples: Position and momentum, time and energy, angle and angular momentum, and so on. Momentum, energy, and angular momentum are called *dynamic* variables or observables, while position, time, and angle are called *kinematic* variables or observables. The *uncertainties* Δx and Δp_x of the one-dimensional position and momentum operators $\hat{x}$ and $\hat{p}_x = mv_x$ (mass m times the velocity v_x along the direction x), respectively, calculated on the state $|\,\psi\rangle$, obey the uncertainty relation

$$\Delta_\psi x \Delta_\psi p_x \geq \frac{\hbar}{2}, \tag{1.20}$$

where $\hbar = h/2\pi$. Needless to say, similar expressions hold for the y and z components. Moreover, analogous uncertainty relations can be written also for other pairs of conjugate observables. The relation (1.20) states that, when one tries to reduce the uncertainty of one of the two conjugate observables, then necessarily the uncertainty of the other increases. In particular, it is possible to have an infinitely precise measurement of one of the two observables, say the position ($\Delta_\psi x = 0$).

[27][HEISENBERG 1927].

But, in this case, the price one has to pay is that the momentum observable is completely undetermined ($\Delta_\psi p_x = \infty$). The value $\hbar/2$ represents then the maximum attainable certainty, i.e. the *minimum uncertainty* product allowed by the uncertainty relations. Let us consider this situation a little.

1.2.5 Features and Quantum State in Phase Space

It is important to stress that, exactly as it happens for classical-mechanical states, we have a description of the state of a quantum-mechanical system represented by the state vector $|\psi\rangle$. That is, the vector $|\psi\rangle$ contains *everything* we may know about the system, i.e. it represents the maximal amount of information about the system at a given moment. However, while in classical mechanics, *all* properties of a physical state are believed to be simultaneously instantiated and therefore can also be in principle *jointly* measured, the uncertainty relations *forbid one to acquire all the information* that is contained in the quantum state vector $|\psi\rangle$. In other words, although the quantum state also represents a stock of complete information about the system in a given moment, this information is not accessible in its totality to any observer and for any possible operation. We must conclude that, while classical mechanics is ruled by the principle of perfect determination and therefore a classical state is characterized by a complete collection of properties, the quantum mechanical state is intrinsically probabilistic and *affected by uncertainty*, i.e. not all the observables can be completely determined at the same time.

The question is why we have such a situation. We have already met the phenomenon of self-interference [Subsec. 1.2.2]. It is a manifestation of a deep quantum reality: Quantum correlations, that is, interdependencies among the components of the state (for instance, paths in the interferometer) or even the subsystems of some quantum system. They may be called *features* of a quantum system (quantum features or simply features when no ambiguity arises), because they are characters of the state of the system that need to be distinguished from true properties, which are always localized.[28] Then, features are responsible for any interference and non-local behavior of quantum systems, and therefore also for the impossibility of having access to the whole of information contained in a quantum state. This is expressed by the uncertainty relations. It is true that we arrived at uncertainties by starting from noncommutativity which in turn we took as a consequence of the interaction between quantum systems. However, we have such discontinuities precisely because the interacting quantum systems *display features*.

The concept of quantum state now has deep implications that can be examined using the phase-space representation of a system. A phase-space representation of a system is given by the representation of its state(s) in a reference frame whose axes are represented by position and momentum: In classical mechanics, position and momentum are the variables that determine the state of the system (at least in the most elementary cases). A classical representation in phase space is necessarily point-like. Indeed, when the principle of perfect determination is assumed, momentum and position always both have a perfectly determined value. If one considers the time evolution of the system, then the point representing the state of the system at a certain time will trace a well-defined trajectory in the phase space [Fig. 1.8].

On the contrary, due to the uncertainty relations, a phase-space representation for a quantum system at a given instant cannot be point-like: Such a representation must reflect the fact that the uncertainties in position and momentum are both finite and that their product cannot be smaller than $\hbar/2$. Therefore, we may depict this circumstance by a spot in the phase space whose horizontal

[28] [AULETTA/TORCAL 2011].

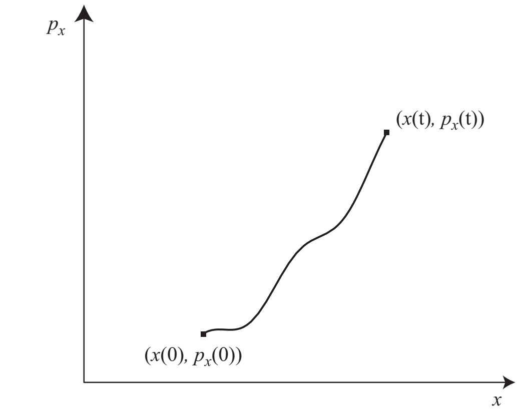

Fig. 1.8 Time evolution of a classical degree of freedom in phase space: At any time t, the state of the system is described by a point.

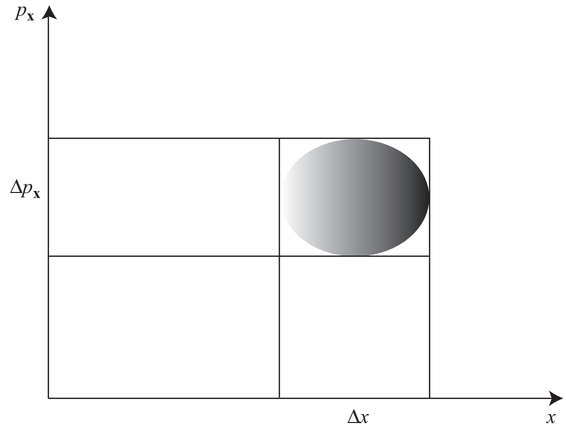

Fig. 1.9 Graphic representation of a "point" in the quantum-mechanical phase space. According to the uncertainty relations, a single degree of freedom should be represented by a spot. Shading has only aesthetic value.

and vertical dimensions are equal to the position and momentum uncertainties [Fig. 1.9]. Moreover, any improvement in the determination of momentum will be paid in terms of a proportional increase in uncertainty for position and *vice versa* [Fig. 1.10].

This has important methodological and philosophical consequences. In fact, since we cannot have perfect determination of two conjugate observables simultaneously, if we wish to know with great accuracy the value of one of the two, then we are obliged to *choose* between measuring position and measuring momentum. It is clear that quantum mechanics forces us to consider knowledge as a matter of choice rather than of a mirror image of a datum.[29] This completely changes the way

[29]On this point see [*WEYL* 1950, p. 76].

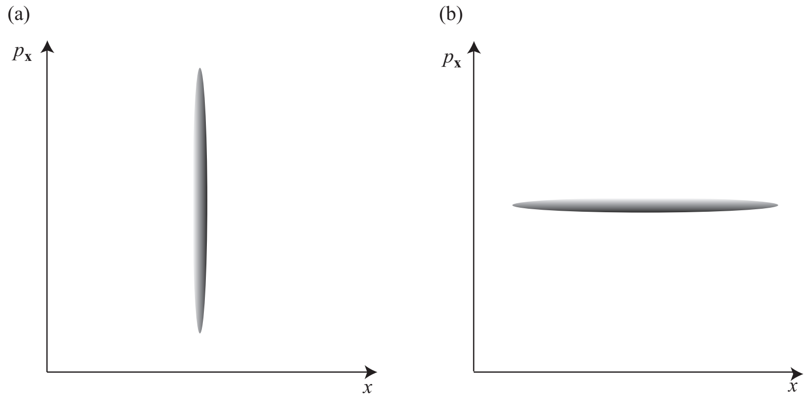

Fig. 1.10 Inverse proportionality between momentum and position uncertainties. When the position is accurately determined, the momentum becomes highly uncertain (a), and *vice versa* (b).

in which we consider our knowledge relative to the world. One might be led to the conclusion that quantum mechanics implies some form of subjectivism. However, this is not the case, as we shall see below.

Another consequence of this situation is that trajectories *do not exist* in quantum mechanics. This is true both in the phase space (for what we have said above) and in the configuration space (the space "in which" quantum systems actually move). In fact, if one could define a trajectory, say, of a one-dimensional particle, then it would also be possible to determine the velocity and therefore the momentum of the particle, violating the uncertainty relations. This has the consequence that quantum systems mostly follow a superposition of different trajectories, a multipath dynamics, which strongly undermines the possibility of making causal inferences.[30]

1.2.6 Complementarity Principle

It is clear from Subsec. 1.2.2 that, e.g., for $\phi = 0$, Detector D2 will never click. This *dark output* may even be used to detect the presence of an obstacle in one of the two paths without directly interacting with it. Let us place an object in the lower arm of the interferometer and set $\phi = 0$. Then, the presence of this object will prevent the interference at the second beam splitter (BS2), shown in Fig. 1.2, and allow, at least with some probability different from zero, that the photon will actually be detected at D2. This phenomenon is known as an *interaction-free measurement*[31]: We can state that, when the relative phase is set to $\phi = 0$, a detection event in D2 tells us with certainty that an object is in one of the two arms and that the photon has taken the other arm to the detector, i.e. it has now been localized. It should be noted that, in some cases, the photon will not be detected at all because it will be absorbed by the object. Still, in those cases when detector D2 clicks, we have learned about the presence of the object without directly interacting with it, something which classically would evidently not be possible.[32] As I have already anticipated, it is clear that interference (and its features) cannot be a manifestation of subjective ignorance: If this

[30] [AULETTA 2005a]. [31] [ELITZUR/VAIDMAN 1993].

[32] This has far-reaching consequences, if one thinks that the 1971 Nobel Prize winner in physics, Dennis Gabor, supported the idea that one cannot acquire information about an object system if at least one photon does not interact with it [KWIAT *et al.* 1996].

were the case, the presence or absence of features would not allow us specific detection events and the objective acquisition of information.

A further consequence is that every time that the photon is localized (i.e. we know with certainty that it has taken either the upper or the lower arm), interference is actually destroyed (since this is a direct consequence of the superposition of $|u\rangle$ and $|l\rangle$). In other words, we cannot acquire information about the path actually taken by the photon without *disturbing* the interference and *changing* the state of the photon itself. As I have said, interference is a manifestation of the presence of quantum *features*. Since, on the contrary, complete information about a quantum system can only be obtained through a measurement *event*, we are then forced to generalize the previous examination as follows

Events and features are complementary.

This principle states that to experience events and therefore to acquire information is complementary to the existence of global features: Information-acquiring implies lowering interference (which is due in turn to the existence of features).[33] A word of warning is necessary here: Complementarity is not a sharp yes/no alternative but rather a trade-off between partial gain of information and partial interference.[34] In other words, full localization ("particle-like" behavior) and full interference ("wave-like" behavior) are only limiting cases of a continuous range of behaviors. Therefore, quantum systems can neither be considered as classical particles, nor as classical waves.

1.2.7 Dynamics and Measurement

The complementarity principle [Subsec. 1.2.6] can be reformulated as a dynamic trade-off between the continuous and interference-like features displayed by the superposition principle [Subsec. 1.2.2] and the discontinuous properties displayed by the quantization principle [Subsec. 1.2.4]. In particular, when a quantum system is free, that is, when it does not interact with other systems, it displays a continuous behavior. When it interacts with other systems, it displays more or less discontinuous properties. Therefore, the complementarity principle tells us that, when two or more quantum systems interact, in the general case *both* quantum features and discontinuous properties are involved.

A particular case of dynamics is represented by measurement. Here, the most discontinuous situation occurs: A detection event [see again Subsec. 1.2.2]. To understand this process, let us consider the vectors $|u\rangle$ and $|l\rangle$ in Eqs. (1.4) or (1.6). They can be understood as possible outcomes when we perform a certain measurement of the path of a photon going through the interferometer shown in Fig. 1.2: The states $|u\rangle$ and $|l\rangle$ represent the upper and the lower paths taken by the photon, and since there is no other possible forms of localization in the interferometer (that is, the photon is found *either* in the upper *or* in the lower path when detected), these two states can be taken as the possible output states when the *position* of the photon is measured. However, what is crucial to understand is that in order to get one of these two possible outcomes we need to prepare the system beforehand in the state $|\psi\rangle$ (i.e. to let it go through a beam splitter). *Preparation* is therefore the first step of measurement.

The states associated with possible outcomes are called *eigenstates* or eigenvectors of the corresponding observable, that is, here $|u\rangle$ and $|l\rangle$ are eigenstates of the position observable of

[33] The complementarity principle was first formulated by Niels Bohr at the Como Conference in 1927 and communicated to a large audience in an article in *Nature* [BOHR 1928].

[34] [GREENBERGER/YASIN 1988].

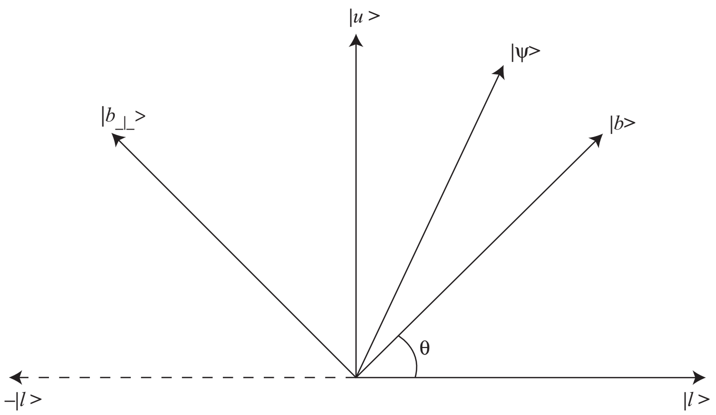

Fig. 1.11 Change of basis. The basis $\{|b\rangle, |b_\perp\rangle\}$ is obtained from the original basis $\{|u\rangle, |l\rangle\}$ by a counterclockwise rotation of 45°.

the photon. Moreover, the two coefficients c_u and c_l represent the probability amplitudes to obtain $|u\rangle$ or $|l\rangle$ as measurement outcomes. *Probability amplitude* (connected with the wave amplitude) means a quantity whose square modulus gives the relative probability (the classical intensity of the wave), i.e.

$$p(u) = |c_u|^2 \text{ and } p(l) = |c_l|^2 , \tag{1.21}$$

or simply the squares of the coefficients, if these are not complex numbers. The square modulus of a number c is the square of the absolute value of c, where by *absolute* it is understood that any minus sign before the *whole* expression c is dropped. The point now is that we could also have decided to measure another observable that is not commutable with position. How can we represent the eigenvectors of such an observable? We know (through the uncertainty relations) that in this case we face a choice. Indeed, the eigenvectors of this other observable can be defined by two vectors, say $|b\rangle$ and $|b_\perp\rangle$, that are different from $|u\rangle$ and $|l\rangle$ and represent therefore an *alternative* basis for the system being in the state $|\psi\rangle$ [Fig. 1.11].[35]

Summarizing, several observables may have distinct eigenvectors and each observable constitutes a different eigenbasis. This has an important consequence: Superposition is indeed a concept *relative* to a given observable. In fact, considering again Fig. 1.11, we immediately see that, for instance, the vector $|u\rangle$, which is an eigenvector of the position, represents simultaneously a superposition of the vectors $|b\rangle$ and $|b_\perp\rangle$, which are eigenvectors of the non-commutable observable we have introduced—let us recall Eq. (1.5). For this reason, when measuring an observable it is very important to know whether the state $|\psi\rangle$, in which the system has been prepared, is an eigenvector of the observable to be measured or a superposition of its eigenvectors. If it is already an eigenvector, then the measurement of a quantum system is an analogue of the measurement of a classical system. That is, the state of the system does not change as a consequence of measurement, and apparently here we limit ourselves to *register* the value that the measured observable already had given the system's state.[36]

[35] For a careful and technical account of these issues see [*AULETTA et al.* 2009, Ch. 9].
[36] The issue is a little bit more cumbersome [AULETTA/TORCAL 2011].

Let us suppose that this state is a certain (arbitrary) superposition $|\psi\rangle$ [Eq. (1.18)] of $|u\rangle$ and $|l\rangle$. If the prepared state is a superposition, the quantum theory of measurement shows features that have no classical counterpart. Indeed, when measuring in this case, we have a discontinuous jump from the initial superposition state $|\psi\rangle$ to a final outcome represented by either $|u\rangle$ or $|l\rangle$. The problem here is that it is impossible to have such a transformation at a general level according to the laws of quantum mechanics. All quantum-mechanical laws, which are in accordance with the superposition principle, are indeed reversible. These laws express a peculiar form of determinism because it is a determinism of probabilities. Indeed, provided that there is no detection event, we have a general delocalized state ruled by probability amplitudes [Subsec. 1.2.2]. Instead, the evolution from $|\psi\rangle$ to either $|u\rangle$ or $|l\rangle$, in all its generality, is irreversible: A detection event tells us that a system is, for instance, either in the state $|u\rangle$ or in the state $|l\rangle$, and this cannot be undone (for instance, a detected photon has been absorbed by the photodetector). The argument can be considered in this way. To have a reversible (a so-called unitary) transformation, it is like rotating a certain basis (from the initial basis $\{|u\rangle, |l\rangle\}$ to the basis $\{|b\rangle, |b_\perp\rangle\}$) of a certain angle. However, such a general transformation providing a rotation starting from *any* superposition and getting one of the two components does not exist, since it would be like rotating the axis of *any* possible angle with a single transformation (I recall that the coefficients c_l and c_u represents the angle of the relative components with respect to $|\psi\rangle$), which is clearly impossible. Resuming, the (superposition) state of a quantum system evolves in time and is ruled by deterministic and reversible laws (according to the superposition principle), while the *result* of a measurement is produced by an abrupt and irreversible change following the quantization principle.

Let us consider how we can obtain this result. After having prepared the system, the second step of measurement consists of establishing some interface between the input (initial) state $|\psi\rangle$ of the system and the final detection event. This is assured by an apparatus, which is coupled with the object system to be measured. In this case, each component of the superposition state of the system must be associated with a component of the apparatus, for instance

$$|\Psi_{\mathcal{SA}}\rangle = c_u |u\rangle |a_u\rangle + c_l |l\rangle |a_l\rangle , \tag{1.22}$$

where $|a_u\rangle$ and $|a_l\rangle$ are the components of the apparatus and $|\Psi_{\mathcal{SA}}\rangle$ describes the states of both the system and the apparatus. Such a coupling allows for the fact that, when having the system in the state, say, $|l\rangle$, the apparatus will be in the corresponding state $|a_l\rangle$. This coupling is called a *premeasurement*. In quantum mechanics, a state of this form is called an *entangled state* and shows characteristic interdependencies that are classically unknown (another manifestation of quantum features). Indeed, if the initial state of the system is a superposition of state of components $|u\rangle$ and $|l\rangle$, given this entanglement the apparatus will also be in a superposition state, namely of $|a_u\rangle$ and $|a_l\rangle$. This is bizarre, since we expect that an apparatus (being often a macroscopic body whose function is to provide a *determined answer*) will be either in the state $|a_u\rangle$ or in the state $|a_l\rangle$. The reason is very simple: The proper job of an apparatus is not only to be a coupling device, but to be able to faithfully indicate which state the system is in. Now, an apparatus in a superposition state indicates no determinate value, and since we have seen that there is apparently no possible transition from a superposition to one of its components, this situation is really puzzling.

However, we have also made an important step. Indeed, the premeasurement (i.e. the experimental context through which we have established a specific connection between an apparatus and a given observable) will at least ensure that *one* specific observable (that is, a basis) is chosen among many ones. It can indeed be shown[37] that either the apparatus cannot work properly (there

[37] [ZUREK 2007].

is no possible transfer of information from the object system to the apparatus) or there is an abrupt change in the state of the system, in which several bases were initially on an equal footing. This means that a slight fluctuation in the environment will determine a preferred basis (and therefore a preferred observable). This is an example of symmetry breaking, since all bases are no longer equivalent (we recall that the superposition principle required the relativity of bases). Obviously, the preferred orthogonal basis will be the one that best fits with the experimental conditions, that is, with the apparatus we have chosen to employ (i.e. with the premeasurement we have performed). This explains why entanglement is both the problem and its solution, since this particular situation ensures that a certain specific basis (and therefore a specific observable), jointly with the action of the environment, will be finally selected.

1.2.8 Decoherence

Finally, to obtain a determinate outcome (*either* $|u\rangle$ *or* $|l\rangle$) and a faithful indication of the apparatus (*either* $|a_u\rangle$ *or* $|a_l\rangle$), we need something additional. It is again the environment that plays an important role: All quantum systems are indeed open to the environment, and this, with its fluctuations, allows for the final determination of the output state. Which particular outcome will emerge in a detection event (the *third* and final step of the measurement process) at the end of this process is completely unpredictable. Indeed, the environment, which is always present in any interaction between quantum systems, *cannot be controlled* and in general it does not enter into our explicit calculations.[38] In fact, the only interaction that we can control is the local interaction between an object system and an apparatus (paradigmatically, in a laboratory). Suppose, instead, that we wish to control the interaction with the environment in order to foresee the result of the measurement.[39] Since any quantum system is open to the environment, this represents therefore a huge complex of entanglements through which, although in different degrees, any system of the universe is directly or indirectly interconnected with any other. Then, the only way to control the environment would be to know about these entanglements, and this in turn implies the execution of many local measurements that would have the effect of changing the interrelations between all these systems and therefore also the interrelation between the measured system and many others.[40] That is, the outcome that will be produced depends ultimately on the way our system is entangled with the rest of the universe, but to know this would in turn imply a change in the state of the universe and also in the outcome of our measurement.

I wish to point out that the centrality of the environment is not restricted to quantum-mechanical systems. On the contrary, the distinction between object system, detection device, and environment is quite general, and to have overseen this point is again a consequence of a simplification introduced by classical mechanics. This means that when measuring we consider the interaction between two quantum systems from a particular point of view, from the *particular perspective* under which we make use of a *local interaction* between an apparatus and an object system, and we obtain a local event. Under this particular point of view, the result that we obtain is an irreversible classical result. However, when considering the *global system* represented by object system + apparatus + environment, all of that follows the ordinary quantum-mechanical, reversible laws, and here non-local features are still present.

The crucial lesson is that the global evolution of the system–apparatus–environment does not determine measurement results in any ascertainable way. Why? Because quantum-mechanical

[38] We will see in the second part of this book that biological systems tend to control the environment, although in a limited way.

[39] [AULETTA 2006b].

[40] As already understood, on a pure classical ground, in [*BOREL* 1920, p. 294].

laws only rule probability amplitudes, and therefore probabilities to obtain a given result in a given environmental context. This shows that local and individual events are a level of physical reality that is different from the quantum-mechanical sources of our acquired information (the systems prepared in a certain state) and correlations (the premeasurement). The consequence is the randomness of quantum events. But this also means that, although any *single* event is random, the experimental conditions in which a measurement happens or a measurement-like dynamical process occurs (which also incorporates quantum-mechanical laws) determine certain regularities (expressed by the probabilities) when *several* systems are considered. For instance, a piece of radioactive matter will decay in a certain amount of time and we can also calculate the *probability* that any atom will decay in a certain time interval, but we cannot tell whether the *next atom* to decay is this or that one.

Another important consequence is that quantum features are never annihilated. From a local point of view (that of the apparatus), due to the complementarity principle [Subsec. 1.2.6], they tend to zero in a very short time, but a form of coherence (correlation) between the different components of the entangled state is still present.

In conclusion, this solution, known as *decoherence*, combines two important aspects:

- It ensures that there is no violation of the quantum-mechanical *laws* regarding the *whole* system comprehending object system + apparatus + environment.
- It points out that the measurement result is somehow an objective but *random* and *local* discontinuous event.

Let me stress that decoherence not only provides a solution to the measurement problem but, quite generally, also describes the way in which open systems spontaneously interact, as we shall see in more detail later.

1.2.9 Summing up

Collecting our results so far, anytime we measure we have[41]

- A selection of the observable to be measured;
- A reduction of the initial superposition to one of its components;
- An unpredictability of the outcome, where only the latter will guarantee that an event happened.

All this means that, when measuring, we have a many-to-one function, that is, from a superposition to one of its components. There are also cases where we have a one-to-many function. For instance, an input photon entering into an interferometer in a state that is "parallel" to the transmitted component (say in the state $|l\rangle$) will be transformed by a beam splitter in the superposition $c_l\,|l\rangle + c_u\,|u\rangle$.

This affects the nature of the chronological order and of causality [see end of Subsec. 1.2.5]. Since we can speak of an event only once we have registered it in a measurement-like interaction, before this detection event we cannot strictly speak of a chronological order [Fig. 1.12]. That is, it is only the detection event itself that eventually allows a univocal temporal structure to be imposed, such that, by means of an inference, it becomes possible to reconstruct what has happened in the

[41] [AULETTA 2005a].

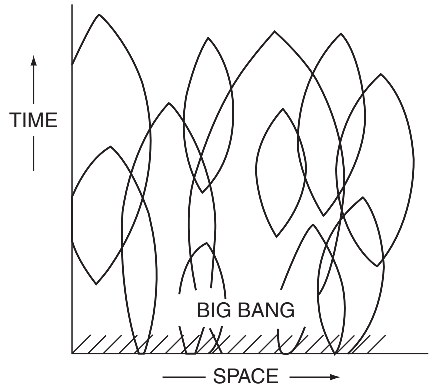

Fig. 1.12 Graphic representation of how all that "has happened" in the past is influenced by choices made in the present as to what can be observed. The upper tip of each "leaf" stands for the act of registration. The lower end of each "leaf" stands for the beginning of the quantum phenomenon being investigated. As is shown, there is not a single path leading to a single event. The leaf structure reproduces the fact that, as explained in the text, in quantum mechanics there can be both a function many-to-one and one-to-many. Adapted from [WHEELER 1983].

past, provided that there are past events.[42] Obviously, since there is no univocal time order, but either a function one-to-many or many-to-one, it is also impossible, in the general case, to speak of a univocal causal chain for quantum systems.

1.3 General Aspects of Quantum Mechanics

1.3.1 Primacy of Dynamics

The dynamical interaction between the measurement apparatus and the object system is fundamental from the point of view of the complementarity principle [Subsec. 1.2.6]. Due to the evasive nature of quantum entities, it is impossible to speak of a quantum system's state without having already somehow interacted with it. Let me summarize the three steps of measurement [Subsecs. 1.2.7–1.2.8]:

- Before the measurement act itself, as we have seen, a quantum system is subject to a *preparation*. In other words, the preparation is a *determination of the state* of the system and we may say that a state is a *class* of preparations.
- When we perform a *coupling* between the object system and an apparatus in the second step of measurement, the so-called premeasurement, we *choose* a specific observable of the object system. Since premeasurement can be appropriate for measuring a certain observable but not other ones, an observable can be understood as defined by a *class* of premeasurements. This is a consequence of the superposition principle and the uncertainty relations [Subsecs. 1.2.2

[42]The fact that in quantum mechanics there is not a univocal temporal order was proposed in [*AULETTA* 2000, pp. 797–802] and independently proved by a theoretical-experimental team [STEFANOV *et al.* 2002].

and 1.2.4]. In doing so, we have determined the way in which the state of the system will be transformed during the measurement procedure.

- Finally, the transition from an initial (superposed) state to a *final detection event* (in the third step of measurement) is largely affected by the open *dynamic interaction* between the apparatus and the object system. We are finally able to assign a property to the system. We can therefore understand a property as a *class* of detection events.

Although the measured observable can have different degrees of determination, according to the extent to which the off-diagonal terms are washed out, the measurement result is in itself a discrete event that either happens or does not (a detector either clicks or does not). This reality of the event is in sharp opposition to any initial superposition, which is characterized by features. The problem is then to properly understand the relationships between the outcome event and features. To this purpose, let us discuss an ideal experiment proposed by J. A. Wheeler,[43] the so-called *delayed-choice experiment*.

1.3.2 Considerations on the Delayed-Choice Experiment

In Wheeler's version of the Mach–Zehnder interferometer, the final detectors may be switched from the ordinary positions DA′ and DB′ to positions DA and DB before BS2 [Fig. 1.13]. This may be done after the beam has already passed BS1. In the arrangement DA–DB we detect the (corpuscular) path of the photon, and this represents an event. Instead, in the arrangement DA′–DB′, we can detect the (wave-like) interference (the features), and this cannot consist in an individual event; in fact, in order to obtain an interference profile, many experimental runs (and detections) are necessary. In general, it is impossible to measure the wave-like features of a

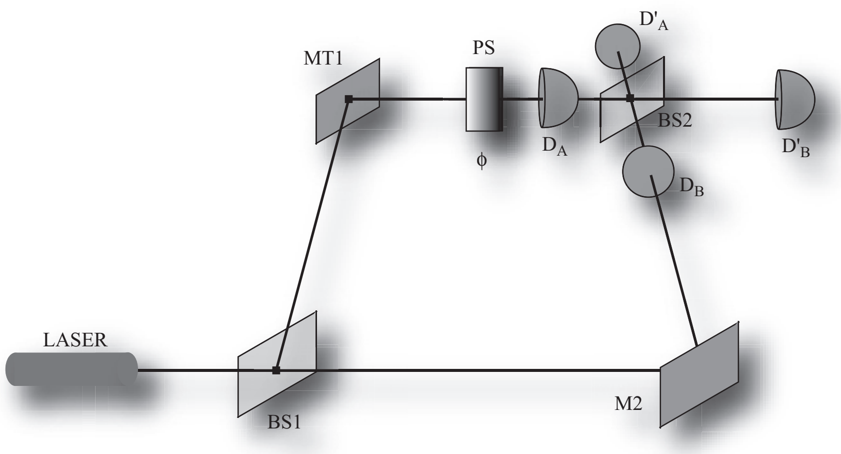

Fig. 1.13 Interferometry experiment for testing delayed choice. The setup is essentially a Mach–Zehnder interferometer in which the two detectors may be switched from positions D_A, D_B (which-path detectors) to positions D'_A, D'_B (interference detectors).

single system.[44] Obviously, the two typologies of detection are incompatible, according to Bohr's prediction [Subsec. 1.2.6]. Now, since we are completely free to displace the detectors until the last attosecond (10^{-18} sec, the time in which the light covers a third of a millionth part of a millimeter) before the photon passes the ideal line represented by DA′ or DB′, a strange situation occurs. The system seems to follow a dynamical evolution, but before passing that line *no event happens*, since, for the nature of any event, we had otherwise obtained some determined reality. This shows that the *only* events that may occur in this experimental configuration are the final detection events (at DA–DB or DA′–DB′). In other words, in quantum mechanics there are time intervals—in our case, the time interval in which the photon travels from BS1 to the detectors—where we cannot assume that "something" happened. This is a strange conclusion, because in our world it is difficult to imagine that nothing happened during the *dynamic* evolution of a system. Then, the question is: Before the final detection events (at DA–DB or DA′–DB′), is it right to assume that there is nothing in the interferometer? If so, we are postulating that the only reality in quantum mechanics is that of detection events and that the initial (superposed) state is only a fiction, a mathematical tool for calculating detection probabilities but not a form of reality.[45]

However, such an idealistic conclusion seems unwarranted. In fact, we have an input photon (in the interferometer) and an output photon (at one of the detectors). Is it reasonable to assume that we have nothing inbetween? Can we assume that a photon disappears and then a second photon comes out from nothing? I think that if one is willing to accept this conclusion, one should also be willing to accept anything else: No science would be possible. Then, *these data* strongly suggest that *there must be some form of reality* before, or independently of, the detection event.[46] What is the minimal—not event-like—reality that we are obliged to assume in any case? This minimal reality consists in the non-local correlations between the components represented by the two arms of the interferometer, i.e. in the *quantum features*. Note that these components, as far as there is no detection, have no reality in themselves (independently), and therefore we are not allowed to assume any reality other than the pure features. One should carefully distinguish here between two different issues:

- The principle according to which nothing can come out of nothing.
- However, this does not imply that the reality from which something comes out of must represent necessary and sufficient conditions in order for this something to come out. In many cases in ordinary life, we understand very well that there can be sufficient but not necessary conditions for a given event. For instance, we all understand that rain can represent a sufficient condition for traffic but that traffic can also be caused by some accident (then, the rain is not a necessary condition of traffic). Now, is it possible that there are only necessary but not sufficient conditions of a given physical event? Any time, the event could not occur without these conditions, but *additional* conditions are also necessary for the event to occur. For instance, a good diet is very important for health, however it is not sufficient if the heart, say, does not pump very well.

Consequently, I assume that, given a certain premeasurement, the features of a quantum system represent a necessary condition for a measurement outcome but not a sufficient one. In fact,

[43] [WHEELER 1978]. [44] [D'ARIANO/YUEN 1996].

[45] A conclusion that is supported by many leading physicists and philosophers [WIGNER 1961, WIGNER 1963][*VAN FRAASSEN* 1991] [ZEILINGER 2004].

[46] [AULETTA/TAROZZI 2004a, AULETTA/TAROZZI 2004b] [AULETTA/TORCAL 2011].

they only represent an insufficient determination for a measurement outcome: Any measurement outcome is already "comprehended," at the level of probability or potential, in a given initial entanglement between object system and apparatus, but this entanglement does not determine *which* of these events will be realized. As such, further conditions are needed, which are provided by the environment. However, these conditions are uncontrollable,[47] as mentioned in Subsec. 1.2.8, and this explains why the measurement outcome is unpredictable.

1.3.3 Some General Theoretical Lessons

Quantum mechanics suggests then the following conclusions[48]:

- There is a primary form of reality: The *dynamical interaction*, which determines the conditions for an event to occur (the essence of the complementarity principle). Besides this dynamic reality, we have two other fundamental aspects.
- The *features* represent a reality that has no analogue in classical mechanics and which only represent a necessary condition for any event (the essence of the superposition principle).
- The *event* is a discrete, unpredictable outcome that eventually represents the annihilation of the initial superposed state (a specific instance of the quantization principle).

A measurement is not the only type of dynamical interaction through which determination may occur. In fact, any time that two systems that are open to the environment interact we can have an analogue of the measurement process. These interactions should happen very often,[49] and, as we shall see, they constitute the necessary condition for the existence of our macroscopic world. This means that the supremacy of dynamics does not necessarily imply a form of idealism, in the sense that quantum mechanics would rely on the role of a human observer who makes use of an apparatus.[50] Neither does it imply the necessity of a macroscopic apparatus as such.[51] What the centrality of interaction dynamics strictly implies is that in quantum mechanics there are no intrinsic properties[52]; rather, any property that may be ascribed to a quantum system must be assigned given the *context* of a given interaction with at least another open system. In this way, quantum mechanics becomes a generalized *theory of open systems.*

The problem is that, in the major part of the literature about quantum mechanics, one does not distinguish between two concepts, *non-intrinsicity* and *subjectivity.*[53] The future of science will depend on the correct evaluation of this difference.[54] In fact, most of the troubles in quantum mechanics derive from a conceptual confusion between these two completely different aspects. Let me therefore clarify that

- A property is *intrinsic* to a given system if it can be assigned without any reference to other systems or higher levels of complexity. The properties of classical-mechanical objects pertain to this kind.

[47]This was already understood by Heisenberg [*HEISENBERG* 1958].
[48]I have already advanced these conclusions [AULETTA 2003a], though still in an unsatisfactory manner.
[49][JOOS/ZEH 1985]. [50]As sometimes Bohr seems to have thought.
[51]Again Bohr seems to think that a classical apparatus is necessary [BOHR 1929]. See also [WIGNER 1963].
[52][*SMOLIN* 1997, p. 51].
[53]It is a very common error that has led to the subjectivist interpretations of quantum mechanics and to the understanding of information in subjective terms. About the latter point consider the book [*VON BAEYER* 2003]. A partially different position can be found in [*VON WEIZSÄCKER* 1971, pp. 347–9].
[54][AULETTA 2006b].

- On the contrary, a property is *subjective* if it can only exist when somebody has it "in the mind." This does not mean that it exists only in the mind, but that without the concourse of a mind thinking about it, this property would not exist at all. The so-called secondary qualities (taste, smell, etc.), introduced in classical physics and philosophy by Galilei and Locke as reducible to physical primary properties, like motion,[55] seem to be of this type.

Now, the properties of quantum systems are in general *neither* intrinsic, *nor* subjective. In fact, for a quantum system to have the actual property of being located somewhere, it suffices that an appropriate detector, or detector-like system, clicks that is somehow connected (through entanglement) with the input system. For this purpose, it is not necessary at all that a human or a mind read the result. Otherwise, after the big bang the world could not have become classical or quasi-classical until at least a mind was there, which seems a little bit absurd. Therefore, what we can say about a quantum system that has been localized through a detection event, is that the actualization of this property ("To be localized somewhere") requires at least the existence of a detector or of an analogue of it (and of its environment), as well as of appropriate correlations between detector and system, i.e. it depends on the existence (and the correct arrangement) of at least another (open) system, and precisely for this reason is *not intrinsic.*[56] This other system (playing the role of a detector) does not need to be a human artifact, but can be an atom (an atom can indeed absorb photons as photodetectors do) or any other system that could, in a given environmental context, produce the result above. In these conditions, a *spontaneous* interaction occurs that is analogous to the procedure we employ in a laboratory.

Things stand similarly for information. As I will show, once it is accepted that the information's properties are not intrinsic, the only way to define them is to consider an environment relative to which they can be described. My guess is that also the properties of macroscopic objects are not intrinsic, notwithstanding the predictions of classical mechanics.

For these reasons, the most basic of all quantum principles is the complementarity principle,[57] since it is also valid for open systems, and, taking into account the previous examination, it turns out to state that the basic form of complementarity is between a global behavior and local events, being the dynamics the joint between these two extreme forms of reality. For this reason, let us reformulate it as follows

Local events and global features are complementary.

1.4 Concluding Remarks

In this chapter, we have seen the three main principles of quantum mechanics:

- The quantization principle, allowing for abrupt localizations when several quantum systems interact.
- The superposition principle that allows us to understand quantum systems as nonlocalized systems.
- The complementarity principle, according to which local events and global features are complementary.

[55][*LOCKE* 1689]. [56][AULETTA 2005b]. [57][AULETTA/TAROZZI 2006].

978–0–19–960848–5 01-Auletta-01-drv Auletta (Typeset by SPi) 32 of 869 June 13, 2011 14:45
OUP UNCORRECTED PROOF – REVISES, 13/6/2011, SPi

Moreover, we have studied the way in which measurement can be understood: As a dynamic process in which systems open to the environment can give rise to certain events. It consists of the three steps: Preparation, premeasurement, and detection. The centrality of environment is one of the most relevant results of this chapter and will show its importance when dealing with biological systems. We have seen that dynamics is here the primary form of reality and that delayed-choice experiments show that quantum processes are not necessarily causal.

In the next chapter, we shall consider how relevant these results are for the matter of information and therefore for the understanding of the whole book.

2

Quantum and Classical Information and Entropy

After having examined some misunderstandings about information, in the present chapter I shall explain why and how quantum systems can be understood as incorporating and dealing with information. The next step is to consider what the relations are between classical and quantum information as well as between information and entropy. Finally, I shall summarize the main lessons of this whole examination for the following analysis.

2.1 Misconceptions About Information

When dealing with information, there are several misunderstandings to be aware of that obscure the whole matter. Since information plays an important role in this book, some of these misunderstandings shall now be examined more closely:

(1) The first is that any dealing-with-information begins with a selection of a given message from the start. This implies that often only selected information is considered information. Selection and variety are confused here. In fact, the informational source only needs to provide the elements that *might* be selected.
(2) Information is contextual [Subsec. 1.3.3] and often its contextuality is mixed up with some form of subjectivity.
(3) Information-acquiring is sometimes meant as building some new information.
(4) Information is understood as entropy or as negentropy (entropy with a minus sign). However, entropy is related to disordered processes (thermodynamically it is to the quality of available energy) and with energy fluxes. Information, instead, is in itself a pure formal quantity. The correct relation between these two aspects can be crucial for understanding the matters dealt with in this book, particularly with the emergence of life and mind.[1]
(5) As a consequence of the previous misunderstanding, one thinks that, when exchanging information, something will be traded. As Jackendoff eloquently put it, information is often misunderstood as a liquid that is poured from one container to another.[2]
(6) This does not mean that information has no physical effect. Formal constraints in general play a very important role in the whole of nature, and as such it is time to awake from the dogmatic dream of metaphysical and ontological reductionism.[3] The standpoint I shall support

[1] Even if in an old-fashioned language, C. Lloyd Morgan seemed to understand this point very well [*LLOYD MORGAN* 1891, pp. 467–8].
[2] [*JACKENDOFF* 1992, p. 2].

in the following is that information is not only a formal structure *but also* a pure physical quantity.[4] We should not, therefore, mix the concepts of the nonmaterial (not composed of material particles) and nonphysical. Features have already provided an example of physical but nonmaterial entities.

(7) As we shall see, there is nonetheless a connection between entropy and information when information is exchanged (acquired). During those processes, several local increases of order or disorder are possible. It is a general principle that it is impossible to create order at a global level. However, as we shall see, the opposite is also true: Order cannot be destroyed at a global level. Any erasure of local order is in fact only a displacement of some ordered relations.

2.2 Quantum Systems as Information

2.2.1 The Problem

Quantum-mechanical systems provide both the informational pool and the basic interconnections of our universe:

- Quantum-mechanical systems can be considered as the *sources* of *any* information in our world. As we shall see, any quantum system, even as elementary as a two-level system, can be considered as coding an infinite amount of potential information. However, as explained, we cannot extract this information as a whole [Subsec. 1.2.5].
- As we have seen [Subsec. 1.2.8], all quantum systems are open to the environment. This is largely an acknowledged fact today and may be a consequence of the common source, as far as we know, of all physical systems, which have been generated from an original big bang. It is true that photons will be continuously created and annihilated, but they are created as open systems, already connected somehow with the other systems of our universe. This conclusion is reinforced by considering that entanglement is not an absolute concept but has degrees.[5]

Let us consider the first point. We must sharply distinguish between the concept of *information source* and that of *determining* or efficient *cause*. A source of information is only something that, provided that there is a channel and under a suitable operation (which, as I will show, is in general a selection), will deliver information (which can eventually be used and then have causal effects). This does *not* mean, however, that it will provide this information *by itself*: Additional requisites are necessary. For this reason, a source of information is not by itself a sufficient cause of information reception. Indeed, we have already seen the troubles in applying a traditional concept of causation to quantum systems [Subsec. 1.3.2]. This is true for information exchange in general. It is true that we have the feeling that a star will provide us with a certain amount of information without those constraints. We actually forget that this information is delivered to us thanks to many additional physical factors and parameters that make it possible that the light reaches us.

I follow some particular physicists[6] in assuming that all systems of elementary quantum mechanics can be considered in terms of information. Before I explain this matter, I wish to point out that there are two major problems in understanding quantum systems in this way.

[3] [ANDERSON 1972]. [4] [LANDAUER 1996a]. [5] [VEDRAL *et al.* 1997a] [VEDRAL *et al.* 1997b].
[6] [WHEELER 1990].

(1) The problem of energy: It seems impossible to have information without energy. Landauer[7] showed, however, that *throwing bits away* (selecting them), not processing them, requires an expenditure of energy. Later, Bennett[8] explained how a computer could be designed that would circumvent Landauer's principle by not discarding information and thus virtually dissipating no energy. Bennett showed that each step in the computation can be carried out in a way that allows not only the output to be deduced from the input but also the input to be deduced from the output—in other words, the machine can also run backwards. Such a machine, after having processed information in the ordinary way, could put itself into reverse mode until each step is undone. No information is erased here and accordingly no energy is lost. This is precisely the way in which quantum systems work when ideally isolated. Instead, during a measurement of a quantum system, information is selected by downloading a part of its initial amount into the environment. Because and only because of this selection, there is energy expenditure. This also means that the local entropy of the system and the apparatus must grow. This is, in fact, what quantum-mechanical calculations show. This energy expenditure together with the increase in entropy and the inaccessibility of the information that has been downloaded makes the measurement process for quantum systems irreversible.

Generally speaking, information is not dependent on the flow of energy.[9] There are even cases where the "flow of information" is opposite to the energy flow, or where information can be transmitted without energy.[10] An example of the first case is a telegraph cable, where a direct current is flowing in one direction, but a message can be sent in both directions by interrupting the current at one point and recording the interruption at another. An example of the second case is a photovoltaic cell, where the interruption of energy flows informs the cell that there is something coming in or going out. Obviously, actual information *acquiring* requires energy, as explained, since it involves selection. Therefore, I am not claiming that quantum systems can be reduced to information. What was previously said shows that other physical parameters as well, especially energy and entropy, are relevant when they interact. My point is simply that information is a basic formal quantity that cannot be reduced to other quantities.

(2) It is also difficult to understand what an interaction between quantum systems may mean if they are considered *themselves* in terms of information. One could be motivated to say that it means exchange of information, but what does exchange of information mean for two systems that are already considered in terms of information? In my opinion, the answer is: When at least two basic quantum-mechanical systems interact, the information they carry or represent will be *modified*. The problem then becomes: How can quantum information be modified? In order to answer this question it is necessary first to briefly recall how the information that a quantum system carries can be understood.

2.2.2 Potential Information

Let us consider a very useful representation of quantum states, the Poincaré sphere of quantum states. Any two-level system may be written as

$$|\psi\rangle = \cos\frac{\theta}{2}|1\rangle + e^{\imath\phi}\sin\frac{\theta}{2}|0\rangle, \tag{2.1}$$

[7][LANDAUER 1961, LANDAUER 1996a]. For a review see [LANDAUER 1991].
[8][BENNETT 1973, BENNETT 1982] [BENNETT/LANDAUER 1985].
[9][*ROEDERER* 2005, pp. 115–16].
[10][VON BERTALANFFY 1955].

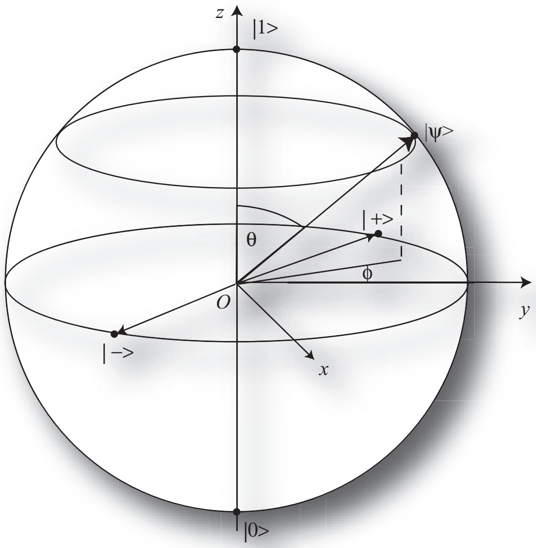

Fig. 2.1 The state vector of any two-level system, for instance $|\psi\rangle$, can be represented as a point on the surface of a sphere of unitary radius. The parameters ϕ (here represented as an angle between the y axis and the projection of $|\psi\rangle$ on the equatorial plane) and θ (represented as an angle between $|\psi\rangle$ and the z axis) are sufficient to individuate its location. The states $|+\rangle$ and $|-\rangle$ lie on the equatorial plane and represent two symmetric superpositions of $|0\rangle$ and $|1\rangle$, located at the south and north poles, respectively. In other words, orthogonal states that were previously represented as orthogonal vectors in the vectorial representation [like in Fig. 1.5], are here represented as the two opposite points of the intersection between a diameter and the surface of the sphere.

which represents a superposition of the states $|0\rangle$ and $|1\rangle$, shown at the south and north poles of the sphere respectively [Fig. 2.1], and is actually a generalization of Eq. (1.6) by considering explicitly the difference of phase $e^{\imath\phi}$ between the component $|0\rangle$ and the component $|1\rangle$ of the superposition state $|\psi\rangle$, where I recall that a difference of phase is the distance among a peak in one component of the superposition and the corresponding peak in the other component [Fig. 1.1 and Fig. 1.4]. The exponential function e is the inverse of the (natural) logarithm function. Its value, for e^1 is $2.718281\ldots$ [Fig. 2.2]. The parameter ϕ covers the parallels of the sphere and represents the relative phase between the components $|0\rangle$ and $|1\rangle$ of the superposition, while the parameter θ covers the meridians and represents the relative contribution of each of the two components to the superposition [see also Fig. 1.5]. This state represents a new informational entity: The quantum bit or, in short, *qubit*.

The amount of information contained in a qubit is infinite because the parameters ϕ and θ allow the state $|\psi\rangle$ to be located *anywhere* on the surface of the sphere (which obviously has an infinite amount of points). However, it is also *potential*, because such an amount of information is not only not accessed at a given moment but is also inaccessible to any observer in *any time window* [Subsec. 1.2.5]. Indeed, any time we try to measure the system we are obliged to *choose*

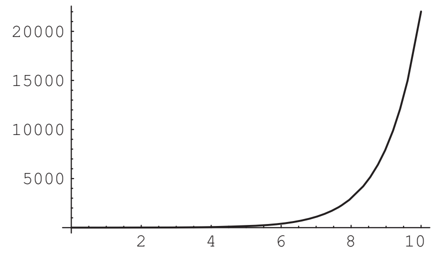

Fig. 2.2 Plot of the exponential function e^n, where the values of n are shown on the x axis.

a given observable [Subsecs. 1.2.5 and 1.2.7]. Suppose that we choose to measure an observable whose eigenstates are exactly $|0\rangle$ and $|1\rangle$. Then, we will obtain either $|0\rangle$ or $|1\rangle$ as a result. However, both $|0\rangle$ and $|1\rangle$ represent one of the two poles on the sphere's surface and therefore, when we obtain one or the other as the outcome of a measurement, we have acquired *only a bit*, a much smaller amount of information than the initial qubit. The reason is that we download into the environment the interference terms (the features) that contribute to this initial amount of potential information by spreading it on the whole surface [Subsec. 1.2.8]. This information is made active (acquiring in this way some classical bits of information) any time the appropriate conditions in the dynamic interaction between two open systems are satisfied (anytime we have a measurement or a measurement-like interaction process).

In classical physics, instead, it is assumed that it is possible to extract all the information contained in the system's state. This is however not supported by facts. Indeed, let us consider the case in which we desire to measure exactly the circumference of a ring (e.g. an ordinary wedding ring). Let us avoid all the complications that derive from the fact that, when the measurement becomes very precise so as to arrive at the molecular or even atomic scale, where matter's discontinuities and even instabilities (indeed, any piece of matter continuously exchanges photons with the environment) prevent us from speaking in any meaningful way about a circumference measurement. Let us, instead, consider a pure ideal case in which matter would be totally uniform (continuous) and static. In this case, we would very soon run into the difficulty that the circumference as well as probably the radius of the ring would be a real number. Now, we cannot (even in principle) exhaust the infinite series of decimals that constitute a real and not rational number, which means that we cannot measure it with infinite precision.[11] In other words, we cannot acquire the whole of the information contained here due to the finite resolution of any measurement (the impossibility to reduce to zero the measurement error). We may reformulate all that in an information *accessibility* principle:

The whole information potentially contained in a system may only be partially accessed.

From a quantum-mechanical perspective, the reason is that we have interference terms (features) that cannot be acquired as (classical) information. Classically, this is due to a finite resolution of measurement. However, in both cases the ultimate source of that impossibility seems to me to be the discreteness of any codification and therefore of *any* information acquiring [Fig. 2.3].

Let us consider the quantum-mechanical case. Assuming that $|0\rangle$ and $|1\rangle$ represent information (they are indeed measurement outcomes), I stress that these states are not essentially different

[11][*MARGENAU* 1950, p. 39].

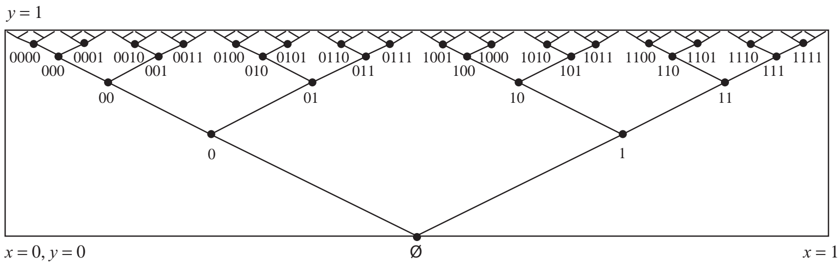

Fig. 2.3 The continuous line $0 \leq x \leq 1$ can be represented as a limiting case when quantized codification tends towards infinite series represented by real numbers.

from any linear combination (2.1) of them. In fact, if one chooses to observe another (non-commuting) observable, whose eigenbasis is represented by $|+\rangle$ and $|-\rangle$ [Fig. 2.1], where $|-\rangle$ is a state orthogonal to $|+\rangle$, the possible outcomes will be $|+\rangle$ and $|-\rangle$. Therefore, they will also represent information. However, they consist of a linear combination of $|0\rangle$ and $|1\rangle$, namely the superpositions

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{2.2}$$

where, for the sake of simplicity, it is assumed that the states $|+\rangle$ and $|-\rangle$ represent symmetric superpositions of $|0\rangle$ and $|1\rangle$, and are located on the equatorial line (equidistant from the poles). What is inaccessible, therefore, is not a linear combination of $|0\rangle$ and $|1\rangle$ *as such*, but the whole amount of information contained in *any* quantum state (independently from any measurement procedure), which also comprehends, beyond the possible eigenstates, any "interference" of them. Therefore, I also assume that information expresses the relation between possible events[12] and that consequently it is not only what is actually communicated and received that can be called information. The reason why many think the opposite is that information-acquiring is meant as constituting information [Sec. 2.1].

We should not forget, however, that there is also an important difference between a prepared state and a measurement outcome. The states $|0\rangle$ and $|1\rangle$ themselves (as well as $|+\rangle, |-\rangle$), taken as *possible* but not actual measurement results, represent *potential* information, i.e. information that has not yet been acquired[13] (and possibly never will be). Let me give further evidence: When any two systems are entangled [Subsec. 1.2.7], they constitute a further informational entity, called an *ebit* (entangled bit). The interdependencies displayed by the two subsystems of an entangled system are not mediated by any detectable physical signal, and therefore there is no causal connection either[14] [Subsec. 1.2.9]; they are nonlocal in character and an immediate consequence of quantum features. The reason is again to be found in quantum information: Since information expresses the relation between possible events, it is also independent of space and time and entanglement is a natural consequence. Therefore, an ebit is a quantum channel, a

[12] Personal communication, A. Zeilinger.

[13] [*VON WEIZSÄCKER* 1972]. See also [*KÜPPERS* 1990, pp. 36–8]. However, these authors speak of a potential information in the sense of an information that can later be received. I agree for classical systems. Here, I speak of potential information in a more radical sense, because the whole of quantum information cannot be received or be accessed. Nevertheless, in my opinion the concept of potential information can cover both the classical and the quantum case, meaning information that has (still) not been acquired.

[14] [*MAUDLIN* 1994].

typical quantum-mechanical information-sharing between systems. Now, an ebit allows things to be done that cannot be performed classically, like encrypting a text or transmitting information in new ways. Thus, entanglement can be interpreted as a potential *resource* to transfer additional quantum information in the future at no further cost.[15] This justifies the notion of potential information.

One may obviously worry about the concept of potentiality as such. However, this worry is again not particularly supported by facts. It is very common in physics to speak (in the presence of some field) of potential energy as a potentiality to do work in appropriate conditions (for instance, a stone that can roll down a hill due to the effect of the gravitational field). It is true that one could object that this is only due to the effect of the external field and therefore has nothing to do with the character of the object as such. This objection does not seem very sound to me however [Subsec. 1.3.3]; moreover, we have wonderful examples of intrinsic potentiality, if this word can pass here. For instance, in chemistry the chemical potential is proportional to the Gibbs free energy and expresses the potential of a chemical substance for undergoing a change in a system.[16] This is interesting because the Gibbs free energy expresses potentiality to do work in pure chemical terms (that is, without considering the work performed when a system expands in the presence of an opposite external force). I recall that, at equilibrium, the chemical potential of a substance is the same throughout a sample, regardless of how many phases are present.

Therefore, the concept of potential information makes perfect sense. We can say that potential information is any string of codified bits (or qubits) considered independently from information-acquiring. Thus, we can state that information codification has several requisites:

- A basis representing (a finite set of) mutually exclusive states that:
- Can be linearly combined (in a superposition state). Note that not all combinatorial rules are linear [Subsec. 1.2.2]. Linearity is, however, necessary for information codification.
- By varying the coefficients (and/or the parameters ϕ and θ), we can obtain an infinite number of possible combinations, i.e. it is in principle possible to express anything by means of these two elements and their combinations. Therefore, the coefficients of any superposition represent the syntactic rules according to which we can combine the elements $|0\rangle$ and $|1\rangle$. It also follows that these elements can be understood as a (binary) code, as, for instance, in quantum computation.[17]
- There are specific rules that allow the translation of a state written in a given basis to its formulation in another basis (that is, there are different codes): This is a necessary requirement for having information. For instance, the basis $|0\rangle, |1\rangle$ is connected to another basis $|+\rangle, |-\rangle$ by the translation rules (2.2).

Therefore, *actual* information is only that information which has actually been received. However, in order to receive something as information, it should already have at least the formal characteristics of information quoted above, i.e., at least mutual exclusivity of the possible unities of the message, infinite potency of expression, and linearity of the combination of the unities. Let us call *codified information* any combination of alternative physical states representing the units of a code according to syntactic structures. Note that codified information is always potential, both classically and quantum-mechanically. For instance, the information contained in a book is not actual until it is read. Indeed, if the texts written in a certain language are no longer read, one can even lose the memory of the meaning of the characters in which that language is written (as happened for ancient Egyptian before the discovery of the Rosetta stone). It is

[15] [HORODECKI *et al.* 2005]. [16] [*ATKINS/DE PAULA* 2006, pp. 122–3].
[17] [DEUTSCH 1985, BENNETT 1995].

not by chance that, in the case of hereditary transmission, one speaks of activating the inherited information by appropriate mechanisms.[18] A good measure of the classical informational content of codified information is represented by the Kolmogorov measure of complexity, namely the measure of the computational resources needed to specify the string of characters instantiating codified information.[19] This measure has also been generalized to quantum mechanics.[20] Therefore, both classically and quantum-mechanically we may assume an information-activation principle:

Any (codified) information is as such potential (or dormant), and only additional (external) conditions may activate it.

For instance, in order to obtain actual bits from a quantum system we need at the very least an environment. Moreover, quantum systems can be understood as *information processors*, since they evolve with time changing their state, that is, performing a particular transformation (mapping) of an initial codified information to some other codified information. I recall that this information-processing is reversible, provided that there is no information selection and therefore no measurement outcome [Subsec. 2.2.1]. This is also the background of the field known as quantum computing.[21]

Summing up, information is *in itself* only a formal quantity (a potentiality). It needs to be activated or acquired. As mentioned and as we shall see in the following, information acquisition is a dynamical process in which entropic fluxes are also relevant. My point, therefore, is that the difference between a qubit and a bit is relative to a measurement procedure within a local context, that is, relative to a given environment: Any bit can be understood as a qubit made active or accessed.[22] Both can be interpreted as minimal information entities. The only difference is that a qubit is a bit from a counterfactual point of view or, in other words, is only potentially a bit: We can speak of a bit only once one had chosen or will choose to perform a possible operation (or if some objective conditions will be spontaneously produced) in which this state could be obtained. This shows that is not the *form* that distinguishes $|0\rangle$ from $|+\rangle$ or $|-\rangle$, but only the *fact* that we have obtained it as actual information $|0\rangle$ and not e.g. $|+\rangle$, given (a) the choice of measuring the observable of which $|0\rangle$ is an eigenstate, and (b) the selection of $|0\rangle$ in certain environmental conditions.

Obviously, this does not mean that the transition from a qubit to a bit is so easily performed from the point of view of the history of our universe. Actually [Subsecs. 1.3.3 and 1.2.8], dynamic interactions happening spontaneously in nature do simulate, to a certain extent, the process of information acquisition in our laboratories with artificial devices (or rather vice versa). As a matter of fact, decoherence provides a necessary condition for classical bits. However, as we shall see, other conditions are also necessary, which are not so immediately given in a prebiotic environment.

2.2.3 Interpretation of Information Modification

We have seen that any quantum system is in itself a reversible information-processing system [Subsecs. 1.2.2 and 2.2.2]. However, its information content can also be changed through interaction with other systems [Subsec. 1.2.4]. My main hypothesis is that such a change of a quantum system (a qubit) can only happen through two types of information modification:

- the constitution of an entangled state (ebit), and
- information selection (bit).

[18] [*PALSSON* 2006, p. 21]. [19] [KOLMOGOROV 1963]. [20] [BERTHIAUME *et al.* 2000].
[21] [*NIELSEN/CHUANG* 2000]. [22] [AULETTA 2005b].

In the latter case, there is a *selection* of the initial amount of information. I will call this process in all its generality (not restricted to quantum mechanics) *canalization* due to the fact that the final output state can be understood as a specific component *chosen* among many that are potentially given in the initial state, so that we may say that the system has been "canalized" into this component. Taking into account both this and what has been said about the randomness of quantum events, I wish to propose here what seems to me the most basic of all nature's principles, a true selection principle:

In appropriate conditions, nature always selects one among several options.

Indeed, this principle covers both the cases of quantum mechanics (where we have quantum events and the quantization principle [Subsec. 1.3.3]) and of the classical world; it also fits perfectly with the nature and character of biological systems.[23] The reason why I consider the selection principle so general will become clear in the following. For the time being, let me give an example. We cannot precisely foresee the fracture line in a given piece of material (like glass) even if a certain shape will emerge after certain local cracks here and there. These cracks are spontaneous selections.

The production of an entanglement of two systems seems to add something new that was not there before: A quantum channel that did not exist between them before is established. However, the information that the two entangled systems come to share is the information that *each already contained*. Let us call this form of information modification in all its generality (not restricted to quantum mechanics) *channelization* because it consists of a reinforcement of a (quantum) channel. Entanglement is a form of information-sharing (technically called mutual information) in which two systems become correlated. This does not imply at all that the two systems exchange signals and therefore have some form of causal connection. Information-sharing can indeed also happen at a distance. In order to understand this point, I wish to introduce an analogy with a classical case. If two persons in two distant cities read the same newspaper, they share information even if they have never met (and never will meet). This can easily be seen, if they indeed meet: In this case they can exchange comments on some article published in the newspaper exactly because both of them already share this piece of information. In other words, the fact that both have read the same news, allows them to talk in a way that they cannot do otherwise.[24]

The reason why there are only these two forms of information modification is that *information* (representing order) *cannot be created*, or, in other words, that entropy (representing disorder) must locally increase or at least remain constant. Quantum systems already possess, in the general case, an infinite—but potential—amount of information. How could further information be added to such systems? It could be said that it is impossible to "put" information into a quantum system from the outside. Any attempt at introducing information into a quantum system is actually nothing more than the enlargement of the original system into a wider system where the environment or the apparatus must also be considered and where the information is differently redistributed among these subsystems (even if the total system, i.e., the whole universe, remains in the same state). I also believe that this impossibility is ubiquitous in nature, as shall be shown below by considering some examples.

Channelization and canalization are of enormous importance and they represent the true informational foundation of the main thesis of this book: A form of generalized Darwinism, as I shall explain shortly. Indeed, thanks to mutual information, systems can be correlated even if

[23][ELSASSER 1969] [PARISI 2006]. See also [*ULANOWICZ* 2009a, pp. 43–7].

[24]Obviously, there are many differences between classical correlations and quantum entanglement, as is nicely shown by so-called entanglement swapping.

they do not interact directly and are even blind to each other.[25] Moreover, they can develop somehow autarchically by selecting information they already contain. We shall see how relevant these issues are for epigeny and in particular for the development of the brain.

2.2.4 Measurement and Entropy

The interpretation of the initial (superposition) state in terms of potential information can be interesting if we wish to give a physical explanation of the measurement process [Subsecs. 1.2.7–1.2.9]. As a result of the measurement process, as described above, one would obtain a state in which the interference terms are negligible but not completely eliminated. The physical meaning of this is that the information contained in the interference terms that express the entanglement between system and apparatus (the features) is *lost into the environment*.[26] This means that part of the initial information becomes obscured or disturbed by some environmental fluctuation, and this part becomes definitively *inaccessible* (it would be accessible only under the hypothesis that we could reconstruct exactly the same situation of the whole universe as existed just before the time of measurement). This is therefore a further application of the information accessibility principle [Subsec. 2.2.2]. Considering the problem the other way around, the reason for the inaccessibility of the complete potential information contained in a system's initial state is that it could be obtained only by measuring the system, but, on the other hand, any detection event consists of a selection of a given subset of the initial potential information, which implies that the rest becomes irremediably lost in the environment. Then, it is precisely this loss or opacity of part of the initial information that represents the necessary condition for obtaining a measurement's result. We may conclude here that, although the problem of accessibility is obviously an epistemological one, this does not eliminate the fact that the environmental disturbance of the initial information is what permits, as a physical result, a measurement outcome.

As mentioned, while *globally* the same quantity of information and entropy is conserved (perhaps the whole universe is a zero-entropy world), locally we may have different repartitions of entropy and information. In other words, the entropy (the amount of disorder) of the object system will in general increase during measurement and other similar dynamical processes. This allows, on a general plane, the possibility that in other locations the entropy of other systems can also decrease in order to balance this increase. For this reason, as I have suggested,[27] the existence of living beings, which display a local decreasing of entropy, is allowed by quantum-mechanical laws (and this is also true for other ordered structures). Here, the term *allowed* means that, for the emerging of local order (decreasing in entropy) somewhere, there are already the *conditions* (increasing in local entropy) *elsewhere*, so that the total imbalance remains constant. This is obviously something very different relative to the classical laws of thermodynamics, which do not provide a justification for those conditions. I also wish to stress that decreasing in local entropy is a more general character of nature that is also present in abiotic self-organizing systems.

[25] A point perfectly understood by D. Hebb when speaking of associations between concepts without having occurred together in the subject's past experience [*HEBB* 1949, p. 132].

[26] It is interesting to observe that the initial potential information contained in an entangled state can be maximal, while the state represented by a state without, or at least with less, interference terms (a mixture) does not represent a maximal amount of information.

[27] [AULETTA 2005b].

2.2.5 Global and Local

Any event is view-dependent, interactional and contextual, because it is local (it can depend on a local measurement). Quantum features are, on the contrary, view-independent, law-like and global [Sec. 1.1]. Then, while events are perspective-like being local occurrences, laws may be very well invariant (since are global). However, an event is actual reality, while features, i.e. quantum interdependencies (like any other relation, structure, or formal constraint), are only potential (relative to events). As quantum mechanics is nowadays the most basic physical theory at our disposal, everything at play here with regard to the notion of features suggests that it is a very fundamental character of our world and that nature is constituted not only by local events but also by global features. This will be explored in the next part of the book. The predominant problem and source of all difficulties in this matter is that the distinction between event and relation is at the same time both relative and absolute:

- It is *relative* as far as the informational content of an event is a cluster of relations (for instance, as we have seen [Subsec. 2.2.2], an eigenstate can be considered itself a superposition from another point of view) and it is unpredictable and surprising *only* relative to a given set of relations or to a previous status quo: A measurement outcome has an informational content and therefore indicates a property only *because*:
  (1) It is a possible outcome in a given set and *not in itself*. $|0\rangle$ without $|1\rangle$ is nothing but a closed monad.
  (2) There is a correlation (coupling) between the object system and apparatus, such that the random detection event allows an attribution to the system of a certain property (the eigenvalue associated to the output eigenstate of the measured observable).
- However, this difference is also *absolute*, as far as there are no means to derive from a *given* set of relations a given event (relations are influential only at a general and potential level and not at the individual one). This is why in quantum mechanics we have true random events, although often also in the macroscopic world random events are important.

In other words, the difference is absolute from the local point of view of the actual information (locally, only events matter) and relative from the global point of view of the potential information (here only the form, i.e. the information codification matters). Therefore, from a local point of view, only an event is absolute: Once it has happened, it has happened; while its associated properties are relational and therefore not absolute. The reason is that, in quantum mechanics, a measurement outcome is simultaneously unpredictable but dependent on the environmental conditions (and for this reason also from the other components of the superposition of which it is a component). On the other hand, features are relative because they encompass possible outcomes, and in this sense are only potential. They only represent necessary conditions for possible outcomes. However, they cannot be accessed from local properties and events only (again the accessibility principle! [Subsec. 2.2.2]), and in this sense they are absolute or indifferent to those outcomes. This means that, if we measure a quantum system, we cannot guess whether it is entangled with other systems or not. In order to know that, we need to *compare* several different measurement outcomes obtained in measuring, for instance, couples of particles prepared in the same state in two different localities. In this case, we shall discover that in some cases the statistics are so extraordinary that they let us infer that there is indeed an entanglement.

What I am saying stresses not only the unavoidable randomness of quantum mechanics (and of our world in general), but also represents the reason why our macroscopic world, which is constituted by actual information, is not a fiction but a solid reality. In a certain sense, it is the only form of reality we understand by this term *actual* reality. This reality is both allowed by quantum-mechanical laws and is not deducible from them because, again, quantum mechanics does not determine actual information (an instantiation of the selection principle [Subsec. 2.2.3]). As we shall see, this is what can truly be called *emergence*. This is the great misunderstanding of the so-called Many-World interpretation of quantum mechanics[28]; in not having considered this point, it implies that macroscopic reality is a species of illusion of the mind.

Summing up:

The global cannot be accessed from the local but the global does not determine the local.

The global is perceived as a background noise from the local point of view (the huge amount of all interference terms that have been downloaded into the environment). The local could probably be imagined from a global point of view as a cluster of anomalous fluctuations.

The complementarity between local (events, atoms, elements) and global (relations, structures) is therefore such that these two aspects are (1) opposite aspects (more local at the expense of global and vice versa) but also (2) tightly interwoven in dynamical processes [Subsecs. 1.2.6 and 1.3.3]. It is a general principle of quantum mechanics (the complementarity principle) and, as we shall see, provides a solid framework for the study of such different phenomena as perception, complex systems, and living beings in general.

2.2.6 Summary

I have pointed out that

(1) Quantum systems are information-processing entities, even if they do not reduce to information when they interact.[29]
(2) It is impossible to "put" information into a quantum system from outside.
(3) Information modification of a quantum system can only happen through either (a) entanglement (or information-sharing) with another system, which in general terms is called channelization or (b) through a type of detection (selection), which in general terms is called canalization. These general characters are true for any information modification, as we shall see.

A consequence is that it is impossible to *transmit* the "content" of an entanglement, represented by the features.[30] This is the reason why features can be understood as potential information. Then, the quantum-information framework presented here sets not only constraints on any exchanging of or dealing with information in our universe but also on the kind of causal relations that are admissible. The so-called information causality principle relates to the amount of information that an observer (Bob) can gain about a data set belonging to another observer (Alice), the contents of which are completely unknown to him. Using all his local resources (which may be correlated with her resources) and allowing classical communication from Alice to Bob, the amount of information that the latter can recover is bounded by the information volume (n) of the communication. Namely, if Alice communicates n bits to Bob, the total information obtainable by Bob cannot be greater than n, otherwise entanglement itself would be transmitted or shared.[31] Consequently, the amazing result that was found by Pawłowski *et al.* is that a hypothetical theory which fulfills the requirements of causality (being a non-signaling theory), if it exceeds the so-called Tsirelson bond, which is imposed by quantum mechanics, also violates the principle of information causality. In other words, the Tsirelson bond sets a limit on the possibility of

[28][EVERETT 1957][DEWITT 1970]. [29][AULETTA 2005b].
[30][CLIFTON *et al.* 2003]. [31][PAWŁOWSKI *et al.* 2009].

acquiring information that is stronger than the simple no-signaling requirement. It is important to understand that this result is of general validity since there cannot be in nature any possibility to violate the principle of information causality. In other words, the result is not confined to quantum mechanics only and clearly justifies quantum information not only as the most general theory of information that does exist but also as the theory ruling and framing any kind of interaction in our universe.

The combination of Theses (2) and (3) above when extended to classical systems can be called, with some provisos, a generalized Darwinism on the line of Lewontin and Plotkin,[32] and represents the core of this book. It is important to stress that the generalized Darwinism as it is proposed here is not concerned with a kind of ideological transubstantiation of evolutionary theory, as should be clear in the following. Indeed, Darwinism can find this general form only by widening itself into a more comprehensive scientific theory. Selective processes are the mechanisms through which new levels of complexity and new features in the history of life can come out. However, at the same time, through these processes, organisms are led both to share information with the environment (even if no Lamarckian instructive mechanism takes place) and to become more integrated systems.

2.3 Classical and Quantum Information Acquisition

In a classical world, information is the systematic dependence of a signal on a source. Classically, the channel is the form under which the receiver becomes and interprets a signal. The channel is not necessarily physical in the ordinary sense of the word. Also classically, it consists rather in the mutual information between sender and receiver. The signal can therefore be considered as a trade-off between novelty at the source and regularity given by the dependency represented by the channel.[33]

2.3.1 The Classical Case

Recall that quantum mechanically, in order to recover the information about an object system, we need the coupling with an apparatus (channelization). Classically, we have a similar situation. Here, we have an unknown parameter k whose value we wish to know and some data d pertaining to a set D at our disposal. This is a very important point, since we NEVER have direct access to events or things (whose properties are described by k) but always to things *through* data.[34] These data can be represented by the position of the pointer of our measuring apparatus or simply by the impulse our sensory system has received, or even by the way we receive information about the position of the pointer through our sensory system. It does not matter how long this chain may be. The important point is a matter of principle:

We can receive information about objects and events only conditionally from the data at our disposal.

This is why I have explained [Subsecs. 1.3.3 and 2.2.5] that properties are always relational as opposed to intrinsic, even in the classical case, and as such quantum mechanics teaches us a general lesson.[35] Let me give a classical example: Suppose that we wished to know exactly what the distribution of matter was in the early universe. We can know this by collecting data about the background radiation we receive now. This again shows a very important common point between quantum and classical physics that is not well understood, and which has been pointed out by

[32][*PLOTKIN* 1993]. [33][*VON WEIZSÄCKER* 1972].
[34][ZUREK 2004] [*BATESON* 1972, p. xxv]. [35][AULETTA/TORCAL 2011].

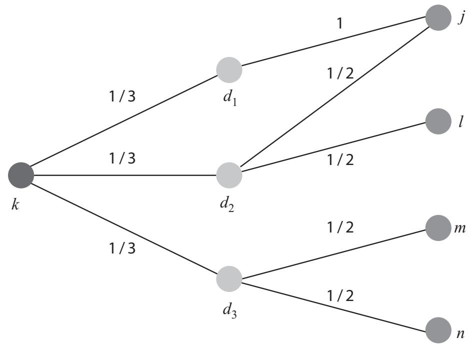

Fig. 2.4 A simple tree for calculating conditional probabilities (from left to right). Suppose that we have an event k, occurring with a probability $p(k) = 1/2$, and a particular set of effects (data at our disposal) d_1, d_2, d_3, each one occurring with 1/3 probability given k. It is easy to see then that the probability to select event j and have event k is $p(j,k) = [p(j|d_1)p(d_1|k) + p(j|d_2)p(d_2|k)]p(k) = 1/3 \cdot 1 \cdot 1/2 + 1/3 \cdot 1/2 \cdot 1/2 = 1/4$ while the probability to select event n and have event k is $p(n,k) = p(n|d_3)p(d_3|k)p(k) = 1/3 \cdot 1/2 \cdot 1/2 = 1/12$. Note that when $p(k) = 1$, we have $p(j,k) = p(j)$ and $p(d_i|k) = p(d_i)$ for any of the data. Note also that the sum of the probabilities leaving any node = 1.

Wheeler's delayed-choice experiment [Subsec. 1.3.2]: We cannot receive any information about past events unless they are received through *present effects* (data). This is an equivalent formulation of what I have said before, since, given the relativity theory, any event, represented by a parameter k, can be known only through a *later* effect, due to the limits of light speed. As a matter of fact, any perceptual experience we have is mediated and slightly delayed in time[36]. Moreover, since we always have experience of a part of the possible effects produced by an event, this is again an application of the principle of information accessibility [Subsec. 2.2.2]. Generally speaking, any spread of a signal from an initial source is subject to some form of dispersion and therefore also of information loss, which could be considered as a spontaneous (random) selection or sequence of spontaneous selections (according to the selection principle [Subsec. 2.2.3]), whose result is therefore a growing noise added to the initial input. To this extent, information acquisition is embedded in a larger and spontaneous behavior of nature, whose consequences will be momentous throughout this book.

Obviously, once we have observed or acquired data, we must perform an information extrapolation that allows us to guess about the value of the parameter k. This is information *selection.* The probability $p(j,k)$ that we select a response j having an event represented by an unknown parameter k (i.e. the probability that both event k and event j occur) is given by

$$p(j,k) = p(j|k)p(k), \tag{2.3}$$

where $p(j|k)$ is the conditional probability that the event j happens given the event k, and $p(k)$ is the absolute probability that the event k occurs. Now, we may expand this probability by taking into account the data d that are somehow the interface between the source event k and our final selection event j [Fig. 2.4][37]:

$$p(j|k) = \sum_{d \in D} p(j|d)p(d|k), \tag{2.4}$$

[36] [*MOUNTCASTLE* 1998, p. 3]. [37] [*HELSTROM* 1976].

where I have made use of a discrete case for the sake of simplicity and the symbol $\sum_{d\in D}$ means a summation over all the data d pertaining to the set D. By inserting the last equation in the previous one we obtain:

$$p(j,k) = \sum_{d\in D} p(j|d)p(d|k)p(k) = \sum_{d\in D} p(j|d)p(d,k), \tag{2.5}$$

where I have made use again of the fact that $p(d,k) = p(d|k)p(k)$. Eq. (2.5) can be considered a generalization of the well-known general formula

$$p(j) = \sum_{d\in D} p(j|d)p(d), \tag{2.6}$$

and it reduces to the latter when $p(k) = 1$, i.e. when the event k occurs with certainty (when k cannot occur or never occurs, $p(k) = 0$). It is important to stress that the two conditional probabilities $p(j|d)$ and $p(d|k)$ are quite different. This can be seen formally by the fact that in Eq. (2.4) we sum over the data d, which represent the conditioned results relative to k on the one hand and the conditions for information extrapolation on the other. This means that the probability $p(d|k)$ represents how *faithful* our data are relative to k, that is, how reliable our apparatus (or our sensory system) is (how good the channelization is). Instead, the probability $p(j|d)$ represents our ability to *select a single j* (this is canalization) able to interpret the occurred event in the best way.[38]

Having made these considerations, we immediately see that Eq. (2.4) or (2.5) represents the classical analogue of the quantum process summarized in Subsec. 1.3.1: The classical term $p(d|k)$ corresponds to the coupling between the object system and the apparatus. Obviously, the difference between the classical and the quantum case is that, when we have an entanglement, we can have a perfect correlation between apparatus and object system, which is hardly the case for classical situations. Finally, the counterpart of the probability $p(j|d)$ is, in quantum mechanics, the probability of a final detection event, given a certain experimental context (i.e. a premeasurement).

This result is very important with regards to the fact that the classical theory of information has supported the idea that information is a two-term process, namely, a process in which we have an input and a corresponding output. This is due to the circumstance that it has been formulated as a communication theory for controlled exchanges of information, so that, in general, we assume a determined output given a certain input. When, on the contrary, we deal with a general theory of information acquisition, we are always dealing with at least three terms: An unknown (not controlled) parameter, some data, and a certain information selection.

2.3.2 The Mechanism of Information Acquisition

As a consequence of the previous subsection, in the most general case, any information acquisition can be thought of as a three-system and three-step process. To shed further light on this point, let us come back to the model of measurement, which is a particular instance of information acquisition (it is also a specific instance of dynamic interactions between open systems). The whole measurement process is divided into [Subsecs. 1.2.7–1.2.8]: A first step in which we prepare a system, the premeasurement or coupling (in the quantum case, an entanglement), and the detection itself, which is a selection act. The detector and the rest of the apparatus can be spatially or temporally separated. Moreover, the final choice can be random or not (in the quantum mechanical

[38] [*FRIEDEN* 1998].

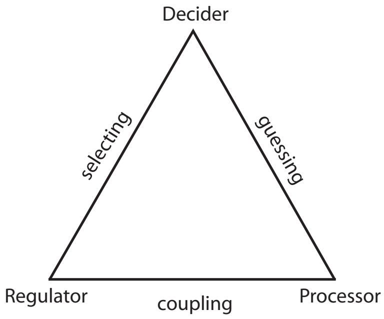

Fig. 2.5 The fundamental informational triad Processor–Regulator–Decider.

case it is random). However, there is always some sort of incertitude affecting the final selection act. Viewing the whole from the point of view of the involved systems, we have: The object (measured) system which represents, as I have stressed, codified information changing in time, and can be considered the *processor* [Subsec. 2.2.2]. Its variations (which can be either random or according to a program) provide the starting point of the whole process. The measuring device that is coupled with the object system is the *regulator*, while the final act of detection is done through a *decider*. The regulator owes its name to the fact that choosing a certain apparatus with a certain experimental set up and with a certain pointer indeed contributes to determine the conditions in which information is acquired. However, since these conditions cannot provide the variation (which is guaranteed at the source by the processor), this determination is rather a tuning of the measurement process. The only codified activity at this level is in the processor, since the regulator connects previously independent systems while the decider can provide a selection that, only thanks to the indirect connection with the processor through the regulator, will finally consist in an option within a set of alternative possibilities. As we shall see, this is also true for higher levels of complexity: Information codification is only in the initial information processing.

Let us try to establish a connection with the previous distinction between event, correlations, and process [Subsec. 1.3.1]: Information selection represents the event, the coupling instantiates the dynamic connection, and the source of variety, which *in itself* is unknown to us, in a quantum-mechanical system contains a set of unknown features, and manifests its information *for us* only through the information selection and coupling, that is, becoming in this way a variety that is accessible *to us* in this dynamical process.[39]

The whole process can then be seen as in Fig. 2.5: The relation established between the regulator and the processor is a coupling, which allows for information to be acquired. The relation between the decider and the regulator is an information selection. Finally, thanks to this two-step process, the decider can acquire information about the processor (or the event resulting from processing), performing in this way the analogue of a *guessing* (which can be a property assignment). In other words, any time that a decider (even randomly) selects and eventually stores some information—which, through an *appropriate* coupling, reveals something about another system—we have something that, at a pure physical level and without any reference or goal-directed

[39][*PEIRCE CP*, 6.97]. I am therefore sympathetic with Bateson's definition of information as "A difference that makes a difference" [BATESON 1971, p. 315].

action, bears some structural relation to what we are authorized, at another level and in another context, to call a true guess. Therefore, this guessing, or the whole measuring process, must not be understood merely in subjective terms [Subsec. 1.3.3]. It is also important to realize that this guessing could be considered part of a further preparation procedure in which a processor is determined. For instance, we may decide to measure our system again starting from its output state. The reason could be that we are not sure of our guess because of some doubts about the apparatus reliability. In this way, the whole process presents a certain circularity, as stressed by Fig. 2.5.

Resuming, the main assumption of this book is that

The above scheme is the basis of ANY dealing-with-information process of our world, and in particular of biological systems.

Also Shannon understood very well that information is concerned with a reduction of incertitude (a choice among alternative possibilities) and that, in order to have an exchange of information, we need both a variety and a channel (an interdependency).[40] However, as I have mentioned [Subsec. 2.3.1], he mainly dealt with engineering problems of communications, in which the task is to increase the match (fidelity) between an input and an output in controlled situations. In this case, the reduction of incertitude already happens at the source (by the sender who chooses a certain message among many possible ones). This is rather a limiting case in ordinary life, and the problem was that Shannon's followers took this specific treatment as a general model of information. The worry consists in the fact that, in the most general case, the reduction of incertitude is *only at the output* and not at the input. This is evident for quantum systems: Qubits are not selected messages[41] (due to non-local features) [Subsec. 2.2.2]; consider again the delayed-choice experiment [Subsec. 1.3.2]. However, this is a truth of general validity: The reason is that, in most situations (even classically), nobody has perfect control of the source, and therefore, even if a determined message has been selected, this remains unknown for the receiver (it is as if it were undetermined).[42] In such a situation, one is obliged to make a guess about the input starting from a certain selection at the output. This is also sometimes true for the sender; if they desired to be certain about the message that has been sent, they need to process it again, and in this way reduce the incertitude that may affect (their knowledge about) the input message. On the contrary, the selection operated by the receiver IS the *received message* (it is the event that has happened). Obviously, the receiver may also try to verify again whether their understanding is correct. However, this understanding concerns the guess *about* the input message and not the *act* of selection, i.e. the event itself by which a reduction of incertitude at the output has been produced for ever [Subsec. 2.2.5]. This is an *irreversible event* and therefore an ultimate fact. We may generalize this by saying[43] [Subsec. 2.2.1]:

It is the final act of information selection that introduces irreversibility in any information exchanging or processing.

Quantum mechanics allows us to understand this point quite well, and for this reason, as already stressed, it is the true generalized information theory. Here, the input information is intrinsically uncertain and the only event we have is at the output. We shall consider the noticeable consequence of this in detail in the next part of the book.

[40] [SHANNON 1948].
[41] Otherwise quantum bits could be cloned, which is impossible [WOOTTERS/ZUREK 1982].
[42] [AULETTA *et al.* 2008]. [43] [LANDAUER 1961, LANDAUER 1996a] [BENNETT 1973, BENNETT 1982].

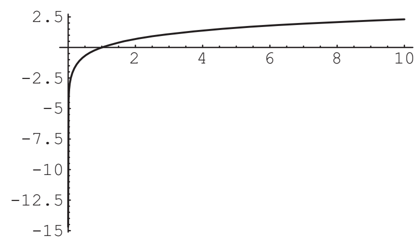

Fig. 2.6 Plot of the log function. Compare with the Plot in Fig. 2.2.

2.3.3 Quantum and Classical Entropy

The information given by the final outcome (selection) j [Subsec. 2.3.1] in a process of information acquisition (reduction of incertitude) as described in the previous subsection is given by[44]

$$I_j = -\lg p(j), \tag{2.7}$$

where

$$\lg p_j = \log_2 p_j. \tag{2.8}$$

The quantity (2.7) was called *surprisal* (or surprise) by Shannon.[45] The logarithm $\log_z x = y$ means that $z^y = x$ [Fig. 2.6]. The number z is called base and the logarithm lg with base 2 is appropriate for binary codification. The properties of the logarithm are

$$\lg(x \cdot y) = \lg x + \lg y \quad \text{and} \quad \lg(x/y) = \lg x - \lg y. \tag{2.9}$$

The classical *Shannon entropy H* is given by the sum of all the possible alternative outcomes or events (pertaining to a set J) weighted by the probabilities $p(j)$, that is,[46]

$$H(J) = -\sum_j p(j) \lg p(j), \quad \text{where} \quad \sum_j p(j) = 1, \tag{2.10}$$

and represents the *incertitude* of the possible outcomes and as such also the randomness of the system from which we extract information. Entropy, therefore, also quantifies the information that *could be* acquired from a certain system. In other words, entropy is strictly connected with how much disorder a system displays. In particular, in a first approximation, increase in entropy means increase in disorder, while decrease in entropy means growing order. For this reason, as mentioned, entropy is a dynamical quantity (which in thermodynamics is connected with the ability or inability to do work) while information is formal. In other words, disorder is always the result of some irreversible dynamical process, while information can be processed in pure reversible way, as quantum systems show. Many even identify an increase in entropy with the increase of heat, which is the thermodynamic expression of disorder. This is, however, not correct, for even at an absolute-zero temperature there is still a residual non-thermodynamic entropy shown by

[44]See [*KHINCHIN* 1957] for a short and effective introduction to these matters. [45][SHANNON 1948].

[46]The properties of logarithm and of probability justify why entropy is defined as in Eq. (2.10). Indeed, the joint probability of independent events A and B is $p(A, B) = p(A) \cdot p(B)$ but the entropy of the whole state obtained by the combination of the states associated or determined by the events A and B is $H(A, B) = H(A) + H(B)$.

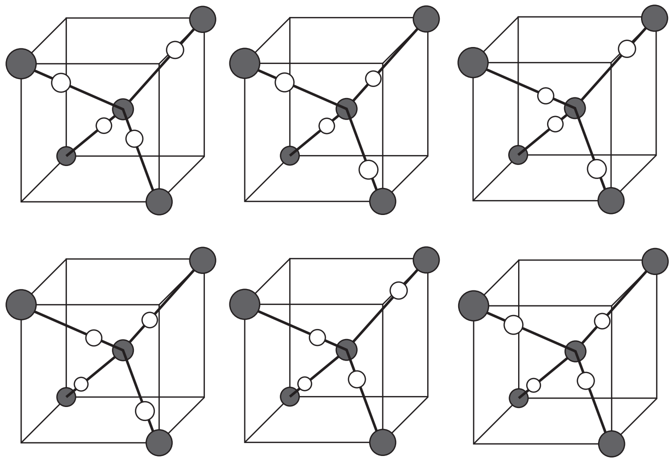

Fig. 2.7 The six possible locations of hydrogen atoms (white circles) relative to oxygen atoms (grey circles) in an ice crystal. Two of the hydrogen atoms must be near and two far away from the central oxygen atom. This incertitude about the location of the hydrogen atoms means that there is entropy even at an absolute-zero temperature.

the different possible arrangements of the atoms. This is evident for ice molecules[47] [Fig. 2.7] and shows that entropy in its general form is not necessarily connected with heat exchanges but with the (dynamic) spontaneous tendency to disorder or to display disorder.

This justifies the connection that I wish to establish between Boltzmann (thermodynamic), Shannon (information-theory), and von Neumann (quantum-mechanical) entropies. It is indeed possible to write down a quantum counterpart of Eq. (2.10).[48] Now, it turns out that most quantum-mechanical systems (some of those showing interference terms) have a zero von Neumann entropy—this is the reason why I have suggested that the universe as a whole may have zero entropy if it obeys to quantum-mechanical laws [Subsec. 2.2.4]. This is due to the fact that features are the strongest forms of interdependency between systems in our universe. As a consequence, quantum systems are the most ordered systems in nature. This confirms the fact that they have an infinite amount of potential information, though it is inaccessible to any information acquisition [Subsec. 2.2.2]. Moreover, this explains why it is necessary to lose features in the environment and to locally increase the entropy of the system when measuring [Subsecs. 2.2.3–2.2.5]: It is indeed impossible to extract information from a system that is too ordered (as well as when it is totally disordered). Information acquisition is possible when there is a sort of trade-off between an entropy that is too high and an entropy that is too low.

We may now use the classical treatment for developing considerations that are also valid for quantum mechanics. When we have two systems characterized by the sets J and K of elements or characters $j \in J$ and $k \in K$, the conditional entropy

[47] [*ATKINS/DE PAULA* 2006, pp. 609–10].

[48] For any density operator $\hat{\rho}$ describing a quantum state, which is a generalization of the concept of projectors [Subsec. 1.2.4], the so-called von Neumann entropy is indeed $H_{VN}(\hat{\rho}) = -\mathrm{Tr}(\hat{\rho} \lg \hat{\rho})$.

$$H(J|K) = -\sum_j \sum_k p(j,k) \lg p(j|k) \tag{2.11}$$

means the incertitude that the set J of the output signals will occur if the set K of the input signals also occur, or, in other words, how much the disorder of the system described by the parameter set J depends on the disorder of the system described by the parameter set K. Another important quantity is the total joint entropy $H(J, K)$, which is given by

$$H(J,K) = H(J) + H(K) - I(J:K). \tag{2.12}$$

This quantity is the sum of the entropy of the two systems separately, minus the information they share given by $I(J:K)$. The latter is the *mutual information* between the sets J and K. Starting by the simple formula

$$I(J:K) = H(J) + H(K) - H(J,K), \tag{2.13}$$

that is banally implied by Eq. (2.12), and using the derivation

$$\begin{aligned} H(J|K) &= -\sum_j \sum_k p(j,k) \lg p(j|k) \\ &= -\sum_j \sum_k p(j,k) \lg \left[\frac{p(j,k)}{p(k)}\right] \\ &= -\sum_j \sum_k p(j,k) \left[\lg p(j,k) - \lg p(k)\right] \\ &= -\sum_j \sum_k p(j,k) \lg p(j,k) + \sum_k p(k) \lg p(k) \\ &= H(J,K) - H(K), \end{aligned} \tag{2.14}$$

where I have made use of Eq. (2.3), of the fact that

$$\sum_j \sum_k p(j,k) \lg p(k) = \sum_k p(k) \lg p(k), \tag{2.15}$$

and of the two properties (2.9) of logarithms, it is possible to define the mutual information as [Fig. 2.8]

$$\begin{aligned} I(J:K) &= H(J) + H(K) - [H(J|K) + H(K)] \\ &= H(J) - H(J|K) \\ &= -\sum_j p(j) \lg p(j) + \sum_{j,k} p(j,k) \lg p(j|k), \end{aligned} \tag{2.16}$$

where I have made use of Eqs. (2.10), (2.11), and (2.13). The two terms in the second (or third) line of Eq. (2.16) are also called input information ($H(J)$) and equivocation ($H(J|K)$), that is, the conditional entropy of an output on a different input.

Note that mutual information is a symmetric quantity, so that we have both

$$I(J:K) = H(J) - H(J|K), \quad \text{and} \quad I(J:K) = H(K) - H(K|J). \tag{2.17}$$

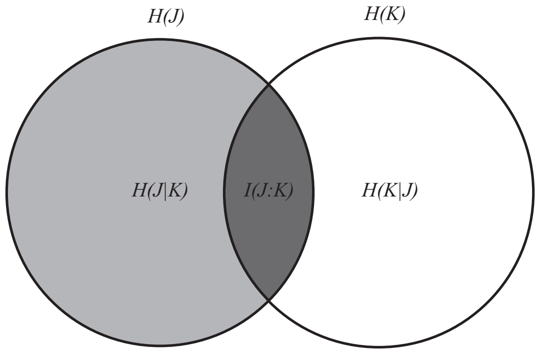

Fig. 2.8 Graphic representation of mutual information: It is easy to verify that $I(J : K) = H(J) - H(J|K)$, where $H(J)$ is the whole set on the left (both dark and light grey regions), while $H(J|K)$ is its light grey part and $I(I : K)$ is its dark grey part, respectively.

Therefore, when the two systems (or the input and output information) are independent, Eq. (2.12) reduces to the sum of the independent entropies of the subsystems:

$$H(J, K) = H(J) + H(K). \tag{2.18}$$

The concept of mutual information is very important because, as we have seen, it covers the domains of both classical and quantum (ebits) information theory, and expresses all forms of interdependencies between systems or parts of a system having informational value. To this extent, it can be seen as a measure of the order of a system. In fact, by rewriting Eq. (2.16), we see that the entropy of a given system turns out to be[49]

$$H(J) = I(J : K) + H(J|K), \tag{2.19}$$

that is, as a combination of order (the first term on the left) and disorder (the second term on the left) relative to a reference system described by K. We can generalize the previous equation by taking the second system as the environment E (the rest of the world), i.e.

$$H(J) = I(J : E) + H(J|E). \tag{2.20}$$

The quantum counterpart of formula (2.16) or (2.12) can be easily written by replacing $H(J, K)$ by the von Neumann entropy of the compound system, and $H(J)$ and $H(K)$ by the von Neumann entropies calculated on the two subsystems taken separately.[50] In other words, both classically and quantum mechanically we can treat any physical system as an open system whose entropy depends on its entropic (dynamic) relations with the environment. We obviously have two limiting cases:

- The case in which the system shares no information with the environment, and can be considered a true monad. Here the entropies of the system and of the environment are separated and the total entropy is just the sum of these separated entropies: In analogy with Eq. (2.18) we have $H(J, E) = H(J) + H(E)$.

[49] See also [*BONSACK* 1961]. [50] [AULETTA 2005b].

- The case in which the system is identical with the environment (it shows no difference). Here the entropy is the same, that is, $H(J) = H(E)$. This is exactly the case for quantum entangled systems with zero entropy. In this case, any subsystem shows the same entropy (zero) as the whole system, that is, $H(J, K) = H(J) = H(K)$.

This formalism can also be applied to a single system and its parts. In the easiest case in which we only have two subsets, J, K, using Eq. (2.14), we can express $H(J)$ and $H(K)$ as

$$H(J) = H(J, K) - H(J|K), \;\; H(K) = H(J, K) - H(K|J), \tag{2.21}$$

so that, thanks to Eq. (2.12), we have:

$$H(J, K) = [H(J, K) - H(J|K)] + [H(J, K) - H(K|J)] - I(J : K), \tag{2.22}$$

which implies [see again Fig. 2.8]

$$H(J, K) = I(J : K) + H(J|K) + H(K|J), \tag{2.23}$$

where, as usual, the first component represents the *order* of the system.[51] Denbigh[52] said that an organized system is one that can perform certain functions by virtue of its particular assemblage of parts. However, at the most basic level, the definition of order is purely physical and does not need the notion of function, which is strictly biological. The second and the third components of Eq. (2.23) represent the *disorder*. This can easily be understood if we consider that the independence of elements is a clear manifestation of the disorder of a given system. This makes my previous statement about entropy meaning disorder more precise: The conditional entropies of the form $H(J|K)$ and $H(K|J)$ represent disorder, while the entropy $H(J, K)$ represents the whole amount of both order and disorder of a system.[53] Generalizations to n-dimensional systems are straightforward but cumbersome when n grows.

A final consideration is the following: We have seen that according to the Landauer and Bennett's theorem only selection (erasure) of information costs energy [Subsec. 2.2.1]. This cost has been precisely quantized: The erasure of any classical bit of information will cost downloading $k_B T \ln 2 = 0.6931$ into the environment,[54] which establishes a strong connection between information and entropy—we shall come back to this in the appendix at the end of this section. Recall that $\ln = \log_e$, where e is the exponential function, k_B the Boltzmann constant (a thermodynamic quantity), and T the temperature of the system. Summing up [Sec. 2.1],

It is information selection that leads to a local growth of entropy.

2.3.4 Stored Information

In the previous subsection we have considered the relations between information and entropy. Let us have a closer look. High entropy in general means a high capacity to transmit (or acquire) information, since the acquired information displays the characters of high surprise, unexpectedness. This does not mean, however, that actual information grows as disorder grows. In

[51] Eq. (2.23) allows us to immediately understand why for pure quantum states (that have zero total entropy) the conditional entropies must be negative (to counterbalance the mutual information among subsystems).

[52] [*DENBIGH* 1951] [*DENBIGH/DENBIGH* 1985].

[53] Also in [LANDSBERG 1984b] a difference between the concepts of entropy and disorder is introduced.

[54] [PLENIO/VITELLI 2001].

978–0–19–960848–5 02-Auletta-02-drv Auletta (Typeset by SPi) 55 of 862 June 1, 2011 17:18
OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

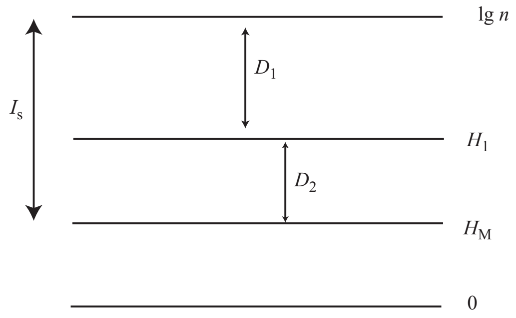

Fig. 2.9 Relationships between maximal entropy ($\lg n$), actual entropy of a single system (H_1), Markov entropy (H_M), mutual information (D_2), and stored information (I_s).

fact, as mentioned, information needs a certain amount of order and structure, that is, low entropy. However, too much order prevents information acquiring, as it is the case for quantum systems.

Stored information varies inversely with entropy.[55] In order to introduce this concept, let us first introduce a very useful concept, that of "entropy difference". This is the gap between the maximal entropy which a particular system (denoted by the subscript 1) could reach and the entropy it actually has [Fig. 2.9]:

$$D_1 = H_1^{\text{Max}} - H_1, \tag{2.24}$$

H_1^{Max} being the maximum value of entropy and $H_1 = H(J)$ being the actual value of the entropy of the system. As is well known, the maximal entropy H_1^{Max} exists when all possible states are equiprobable (maximal randomness), that is, when $p(j) = 1/n$, where n is the number of the possible states. In this case we have [Eq. (2.10)]

$$\begin{aligned} H_1^{\text{Max}} &= -\sum_n \frac{1}{n} \lg \frac{1}{n} = -\lg \frac{1}{n} \\ &= \lg n, \end{aligned} \tag{2.25}$$

since $\sum_n 1/n = 1$. Recall that we have defined the entropy of a system as in Eq. (2.19). Consequently, the information that is *stored* by a system, which is particularly important for biological systems, is also called information density and is given by

$$I_s = D_1 + D_2. \tag{2.26}$$

In other words, in considering information storing, we must also take into account at least a second system (indicated by the subscript 2). Information storing represents the necessary condition for structural information. D_2 is given by the difference between H_1 and the conditional entropy, $H(J|K)$ (which I have also indicated as H_M because it is the so-called Markovian entropy), that is,

$$D_2 = H_1(J) - H_M = H_1(J) - H(J|K), \tag{2.27}$$

in accordance with the above formulation of mutual information (2.16), that is, $D_2 = I(J : K)$. Then, the stored information may also be written as [Eq. (2.24)]

[55] [*GATLIN* 1972].

$$I_s = D_1 + D_2 = H_1^{\mathrm{Max}}(J) - H_1(J) + I(J:K), \tag{2.28}$$

where I have explicitly indicated the dependence on J. In the previous formula, I have only taken dependencies between two possible systems into account. This formalism can be generalized to a multidimensional entropy (when many systems are involved). In this case, Eq. (2.26) may be generalized as $I_s = D_1 + D_2 + D_3 + \ldots$, where the rising index indicates the growing number of interrelated systems. This leads to a simple generalization of Eq. (2.28) that considers the mutual information between a system and the entire universe, on the same line of Eq. (2.20).

Let us consider these quantities from the point of view of the transmission of a coded message. D_1 is a measure of the classical *variety* of the message, while D_2 is a measure of its *reliability*. D_2 is not allowed to drop below a certain limit because of the resulting inability to detect and correct errors, due to the lack of context. The entropy

$$H(J|K) = H_M = H_1(J) - D_2 = H_1^{\mathrm{Max}} - D_1 - D_2, \tag{2.29}$$

where I have used formula (2.24), is the entropy of a language possessing D_1 variety and D_2 reliability,[56] in accordance with what was said in Subsec. 2.3.2. The most important point of this examination is that, in any classical transmission of information, an optimal trade-off between these two forms of information must be found, i.e. between (a) a high capacity to send information, which is proportional to high entropy (i.e. to D_1), and (b) a high capacity to receive information, represented by D_2.

Therefore, I wish to stress that the concept of stored information (and also that of structural information, as we shall see) does not reduce to that of mutual information, but also comprehends an entropic part (in the case of the stored information this is represented by the component D_1). The entropic component is necessary for a structure to emerge. Indeed, quantum systems that have zero entropy [Subsec. 2.3.3] and thus show a high degree of mutual information through entanglement, do not possess any structure in the proper sense of the word.

2.4 A General Framework

2.4.1 Physical Processes

An important theoretical consequence of the previous investigation is that we should distinguish between (i) pure disruption of an existing order representing entropy growth, which is irreversible, and (ii) a reversible and deterministic (information) processing[57] [Subsecs. 1.2.7–1.2.8]. Most interaction processes show both reversible (pure formal) and irreversible (entropic) aspects and their roots can be found in the quantum theory of open systems. We have indeed noted that when a measurement-like interaction occurs, the local entropy of the object system or the apparatus increases, while the entropy of the rest of the universe decreases in order to have a balance [Subsec. 2.2.4]. We are then authorized to point out three main types of physical processes.

- When we have local increase in disorder, accompanied often by a structure's breakdown and by the mixing of some previous stuff, we speak of mechanical processes.
- When there is a decrease in local entropy which promotes the constitution of new regularities, interdependencies, and structures, we speak of an *order-building* process (which may consist information acquiring and exchanging between physical systems) [Subsec. 2.3.3].[58] W. Wundt

[56] [*NICOLIS* 1986]. [57] [LAYZER 1990]. [58] [*SPENCER* 1860–62, pp. 250 and 464].

Fig. 2.10 A constant-entropy (perhaps a zero-entropy) universe in which local disruptive and order-building processes do not affect the configuration of the whole. Actually an adiabatic expansion of our universe can keep the entropy constant and at least at a value very near to zero.

spoke in this case of positive molecular work denoting the disorder-increasing process as negative molecular work.[59]

- The universe as a whole, at a very *global* level, proceeds in its reversible, law-like quantum evolution [Fig. 2.10]. This is strongly suggested by the fact that for every irreversible local (computational) process, a larger reversible context can be found.[60]

Entropic processes are irreversible, so that here we can clearly establish what exactly the succession is between events. However, in principle we cannot predict single events, even in many classical cases. For instance, how many ways are there to break a cup [Subsec. 2.2.3]? It is likely that there are infinite ways. Let us simply take one of them, say, smashing a cup on the ground; how many ways are there to break it now? Again the answer is infinite. Even if we perfectly know the imperfections of the material, so that we can foresee that a break line will pass through such a point, we cannot foresee the whole breaking pattern. Some of these incertitudes derive ultimately from quantum mechanics. Indeed, there are as many possible outcomes of a measurement process as those representing an eigenstate of the measured observable. Statistical fluctuations are ultimately a consequence of this fundamental principle. This is the reason why quantum events are the source of any novelty in our world.[61] On the contrary, lawful processes are reversible and determine exactly what the next state will be (or what the previous one has been) given the knowledge of the actual state and its environment. However, by observing two states, we cannot predict what comes first or next. This is another expression of complementarity [Subsecs. 1.2.6 and 1.3.3].

The word *mechanical* has here a certain ambiguity. It can be understood either

(1) In the sense of mechanical engines, which are subject to the laws of thermodynamics or
(2) In the sense of the theory called classical mechanics, which arose between the end of the 18th century and the beginning of the 19th century, and described systems evolving in a perfect, frictionless, reversible, and independent way.

[59][*WUNDT* 1893, Ch. 3]. [60][TOFFOLI 1980] [FREDKIN/TOFFOLI 1982]. [61][*SMOLIN* 1997, p. 44].

In the second sense, classical systems defined in this way represent a pure idealization. Quantum mechanics (as far as the interference terms of a system are not destroyed) fulfill the dream of classical mechanics insofar as quantum systems realize a perfect example of law-like and reversible physical processes. Obviously, existing classical or semiclassical systems fall under category (1), and it is in this sense that I am using the word *mechanical process* here.

With regard to the order-building process, some have believed that it can be cast in a general principle of creation of structures through combination of some discrete unities (the code), by generalizing what was said in Subsec. 2.2.2 also to classical systems.[62] This is known as the *particulate principle* and was first proposed (for explaining biological heredity) by Ronald Fisher.[63] However, it is difficult, as we shall see, to extend this principle to *any* creation of order. It remains true that such a principle applies to a lot of systems. All physical systems that make infinite use of finite means (according to von Humboldt) conform to the particulate principle[64]: Discrete units from a finite set of meaningless elements are repeatedly sampled, permuted or combined in larger units that are higher in the hierarchy (in general, independence of the different levels of the hierarchy should be maintained).

The particulate principle rationalizes both the combinatorial mechanisms and the hierarchical structure that is ubiquitous in nature. However, this does not mean that such a combinatorics can be identified with information codification, which occurs *either* at a very basic level (as in quantum systems) *or* at a very high level (as in living systems). An application of the particulate principle in a prebiotic domain, according to Abler, is represented by chemistry: Here, electrons, protons, and neutrons could be understood as a code whose syntactic combination (corresponding to the physical location in the nucleus and orbitals) give rise to the different sorts of elements. Since these elements are finite (even if they are of a far greater number than electrons, protons, and neutrons), they could also be understood as a code giving rise, by combination, to all different sorts of stuff. However, the matter structure as such does not show the typical aspects of information codification. Indeed, the only character of codified information that is retained here is the combination of a finite set of elements, but it is difficult to see what the other alternative codes are; we have neither the combination nor the translation rules that are necessary for speaking of information codification. As mentioned, we must wait for living systems in order to see coding activity again, i.e. the genetic code. The reason why the typical organization of codified information becomes lost in the passage from quantum mechanics to the molecular and chemical level and also affects many systems of our macroscopic world is a very great and difficult problem. With the growing complexity of the interacting systems, it is possible that the simplicity of a linear combinatory of a small set of elements gets lost. Here, non-linear effects may also play an important role. My guess is that the domain between quantum mechanics and living organisms is simultaneously too complex for quantum information (interference terms get lost) and too elementary for classical information coding and processing, since the conditions through which these processes may be shielded against environmental perturbations have not yet arisen. I shall come back to this point in the second part of the book. So, order and information, as explained [Subsec. 2.3.4], are not equivalent concepts, even if any order has its roots in the quantum or any other higher-level information codification.

Summing up, here and in the following chapters, I try to support a more generalized understanding of dynamics in which the interplay between global and local aspects is taken into account. Its utility could be measured by its ability to help the future progress of biology.[65]

[62][ABLER 1989]. [63][*FISHER* 1930]. [64][STUDDERT-KENNEDY 1998, STUDDERT-KENNEDY 2000].
[65][THOMPSON 1995b].

2.4.2 Emergence

Laws are underdeterminate relative to the domain where they apply.[66] The more they are general, the more they are underdeterminate. This is clear for quantum mechanics, where laws have a general character [Sec. 1.1; Subsecs. 2.2.5 and 2.4.1], but this is a lesson of general validity.[67] This underdetermination allows for novelty: Novelty is something that cannot be foreseen or even understood only on the basis of some previously known laws and which notwithstanding happens within a general framework provided by those laws [Subsec. 1.2.1]. It comes out from the intrinsic variety of the lower-level entities from which a new configuration is built. For instance, organisms do not violate the second principle of thermodynamics, according to which entropy grows or remains constant in an isolated system.[68] However, the way in which organisms build order is not simply a consequence of thermodynamic laws even if it happens in accordance with them.[69] Every time that such a phenomenon happens, we must have the emergence of a new structure that somehow represents a new fundamental property or behavior. Such a property or behavior is then robust to fluctuations or variations of variables or conditions out of which is emerged, becoming in this way independent of the details determining the latter.[70] This new property or behavior can become in the long run a new "regional" law. As I shall explain in the following, the *driving* force of emergence is selection operating canalization, while the *adaptation* is represented by different sorts of constraints contributing to channelization [Subsec. 2.2.6]: This is the essence of generalized Darwinism that I have introduced. We shall indeed learn that features are only a special (quantum-mechanical) case of formal constraint.

Emergence is a widespread phenomenon that covers many physical situations of our world.[71] Among the first scholars to have thought about it was John Stuart Mill, who affirmed that the addition of causes does not necessarily imply the proportional addition of effects, but may result in a new configuration.[72] Unfortunately, traditional philosophical positions seem to be strongly committed to a radical reductionism, to a metaphysics of elementary particles that, after quantum mechanics, seems to be anachronistic.[73]

We can speak of a *direct* emergence which relies on properties and relations between individual elements and of *indirect* emergence when the relations between individuals are mediated by active and often complex environmental structures.[74]A case of direct emergence is the traffic jam. Here we see that complex structures are in a constant flux. An example of indirect emergence are termites' buildings[75]: A modification of the local environment is a response to the triggers provided by previous alterations to the environment (made by termites at an earlier time). This is called a stigmergic signal (usage of work as the signal for more work). Indirect emergence is very important for living beings due to their capability of canalizing environmental cues in an appropriate way.

When one speaks of emergence, one often takes into consideration the problem of levels of reality. This is not the place to extensively discuss this point, but I wish to say that the term "levels of reality" is sometimes ambiguous, while emergence is certainly concerned with *properties*

[66][PEIRCE 1891, p. 296]. [67][*BOREL* 1920]. [68][*ATKINS* 1994]. [69][LOTKA 1922b].

[70][*BATTERMAN* 2002] [*MITCHELL* 2009, pp. 14–15 and 21–6]. In this book I shall be concerned with all five forms of emergence enumerated in [KIM 1999].

[71][LAYZER 1976] [*CHAISSON* 2001] [*MOROWITZ* 2002] [ELLIS 2004, ELLIS 2005a, ELLIS 2008b]. See especially the very useful book [*MURPHY/STOEGER* 2007] as well as [*CLAYTON/DAVIES* 2006].

[72][*MILL* 1843, III, Ch. 6].

[73]J. Smart ones said that he could not "believe that ultimate laws of nature could relate simple constituents to configurations consisting of perhaps billions of neurons ... all put together for all the world as though their main purpose in life was to be a negative feedback mechanism of a complicated sort" [SMART 1959]. This turns out to be almost literally true for life.

[74][CLARK 1997, pp. 73–75]. [75][BECKERS *et al.* 1994]

that, as mentioned previously, are neither a simple consequence of other laws nor in contradiction with them. Quantum mechanics is a general paradigm that allows for the emergence of both classical or semiclassical systems and biological systems.[76]

In this way, as suggested in Sec. 1.1, we can understand the emergence of biological systems as both something new (and therefore displaying properties and principles that are additional to) and in accordance with the laws of physics (as well as of chemistry).[77] I shall define the philosophical point of view supported in this book as an *emergent monism*.[78]

2.4.3 Structures and Elements

A superposition of different eigenstates of an observable that is measured only represents potential information [Subsec. 2.2.2]. It displays how the system being in this state *could* behave if it interacted with an apparatus and the environment. *A priori*, what guarantees the connection between global features and local events are the coefficients entering into an expansion of a quantum state in a given basis (and therefore also the probabilities of those events that are calculated by performing the square modulus of these coefficients). *A posteriori*, this connection is provided through the interaction with at least one other open system.

This essential character is also true for any structural relation in the classical world. Any structure, in fact, represents a potentiality to behave in a certain manner *given* certain environmental conditions. For instance, a crystal structure in itself only represents a disposition to behave in a certain manner and to produce certain effects when stimulated by scratching, sounds, light, and so on. Peirce had already individuated this point, although he seems to have interpreted it in a purely epistemological style.[79] Another example can be represented by a forest. From a certain point of view (fine-graining), only the individual trees seem to be ontological realities and the forest, on the contrary, seems deprived of any reality apart from that conferred by the human observer. As we shall see, this is true to a certain extent. However, it is also true that the forest (that is, the specific disposition of the trees) will have a crucial role when purely physical agents are in play, e.g. against wind or fire. For this reason, structures cannot be dismissed as illusionary phenomena.[80] In other words, I am suggesting that any structure or arrangement of elements should be considered as a complex of *formal constraints* that can be *activated* in certain conditions. Now, from the point of view of the structure, the individual existence of one tree or another is completely indifferent. In fact, nothing essential would change if, in place of a given tree, there were another of the same species or at least one of a similar shape and size. This is almost true for any tree of the forest. On the other hand, the existence of the individual trees is the only feature that guarantees that we can speak of a forest. We see, therefore, that there is a certain asymmetry between global structures and localized objects and events [Subsec. 2.2.5], but also that both, in their manner, can be perceived as constitutive of our world.

The reason why structures have this character is that they can be thought of as based on the mutual information or on any relation between connected elements.[81] The concept of potentiality

[76][AULETTA 2005b]. [77][*GOULD* 1985, pp. 379–80].

[78]A concept developed by Lloyd Morgan [*LLOYD MORGAN* 1891, pp. 471–2][*LLOYD MORGAN* 1923]. See also [*CLAYTON* 2004]. For the concept of monism see [*HAECKEL* 1899]. See also [PEACOCKE 1999] and [*BUNGE* 1980] for a materialist understanding. For a more recent discussion see [ELLIS 2005a].

[79][PEIRCE 1872]. Nevertheless, the reduction of pragmatism to a species of relativism without ontological import does not seem very appropriate [*MARGOLIS* 1986] and has not produced any relevant result.

[80][*LAUGHLIN* 2005].

[81]This connection between information and structure was stressed in [*VON WEIZSÄCKER* 1971, pp. 50–5, 346–48].

may sound strange to scientifically educated ears, even if I have followed Heisenberg in applying it to quantum mechanics [Subsec. 2.2.2]. In all its generality, let us define potential as something that[82]

- Can be an ingredient of certain dynamical processes in which local interactions also occur,
- Contributes to the final outcome of the dynamical process,
- Needs some additional condition in order to be "activated," that is, to actually concur in determining that outcome.

Natural sciences before the end of the 19$^{\text{th}}$ century and the birth of quantum mechanics were strongly reductionist. For instance, the dominant point of view in chemistry was the atomistic one offered by Dalton, according to which the goal of chemistry is to individuate the elementary substances. However, in 1825 Liebig and Wöhler had already discovered that two different silver compounds, silver cyanate and fulminic acid, though having the same chemical composition (the same proportion of carbon, nitrogen, and silver), showed different observable properties. The silver compound of fulminic acid was explosive, whereas silver cyanate was not. These and similar substances, called "isomers" by Berzelius, led chemists to suspect that substances are defined not simply by the number and kind of atoms in the molecule but also by the arrangement of those atoms, that is by their spatial relations.

2.4.4 System Theory

In order to avoid both dangers of an ontological reductionism and ontological holism,[83] this book can be seen as a treatise in applied system theory (a System biology[84]). The reason for system theory's birth is that the scheme of isolable causal trains and meristic treatment has proved insufficient in many fields and especially in biological and social sciences.[85] System theory attempts to provide theoretical explanation and scientific interpretation where previously there were none, anyway with higher generality.[86] Traditionally, analytic procedure in science supposed that interactions between parts of a system are non-existent or weak enough to be neglected, and that relations are linear. I invite the reader to consider system theory together with information theory as the formal tools that may ground biology and cognition as traditional mathematics grounds physics.

As we have seen, a system can be broadly defined as an ordered complex of interrelated elements [Sec. 1.1], rather than "interacting," as originally defined by von Bertalanffy. When dealing with systems, three features must be considered: The number of elements, their species, and their relations (structure).[87] When systems like the organisms evolve, there is increasing modularization, which means increasing determination of the functioning of elements which are only dependent upon themselves, and consequently a loss in regulability which rests on the system as a whole, owing to the interrelations that are present. As we shall see, this phenomenon is typical in biology. Only this differentiation,[88] or decrease in interaction between parts, can determine progress. However, this process can never be fully accomplished without destroying the system: There is a tension between wholeness and pure summation (atomism). Therefore,[89]

[82] [AULETTA/TORCAL 2011]. [83] [*ROBERT* 2004, pp. 69–70]. [84] [BIZZARRI *et al.* 2008].
[85] [*VON BERTALANFFY* 1969b, pp. 3–29]. [86] [AGAZZI 1978]. [87] [VON BERTALANFFY 1950].
[88] [*SPENCER* 1860–62, pp. 277–95].
[89] [*SPENCER* 1860–62, pp. 291–5 and 324–7]. See also [*LLOYD MORGAN* 1891, p. 241].

Progressive segregation must be accompanied by at least some form of integration or even a progressive centralization of the parts.

This statement, which may be called Spencer's law, is in accordance with a quantum-information point of view [Subsec. 2.2.5]. This is also the issue of complexity, as we shall see.

In the next chapter I shall begin to develop the specific contents of this system theory. In the present, very abstract context, it is useful to consider system theory as a generalized theory of relations and interactions. Let us come back to Fig. 2.5. The problem can also be seen at the abstract and general level: Why three interrelated systems? For our examination, three systems, as we have seen, is the minimal level of complexity required [Subsecs. 2.3.1–2.3.3]: Indeed, any information dependency that is not a pure covariance needs at least three elements. The problem then, is why 3 and not 4, 5, or even 20? This is a problem of the theory of relations. Peirce distinguished three types of relations[90]:

- *Monadic* relation: This is a zero degree of relations, or a pure absence of relation. Mathematically, it can be represented by pure numbers. It can be geometrically represented with a point, which has zero dimensions.
- *Dyadic* relations: This is the pure (static) interdependency. In physics it expresses the covariance between several systems. Entanglement is a specific form of this interdependency. In mathematics, it is the dependence of a function on a variable or inverse relation. Geometrically, it can be represented by a line (or a vector).
- *Triadic* relation: This is a true three-term, dynamical relation. Here any element is dynamically connected with another through a third one (often a medium). Mathematically, it can be represented by matrices. Geometrically, by surfaces.

According to Peirce, any higher-degree relation can be reduced to some combination of these three. Therefore, in the most general case, we have three elements, A, B, C. These can be (1) unrelated, (2) dyadically related (AB, BC, CA: since dyadic relations are interdependencies, they are symmetric), or (3) triadically interrelated: ABC. Since any element here is dynamically connected through the other ones in a (feedback) circuit, any combination of the three is equivalent to ABC. This triadic relation is irreducible to binary relations. When we consider that the zero-degree relations consist exactly in each of the 3 elements (A, B, C) taken apart, we obtain 3 elements, 3 dyadic relations, and 1 triadic relation, the magic number seven.[91]

This can also be seen from a purely physical point of view. It is well known that forces can be vectorially added. Now, it does not matter how many forces there are in play; any pair of forces can be reduced to a single one (the resultant) through summation. By reiterating this operation we finally arrive at two force vectors (which are eventually the sum of many other vectors) and their resultant. This means that we can consider the resultant either alone in itself (monadic relation), the two forces giving rise to the resultant (dyadic relation), or the resultant *as* a resultant of these two forces (triadic relation). Things stand in a similar way for quantum-mechanical systems, since their state can be vectorially represented [Fig. 1.5]. The same is true for information codification. Any n-dimensional code can be reduced to a binary one in the same way [Fig. 2.5]. However, any binary code is really ternary: 0, 1, and their combinations. I hope to have clarified here the systemic reasons for the scheme in Fig. 2.5, which will be reiterated throughout the whole book. Under this point of view, I follow Carnap's fundamental idea that science essentially consists in description of relations and structures.[92] It is also true that structural descriptions, although basic,

[90][*PEIRCE CP*, 1.293; 1.303–32; 3.472–3]. [91][MILLER 1956]. [92][*CARNAP* 1928, Ch. 2].

978–0–19–960848–5 02-Auletta-02-drv Auletta (Typeset by SPi) 63 of 862 June 1, 2011 17:18

OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

do not cover the complexity of biology and cognition (they describe the physical level even if they are also used in biology and cognition). For this reason, in the next part of the book, I shall introduce functional descriptions, and in the third part mental and social acts. Moreover, I stress that scientific explanations also rely on mechanisms and not only on descriptions.

2.5 Concluding Remarks

In this chapter, we have seen that quantum mechanics allows us to think in a new and original way about some very important physical issues. With regard to the issue of information, we have learned that:

- Information and entropy are not the same quantity: Information is a formal quantity, while entropy (especially as conditional entropy) is a measure of the disorder of the system.
- Information is basically codified information. In this form it is essentially potential information (information that in certain contexts can be acquired). However, there is never access to the whole information potentially contained in a system.
- Information can be processed, shared, or selected. An appropriate combination of these three aspects gives rise to information acquisition.

Moreover, I have shown that:

- There is a fundamental complementarity between global and local behavior. This complementarity should be understood in dynamic and smooth terms.
- Global behavior is, according to the laws of quantum mechanics, continuous and wave-like. Global behavior does not determine local behavior.
- Local behavior is random, discontinuous, jump-like. Local behavior is blind to global behavior.
- This complementarity is the source of any emergence in our world.
- We probably live in a universe with constant (perhaps zero) entropy with two different processes: An order-building one and an order-destroying one.

Apart from these specific results, I wish to recall here a general methodological (and epistemological) lesson and the first fundamental principle (valuable not only for quantum systems) we have drawn. The methodological lesson is the following: We need to consider problems both from a global and local perspective and these two different views need to be integrated. This means that I shall use both a reductionist methodology and an approach making use of top-down processes, as will be explained in the following. This also means that I shall avoid a pure subjectivist or idealist approach but also a traditional objectivist one, putting the stress on both local perturbations and interdependencies among systems. The principle can be formulated as follows: In certain conditions, especially when there are interaction processes, nature can operate selections among several possibilities.

Appendix: The Relation Between Shannon and Boltzmann Entropies

An important relation is that between the Shannon (informational) and the Boltzmann (thermodynamic) entropies. Let us consider the arrangement shown in Fig. 2.11.[93] The Boltzmann entropy of all molecules is given by

[93] [*ROEDERER* 2005, pp. 173–87].

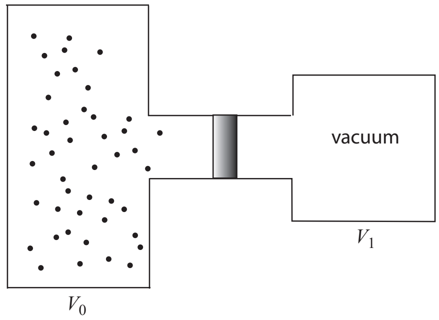

Fig. 2.11 Two vessels, with volumes V_0 and V_1, are initially separated by a bar that is removed in order that the gas contained in the first vessel can flow toward the second one.

$$S = k_B N \ln W, \tag{2.30}$$

where k_B is the Boltzmann constant, N the number of gas molecules, W the number of possible configurations of the gas particles, and the natural logarithm is given by $\ln x = \log_e x$, where e is the exponential (Euler) function [Fig. 2.2]. I am assuming here that in the final configuration every particle can be either in the vessel with volume V_0 or in the vessel with volume V_1. Once the bar separating the vessels with volumes V_0 and V_1 is removed, the Boltzmann entropy is increased by

$$\begin{aligned}\Delta S &= S_f - S_i = k_B N(\ln W_f - \ln W_i) \\ &= k_B N\left[\ln(V_0 + V_1) - \ln(V_0)\right] = -k_B N \ln \frac{V_0}{V_0 + V_1},\end{aligned} \tag{2.31}$$

where S_f and S_i are the final and initial Boltzmann entropy, respectively. The number W_f of possible final configurations (states) in which the gas can be is proportional to $V_0 + V_1$ (since the gas occupies both vessels) while the number W_i of the initial one is proportional to V_0 (the gas is here confined to the left vessel).

On the other hand, the initial Shannon entropy per molecule when it is maximal with some simplifying assumptions that do not deprive this formalism of its generality, can be taken to be $H_i = \lg W_i$ as well as the final one as $H_f = \lg W_f$ [see Eq. (2.25)], since the probability as inversely proportional to the number of possible configurations. Then, the increase of Shannon entropy is given by

$$\Delta H = N\left(\lg W_f - \lg W_i\right) = -N \lg \frac{V_0}{V_0 + V_1}. \tag{2.32}$$

By comparing the latter two equations and considering the relations between the natural and the binary logarithm ($\ln x = \lg x \ln 2$), we have

$$\Delta S = k_B \Delta H \ln 2, \tag{2.33}$$

or

$$S = k_B N \ln W = k_B N \lg W \ln 2$$
$$= k_B N H \ln 2. \tag{2.34}$$

It is obviously meaningful to associate a Shannon entropy to a thermodynamic entropy (since any increase in thermodynamic entropy also means increase in disorder), but the reverse is not necessarily the case, as Fig. 2.10 nicely shows.

3
The Brain: An Outlook

In the next three chapters I shall show some basic information-acquiring processes and systems that are relevant for biology and cognition. The subject of this chapter is the brain, which is the most studied system for information-acquiring. In Ch. 4 we shall investigate a specific information-acquiring modality, namely vision. In Ch. 5 we shall deal with the issue of motion. As we shall see, motion is concerned with both motion perception and movement controlling. This will lead us further than pure information-acquiring and open the path to the next part of the book, devoted to information control.

Now, let us briefly consider the contents of the present chapter. After having considered the reasons for dealing with the (human) brain as an information-acquiring system, I shall consider how information is processed from the peripheral sensory system to the central nervous system. I shall then give some basic information about the brain's structure. Thereafter, I shall deal with three specific problems: The brain's modularity, the stages of information-processing, and the brain's mapping activity. Finally, I shall introduce the neural networks approach.

3.1 Biological Systems Acquire Information

The archetype of any information treatment by biological systems[1] was traditionally considered as the brain, in particular the human brain, and it is until now the best known biological system which is supposed to be an information-processing or information-acquiring device. Indeed, in a biological context, the theory of information-processing was historically first applied to this complex organ. This is due to the fact that cognitive sciences were dominated (and still are in part) by a functionalist approach that underestimates the centrality of the biological dimension for the treatment of information and considers the brain as the proper place for dealing with information-processing. Obviously, information-processing does not exhaust information-acquiring [Subsec. 2.3.2], but at this stage of the examination the stress will be put on this aspect. Therefore, it should not sound strange that, in studying biological systems, we first begin with the brain in its information-processing function and then, in the next part, we shall deal with organisms (as well as with the higher functionalities of the brain itself). This allows us also to formulate many basic distinctions that will turn out to be useful in the following as well as to become acquainted with the basic schools dealing with cognition. The brain is not the sole organ to deal with information but in higher animals there are essentially three systems that treat information[2]:

[1] I use here the term biological system simply for indicating a system in the sense previously defined that represents a living organism or a part of a living organism. In the second part of the book, we shall learn a rather technical understanding of this term.

[2] [*ALBERTS et al.* 1983, pp. 879–961].

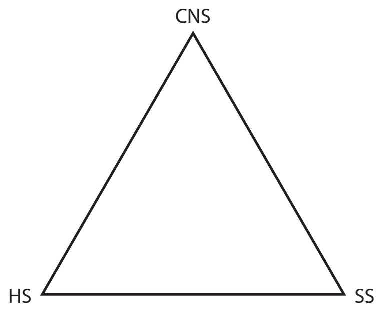

Fig. 3.1 The three systems dealing with external information and the way it is integrated in the organism and regulated by it: The central nervous system (CNS), the peripheral sensory system (SS), and the hormonal system (HS).

1. The peripheral sensory system (SS)—that is part of the peripheral nervous system, which also comprehends the motor connections to muscles—or any other elementary interface with the environment,
2. The regulative systems and in particular the hormonal system (HS), and
3. The central nervous system (CNS) or any more elementary analogue.

The three systems may be cast as in Fig. 3.1, in agreement with our previous analysis [Subsec. 2.3.2]. Speaking of vertebrates at least, the function of the sensory system is to acquire and—after a first processing—transmit information to the CNS; the specific function of the hormonal system is to regulate the transmission and the exchange of information (thus regulating both metabolic and CNS processes); the specific function of the CNS (consisting essentially of the brain and the spinal cord) is to acquire information (from the sensory system) and, in the lowest sensory and motor areas, process it, as well as to control information exchange and select specific options. The hormonal system transmits information through the hormones released by the glands. In particular, it mainly regulates the circulatory system, the digestive system, and the reproductive system, apart from its influences on the brain and the CNS. It is interesting to observe that information in the brain is also transferred through local diffusion of neurochemicals in a paracrine way.[3] Also the CNS transmits and exchanges information as well as regulates it. On the other hand, the sensory system is not able to exercise regulation or control, and, as a consequence, a very violent stimulus (a very high-pitched sound, a sudden intense emission of light, and so on) can have disruptive consequences on it. The relationships among these three systems will be clarified especially in the second part of the book. Let us now introduce some short remarks on the sensory system and the CNS:

- The peripheral sensory system is an information-acquiring system and also, more specifically, an information codifier and processor relative to the whole organism. Allow me to give evidence of how information is processed via the visual system from the sensory systems to the higher brain areas through specific algorithms.[4] As we shall see more clearly below, all visual information starts with two paths, the first for color processing (from visual area V1 to area V4), the second for motion processing (from area V1 to area V5). (I am not considering here other aspects like form processing.) The first step in color processing is provided by lightness

[3] [*PANKSEPP* 1998, p. 68]. [4] [ZEKI 2001].

algorithms to determine the lightness (or relative brightness) record of a complex scene in different wavebands. The lightness of a surface will not change with changes in intensity of light. The second step is a comparator which, confronting at least three different wavebands, is able to construct what we call color.

- Not all types of CNSs have the same structure. Indeed, here evolutionary development took place twice: In vertebrates the central nervous system forms a single mass of brain and spinal cord, whereas in the invertebrate it consists of a series of separated masses (ganglia) connected by relatively thin neural strands. Although in this chapter I will consider mainly how the human brain processes information, some of the following considerations have a more general scope, and, unless noted otherwise, they can be applied to almost any vertebrate.

3.2 Classical Computation, Cybernetics, and the Brain

Two main developments in cognitive sciences must be taken into account here:

- Traditionally, at the beginning of cognitive science, parallel to the first steps of AI, a central-processor model of brain activity dominated the classical theory of computation and the brain.[5] According to the model of the central-processor computation, external inputs are linearly and logically (syntactically, following logical laws) processed by a unit, and the result is constituted by the outputs of this central process. Representations (in the sense of explicit or implicit descriptions of objects) are assumed to be the result of a passive reception of inputs. Von Neumann was one of the first to notice the characteristic digital functioning of the nervous system.[6]

  This initial assumption was subsequently rejected along two different lines of investigation. (i) In strong opposition with the idea of a central processor, Chomsky and his followers proposed a modular model. (ii) Parallel studies in neurology (in particular due to Hebb's contribution) and in computer science (connectionism) have changed our vision of the brain toward a net-like, parallel processing device. I also mention that a further correction to pure classical information-processing in both variants—the central-processor model (as it was understood by von Neumann) and the modular computation model (as it was understood by Chomsky and Fodor)—is the introduction of feedback circuits: This was the great contribution of the father of cybernetics, Norbert Wiener.[7]
- From the perspective of a classical-mechanical physics, a strongly local and reductionist understanding of the brain initially dominated, in particular in neurology and neuropsychology. Later developments have suggested the necessity of the opposite point of view, that is, of a holistic understanding. This strong dichotomy cannot be overcome from the classical perspective. On the contrary, quantum information suggests from the start the necessity to adopt BOTH points of view and to integrate them into a wider framework.

3.2.1 Early (Classical) Models of Brain

As already mentioned, classical computation and cybernetics have determined early models of the brain. Though these two traditions gave rise to the common scientific enterprise of cognitive science and, to a certain extent, they also mixed and superposed, there are important tensions between

[5] [*GARDNER* 1985]. [6] [*VON NEUMANN* 1958]. [7] [*WIENER* 1948].

them. Very early in the history of cognitive sciences, the brain and the cognitive activities were modeled following these assumptions:

- The brain is centrally guided exactly like a classical computer with its central processor. We have here a *central regulation assumption*, strongly supported by classical computation.
- Any cognition is symbolically mediated. We have here the *symbolic assumption*, a view strongly supported by cybernetics.
- The activity of the brain essentially consists of computation, that is, of a syntactic and logical combination of symbols. This is the *syntactic assumption*, again strongly supported by classical computation.

As a consequence, like a computer, the brain was thought of from the start[8] as a *general problem-solver*. For Newell and Simon there are four kinds of elements in problem-solving: An initial state, a goal, operators to undertake actions, and path constraints (on how many steps there are to reach the solution). The difficulty here is that the size of the space of the computational paths grows exponentially with the depth of the search (the so-called combinatorial explosion).[9] For this reason, it was assumed that human beings use heuristic searches,[10] that is, employ only loosely defined rules [Sec. 1.1]. For instance, Simon and Chase remarked that expert chess players had acquired the ability to recognize large and meaningful perceptual units instead of serially searching through their memory, like conventional computers do. Often, humans perform a means–ends analysis by swinging back and forth between actual state and goal.

Another consequence of the above assumptions is functionalism.[11] In this context, it suffices to say that functionalists assume that there is a sharp separation between hardware and software (as well as between brain and body) and that several different hardwares may run the same program or algorithm. In the next part of the book we shall deal extensively with this kind of problem.

3.2.2 Cybernetics

Wiener gave the following definition of *cybernetics*: It "is the science of control and communication, in the animal and in the machine." Among the originators of the science of transmission and control of information, Wiener quotes Fisher, Shannon, and himself.[12] Three ideas were prominent for him:

(a) Cybernetics has to do with an irreversible time flow,
(b) Any process and control of information is based on feedback circuits, and
(c) Any control of information has to do only with the form (structure) of the signal, neither with related physical quantities [Subsec. 2.2.1], nor with its contents.

All three ideas play an important role in this book. They allowed Wiener to interpret information theory in a very original way that is consonant with the analysis developed in the previous chapter. According to Wiener, in all classical phenomena where considerations of probability and prediction enter, the problems become asymmetrical. In fact, one can bring a system from the past into the present in such a way that one determines certain quantities (preparation), assumes that other quantities have known statistical distributions (premeasurement), and then observes the statistical distribution of results after a given time (final detection) [Subsecs. 1.2.7–1.2.8]. This process cannot be reversed since, in order to do so, one should pick up a fair distribution of systems which, without intervention on our part, would end up within certain statistical limits (a sort of fore-preparation),

[8] [*NEWELL/SIMON* 1972]. [9] [HOLYOAK 1995]. [10] [SIMON/CHASE 1973].
[11] [*PUTNAM* 1981]. [12] [*WIENER* 1948, p. 10].

and find out what the antecedent conditions were at a previous time (a sort of retro-selection). However, for a system starting from an unknown position to end up in a tiny statistical range is so rare an occurrence that it may be regarded as a miracle. In other words, one can prepare a system in a certain way and then measure it, but not *vice versa*: Selection comes *after* preparation and regulation, as I have said in Subsec. 2.3.2. We shall see the consequences of this important aspect of the cybernetic theory of information later on.

Another father of cybernetics is W. Ross Ashby. He affirmed that cybernetics is not concerned with (mechanical) objects but with *ways of behaving*.[13] It is a functionalist and behaviorist approach from the start. Moreover, it does not depend on the property of matter or on the laws of physics as such. Cybernetics stands to real machines as geometry does to extended objects. Therefore, it is concerned with a set of possibilities: What are all the possible behaviors that can produce a given result? While biology traditionally took into consideration the problem of the available free energy for producing determinate results, Ashby, in the same way as Wiener, stressed that cybernetics is interested in the form of signal exchanging: It studies systems that are energetically open but informationally closed. This is in accordance with our previous analysis of quantum information [Subsecs. 1.2.8 and 2.2.3–2.2.4].

Cybernetics also assumes that any dynamical process develops step by step (discontinuously). This is another important point in common with quantum information, and is true for all biological systems. A change is a transition from an operand (initial set or state) to a transform (final set or state) induced by an operator. It is therefore a matricial theory [Subsec. 1.2.4]. When every element of the operand also occurs in the transform, the transformation is closed. A determinate machine is that which behaves with a closed single-valued (one-to-one) transformation. If the state of a whole system can be decomposed in the state of the parts, we can describe it with a vector. Not all transformations have immediate effect on all elements, and there are transformations in which some elements are transformed in a way that is completely independent of the whole. When the whole system can be decomposed in functionally independent parts, then it can be *reduced* to these parts.

Focusing on the organism, we may define the environment as the set of both those variables whose change affects the organism and those variables which are changed by the organism's behavior (it is a feedback system). However, before any regulation can be undertaken or even discussed, one must be able to recognize which are the essential variables of the problem faced in a concrete situation and which is the set of states permissible for the organism, that is, the organism must somehow "know" what is important for its survival and what is its goal.[14] Therefore, the problems of any organism are:

(1) How to block the flow of variety by transforming it from a disturbance into variables essential for its needs[15] (as we shall see, this is the problem of teleonomic causation and, at a higher level, of representation); and
(2) How to act on the environment according to these needs or representations. This is the issue of teleology and information control on the environment.

Being able to solve these two problems, organisms are self-regulated systems. Therefore, cybernetics provides the main systemic foundation upon which this book is grounded, and, as far as I can judge, it is a very solid one.

[13] [*ASHBY* 1956, pp. 1–72]. This is in general an excellent introduction to the problem.
[14] [*ASHBY* 1956, p. 219]. [15] [*ASHBY* 1956, p. 201].

3.3 Information from the Senses to the Brain

Throughout the whole previous chapter, we have learned through quantum mechanics that each information transmission or acquisition will be the combination of a discrete, local behavior (information selection) and of a continuous, wave-like, global behavior (coupling). The brain is one of those classical systems displaying both aspects. Indeed, it is characterized by two phenomena: Information is acquired in a *spike-like* way, in discrete terms, while all global (processing) activity, which involves many neurons or several areas, is treated in a *wave-like* way.

3.3.1 The Sensory System and the CNS

Sensation is how sense organs (of the sensory system) respond to external stimuli, while *perception* is the processing, organization, and, in a wider sense, interpretation (initially in the sensory system itself but mainly in the CNS) of sensory signals.[16] Another way to understand the distinction between sensation and perception is to introduce that between receptor (which is purely passive and stimulated) and perceptual organ (which is activated during the perception process).[17] Therefore, I understand by *sensation* the pure process of information acquisition from the external environment, and I understand *sensory coding* as the way in which our senses translate the physical properties of a stimulus into neural impulses. This is necessary, since an organism can acquire information only by a digital encoding process [Fig. 3.2]. On the contrary, as we shall see, representations as well as other higher brain activities and functions are analogical. Sensory qualities are coded by a

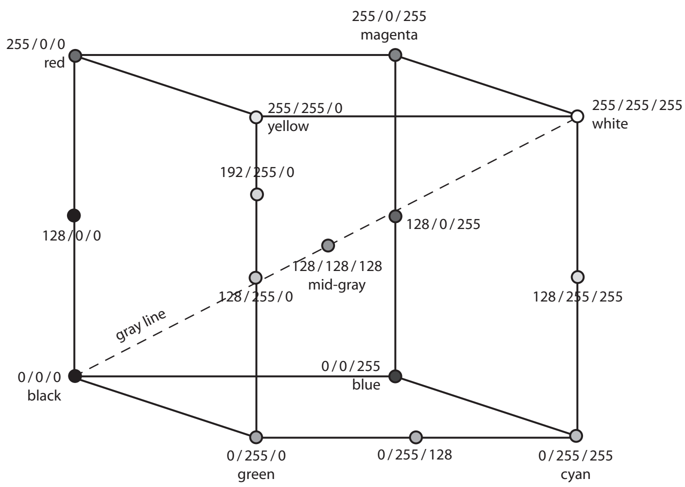

Fig. 3.2 An example of digital encoding. The color (RGB) cube. Color intensities vary from 0 (no color) to 255 (full color). White can be understood as the maximal intensity of red, green (lime), and blue, while black is the minimal intensity (no color at all). Some combinations are shown. Note that the human-vision red–green–blue system is only one of the possible color coding possibilities. Theoretically speaking, also a magenta–yellow–cyan encoding would work. (This figure is reproduced in color in the color plate section.)

[16][*GAZZANIGA/ HEATHERTON* 2003, pp. 69–94]. This is a traditional distinction that can be found, for instance, in [*CABANIS* 1802, pp. 103–5] [*HERBART* 1816, Sec. 73] [*ROMANES* 1884, pp. 125–7].

Table 3.1 Possible reaction when the signal is present or absent.

Signal	hit	miss
No Signal	rejection	false alarm

few *receptors* (coarse coding). However, most single receptors are tuned on single sensory elements and properties. For instance, each taste quality (bitter, salty, sweet, umami, and sour) is encoded by a specific receptor cell.[18] Further evidence for sensory information encoding is given by the way the skin works as a detection system: It is, from this specific point of view, a kind of computer.[19]

Therefore, receptors are specialized pseudoneurons in the sense organs that pass impulses to connect neurons when receiving physical or chemical stimulation. This activity is called *transduction*: It is a signal transformation through which the organism assimilates the external environment to itself. At a general level, organisms are very selective in the way they interact with external physical conditions.[20] In particular, all animals show selective responses to some stimuli (lower organisms show a sensitivity to a smaller range of signals than higher organisms do) irrespective of the mechanical intensity of the stimulus.[21] For instance, not all electromagnetic energy causes visual perception as such, but only visible light.[22] We re-cover here the main aspects of information acquisition:

(1) An unknown environmental stimulus will be, in a certain sense, guessed [Subsec. 2.3.2] by the sensory system through:
(2) Information-sharing with the environment that allows sensory coding and
(3) Information selection able to individuate the specificity of the signal.

To detect a stimulus requires an evaluation of its presence or absence, based on inherently ambiguous information [Subsec. 2.3.2]. There are four critical variables: Signal, response, noise, response bias. If there is a signal, a hit and a miss are possible, and, if the signal is not present, a false alarm or a correct rejection are possible [Tab. 3.1]. Noise is therefore continuously mixed with information in the brain [Subsec. 3.2.2]. The most effective way to remove or attenuate noise is by averaging. Averaging is based on the concept of additive noise: Random (noisy) fluctuations above and below actual signals will gradually even out as one averages more and more percepts. Here, perception is considered as the extraction of forms or patterns from stimuli as if they were signals embedded in noise.

I wish to briefly stress an important point upon which I shall elaborate in due course: Sensation and perception are possible only because the perceived object and event are in an *environment* [Subsec. 1.2.8]. Any perceived item is an item because it is an item *against* a given background. However, for the same reason also *without* the given background, the properties of the item would not be contrasted with anything else, and therefore would never be segregated on the background noise. We see again the positive role of the environment whenever information is acquired. It should be noted that in the proceeding discussion the word *environment* is not always explicitly used to denote this fundamental role. Sometimes I also use the word *context* for referring to it.

[17][*GIBSON* 1979, pp. 53–5]. [18][CHANDRASHEKAR *et al.* 2006].
[19][LIEBERMAN *et al.* 2007]. See also [LUMPKIN/CATERINA 2007] for a review of the several mechanisms of sensory transduction in the skin.
[20][*LEWONTIN* 2000, p. 65]. [21][*ROMANES* 1884, p. 49]. [22][STOFFREGEN/BARDY 2001].

978–0–19–960848–5 03-Auletta-03-drv Auletta (Typeset by SPi) 73 of 862 June 1, 2011 17:18

OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

The smallest increase in stimulus strength that can be detected is proportional to the initial strength of the stimulus, i.e. the sensitivity is adjusted to the task at hand. The fact that information is coded at the start of sensation accounts for the fact that there seem to be no nonlinear transformations in sensation and in the first steps of perception [Subsecs. 1.2.2 and 2.2.2], leading to Stevens's law for the stimulus intensities,[23] which represents a further development of Weber's and Fechner's laws. It can be formulated as

$$S = kI^p + c, \tag{3.1}$$

where S is the numerical report of the sensation magnitudes evoked by the stimuli I, k is a constant, p is the slope parameter of the function plotted in double logarithm coordinates [Fig. 2.6], and c is the intercept. Our receptors respond to sudden changes and then they *adapt* (this is the basis of habituation) to constant stimuli. Therefore, only the change is important here.

Since three parameters are relevant in neural communication of sensory information (modality, location, and intensity of the stimulus), different codes are used[24]:

- The labeled-line code that signals with certainty the quality of the stimulus (the type of the stimulus, for instance whether it is painful or friendly), and
- The frequency code, i.e. by making usage of differences in frequency (it is good for the energy or intensity of the stimulation: On average we only make about seven such categories[25]).

The location of the stimulus is determined by the spatial distribution of sensory neurons. The labeled-line code is point-like, the frequency code wave-like. Obviously, as said already, the CNS too acquires information, this time from the sensory system. And therefore, relative to the latter, it shows all the above characters of information acquiring:

(i) The character (its survival value for the organism) of the peripheral stimulus (for instance, whether a visual shape or an odor) is guessed thanks to
(ii) Information sharing with the sensory system through specific sensory pathways leading to the CNS, and
(iii) Selection of some specific aspects of the sensory information that will be integrated in higher sensory areas of the brain.

Indeed, as mentioned, all messages received by our sense organs and transmitted to the CNS consist of streams of very similar impulses[26]: All information coming from the senses is subjected to a digital all-or-nothing law—a neuron receptor either fires or does not—and all signals consist of streams of very similar impulses that are further transmitted essentially through the same mechanisms.[27]

Finally, information coming from different sensory sources is integrated into the brain, and here true representations of external objects and events are produced. The information integration in the brain is optimized. For instance, the nervous system seems to combine visual and haptic information in a fashion that is similar to a maximum-likelihood integrator[28]: Visual dominance occurs when the variance associated with visual estimation is lower than that associated with haptic estimation. There are two ways in which information is finally represented in the brain: Transient spatial-temporal patterns of electrical impulses and strength of interneuron connections, which still changes

[23] [*MOUNTCASTLE* 1998, pp. 9–11]. [24] [GARDNER/MARTIN 2000]. [25] [MILLER 1956].
[26] For more on the subject of this section the interested reader can turn to a good synthesis like [*GLYNN* 1999].
[27] [MOUNTCASTLE 1978]. [28] [ERNST/BANKS 2002].

with time but has more stability. Therefore, in the brain there is a hardware, but no software can be found, apart from the transient patterns. This runs against the functionalist approach and deals with the tremendous problem of why the brain, though dealing with information, does not represent the world in a codified way; a problem with which we will deal in the second and third parts of the book.

To resume, any stimulus determines a twofold change in the nervous system[29]:

- One is the pure reaction (sensation), which gives rise to changes due to excitability,
- The other one is the result of an appropriate combination of stimuli and consists in the plastic changes of the nervous system (perception).

Therefore, the translation of the external signal as a whole actually consists of an autonomous *production* of the response to a given stimulus rather than of the pure "reproduction" of the external stimulus[30]: The organism is informationally shielded. This is evidence for the generalized Darwinism [Subsec. 2.2.6] I support in this book (there are no "Lamarckian instructions"). Let us consider two examples:

1. Mammals like cats show spontaneous (endogenous) cortical activity—i.e. independently from the stimulus—that develops patterns that are typical of orientation maps. This activity is not random and has a typical cycle, in which there is not only stimulation of the nerve activity but also periods of inactivity, a pattern that optimizes coordinate connections.[31] The same results can be found for rats.[32] Here, the activity of running has been considered and striking correlations in excitation patterns between running and REM sleep have been found. These results have been confirmed by similar studies on primates under anaesthesia.[33] In my opinion, all this displays evidence that the structures here associated with vision are spontaneously produced by the brain, then used and eventually reinforced, in the interaction with the external conditions, depending on the type of feedback received from the environment.[34]
2. Another example is the computation of probabilities of events, which can be shown to happen also in rhesus monkeys.[35] In this case, neurons are able to sum and integrate probabilities and likelihood according to the logarithm rule of product–sum. It is a typical stimulus-free, information-processing activity.

I also observe that the connections within the prefrontal cortex (both within and between layers) are far more numerous than the connections coming in from other areas, including sensory processing systems.[36] This means that there is heavy information-processing (as well as information-acquiring) activity in the CNS, especially when endogenous activity (without referential and therefore representational import) is developed.

3.3.2 Neurons

Let us consider the cell unity of the brain: The neuron.[37] The following pages should be understood as a quick reminder of some basic notions about the neuron that could become useful for the following. The informed reader may also skip this subsection. In the human brain, there are 10^{11} neurons and almost 10^{15} connections among them. Contrary to what was previously supposed,

[29][KANDEL 2000a, p. 34]. [30][*VON HELMHOLTZ* 1867, p. 586].
[31][KENET *et al.* 2003] [*WEST-EBERHARD* 2003, p. 111]. [32][LOUIE/WILSON 2001].
[33][PINSK/KASTNER 2007] [VINCENT *et al.* 2007]. [34][AULETTA 2002].
[35][YANG/SHADLEN 2007]. [36][*LEDOUX* 2002, p. 188].
[37][*GOLGI* 1995] [*CAJAL* 1899–1904]. See also [KANDEL 2000a] [*LEVITAN/KACZMAREK* 1991].

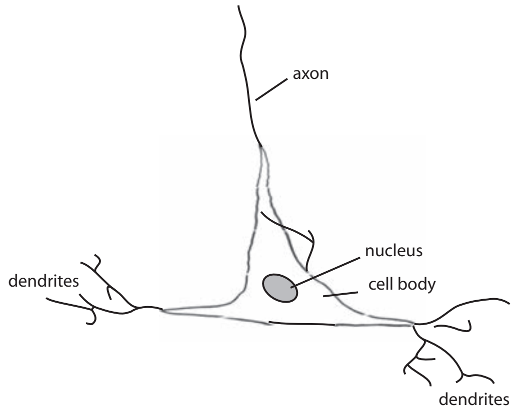

Fig. 3.3 Essential components of a neuron: The cell body with its nucleus, the axon, and the dendrites.

neurons are continuously replenished in some areas of the brain and this turnover may play a crucial role in learning and memory.[38]

We examine neurons according to their functionality, their anatomy, and their structure. All of the body's neurons are of three different *functional* types:

- Motor neurons for controlling motion, mostly located in the spinal cord and in the brain stem,
- Sensory neurons for receiving sensory inputs, mostly located in the peripheral nervous system, and
- Interneurons, mostly located in the cortex and mediating between the other two types of neurons.

Anatomically, a neuron is composed of *dendrites*, which receive several inputs from other neurons, a *cell body*, which integrates these inputs, and the *axon* (equipped with one or more terminals) for transmitting outputs to other neurons without attenuation over distance [Fig. 3.3]. Many neurons have a single axon. However in general, an axon branches many times so that the output will be transmitted to many neurons, as well as a single neuron being able to receive an input from many other ones. Many dendrites have little knobs that are called *spines*. Sometimes, axons can be directly connected with other axons as well as dendrites with other dendrites. However, in most cases the terminal of an axon is connected to a dendrite of another cell.

The signals are sent and received through a specialized structure that is called a *synapse* [Fig. 3.4]. The synapse consists in a *presynaptic membrane* at the end of the axon presynaptic terminal, a *postsynaptic membrane* (pertaining to another neuron), generally attached to a dendrite, and, between these two structures there is a small gap (200 armstrong) called a *synaptic cleft.* All the transmitter-gated channels are constituted by proteins. The information is transmitted through small signaling molecules known as *neurotransmitters* (chemical signals mostly of the family of amino acids) that are able to jump the gap of the synapse and be gathered by appropriate receptors in the postsynaptic membrane.[39] Neurotransmitters can also be biogenic amines (amino acids enzymatically modified), short proteins called neuropeptides (they are essentially modulators),

[38] [GROSS 2000]. [39] [DEUTCH/ROTH 1999].

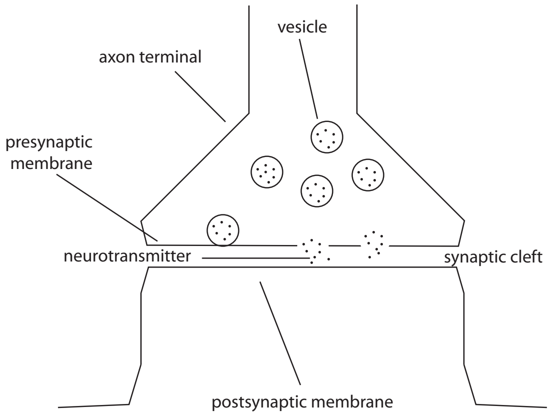

Fig. 3.4 Essential components of a synapse: Presynaptic and postsynaptic membranes, synaptic cleft, together with neurotransmitters.

or a miscellaneous group including the first neurotransmitter discovered, namely acetylcholine (ACh). Neurotransmitters induce a flux of ions.[40] Ion (Na, K, and Ca) channels are ion selective and fluctuate between open and closed states.[41]

Structurally, we mainly distinguish between pyramidal cells, spiny stellate cells, and smooth (or sparsely spinous) cells [Fig. 3.5]:

1. The *pyramidal cells* are always involved in excitatory (asymmetric) synapses. Excitatory cells release transmitters at their synaptic ends that, on contact with the postsynaptic membrane, create currents that depolarize the postsynaptic cell. The apical dendrite of pyramidal cells extends until the cortical surface, while the other dendrites (the basal dendrites) and the axon grow from the base in a downward direction toward the white matter.
2. The spiny stellate cells are mainly excitatory. They are concentrated in the middle layers of the cortex and are most abundant in the visual cortex.
3. The smooth stellate cells are inhibitory (symmetric), i.e. they release transmitters that hyperpolarize the postsynaptic cell, diminishing the effect of depolarizing currents.

The tissue formed by the neurons is called neuropil—it is probably the most complicated structure of our universe [Fig. 3.6]. The cortical tissue is composed of two types of cells, neurons and neuroglial cells. The latter does not take part in the interactions between neurons, although they may play a role in the slow modulation of neural function as well as for synaptic formation, maintenance, and efficacy.[42] Neuropil has emerged independently three different times in the history of our planet, in molluscs, crustaceans, and vertebrates.[43]

As we have seen, unlike other cells of the body, neurons directly communicate with one another. In the 19th century it was believed that this communication was mechanically produced through the propagation of an electric impulse. However, it was shown by von Helmholtz[44] that the actual speed of transmission is too slow, and this in turn suggested that each neuron somehow mediates

[40] An ion is basically an element which is negatively (anions) or positively (cations) charged, that is, that presents more or less electron relative to the referent chemical.
[41] [*ALBERTS et al.* 1983, pp. 651–92]. [42] [GALLO/CHITTAJALLU 2001].
[43] [*FREEMAN* 1995, pp. 38–39]. [44] [*VON HELMHOLTZ* 1883, pp. 663–79 and 881–85].

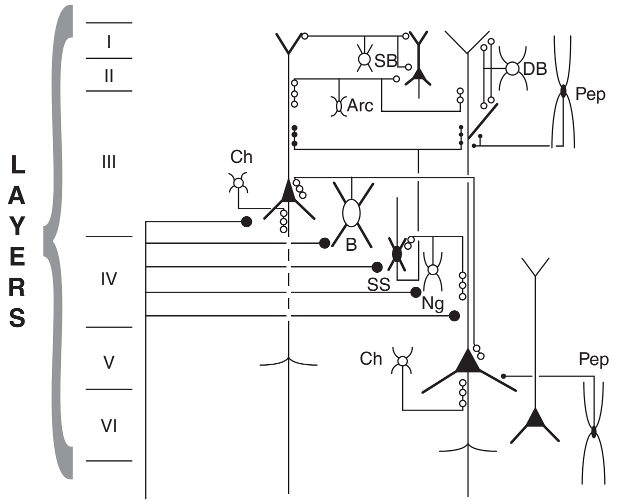

Fig. 3.5 Interconnections between neurons, including terminations from the thalamus. Solid circles are excitatory, open circles inhibitory connections, solid cells excitatory, open cells inhibitory. One can also distinguish the 6 traditional layers of neocortex. The pyramidal cells are easily recognizable. Other neurons are flagged by following denominations: Arc = arcade cells, B = large basket, Ch = chandelier, DB = double basket, Ng = neurogliaform, Pep = peptide cell, SS = spiny stellate, SB = small basket. Adapted from [*MOUNTCASTLE* 1998, p. 62].

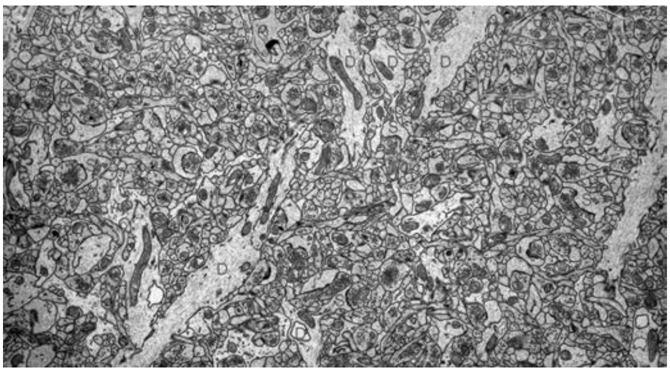

Fig. 3.6 Electron micrograph of neuropil, a tissue including axons (Ax) which are forming synaptic contacts (Sy) on dendritic shafts (D) or spines (S) and are intervened by glial processes of astrocytes (Ap). Also cell bodies and processes of oligodendroglia and microglia can be sporadically present as well as numerous blood capillaries. Adapted from http://synapses.mcg.edu/anatomy/neuropil/neuropil.stm.

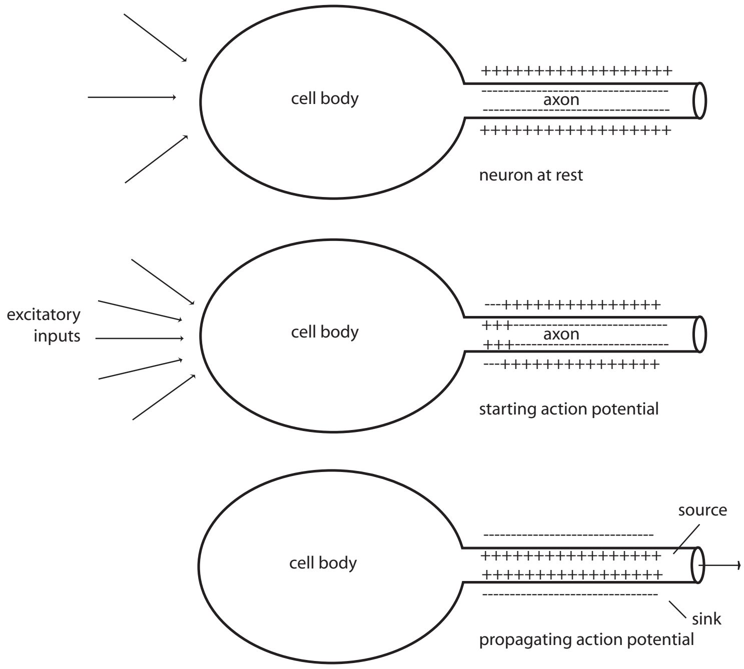

Fig. 3.7 A neuron is initially at rest (the inputs coming from other neurons' action potentials are not sufficient to activate the cell). After the activation threshold is passed, the cell's own action potential starts and is later transmitted.

and regulates the impulse it receives before further transmitting it.[45] When a neuron is at rest, the electric charge inside its axon is negative, while it is positive outside. The disparity between sink and source establishes an electric dipole. An electric dipole is a pair of opposite electric charges usually separated by a small distance. A change of electrical potential in the presynaptic cell triggers the synapse to release a neurotransmitter: When it receives sufficient excitatory input from other neurons, electric activity may begin through which the interior and exterior charges are inverted.[46] An *excitatory* stimulus (eventually inducing depolarization) is made by positively charged ions that flow into the cell (the source) and determine a negative charge in the contiguous extracellular space (the sink) [Fig. 3.7]. Therefore, an electrical stimulus that exceeds a certain threshold strength triggers an explosion of electrical activity that is rapidly propagated and amplified along the neuron's axon—the so-called *action potential* (AP). The grand postsynaptic potential (grand PSP) in a neuron represents a spatial and temporal summation of many small postsynaptic potentials. While the grand PSP is a continuous graded variable, action potentials are always all-or-nothing and uniform in size. Following the emission of a spike, a neuron needs time to recover: There is a period of 1–2 milliseconds in which the neuron cannot emit a second spike. This is called the absolute refractory period.

Therefore, the synapse can be considered a converter from the input represented by the spike train of frequency-modulated signals to the wave-like amplitude-modulated signals [Fig. 1.1]. Furthermore, the axon can be considered an inverse converter from the amplitude-modulated

[45] [*LEDOUX* 2002, pp. 43–45] [*CHANGEUX* 2002, pp. 14–18]. [46] [KOESTER/SIEGELBAUM 2000].

signal to the output represented by frequency-modulated signal, e .g. the intensity of the peripheral stimulus is transmitted in such a way.[47] Consequently, the signal goes through the axon of the presynaptic neuron as a frequency-modulated signal, eventually passes the synapse and is converted into an amplitude-modulated signal, then eventually, before leaving the postsynaptic soma, is again converted into a frequency-modulated signal and goes through the postsynaptic axon.

Therefore, neurons also show analogical aspects. Also, from another point of view, neurons can be considered as complex analogical devices rather than digital ones. Indeed, the fundamental output information of a neuron is not encoded merely in the form of individual AP signals, but also in their temporal sequence.[48] Moreover, when the neuron is "at rest," it displays a highly developed, wave-like activity. Indeed, as I have said, it receives a lot of (excitatory and inhibitory) inputs, although often they are not sufficient to overcome the threshold of activation. Neurons also "talk" in between two subsequent action potentials. Specific intracellular–calcium sensors regulate these interactions.[49] Summing up, from the point of view of information acquisition (of the neuron informational activation), the neuron can be conceived as a digital device, but from the point of view of the way it treats and processes this information together with other neurons, it is rather an analogical and wavelike device. As we shall see, this suggests that the brain as a whole is a pure representational device that is able to treat information but does not make a general use of a linear combinatorics, which is a necessary requirement for coding [Subsec. 2.2.2].

Excitatory effects that are passed to the postsynaptic cell are mainly transmitted through glutamate, while inhibitory effects mainly through gamma-aminobutyric acid (GABA) [Fig. 3.8].

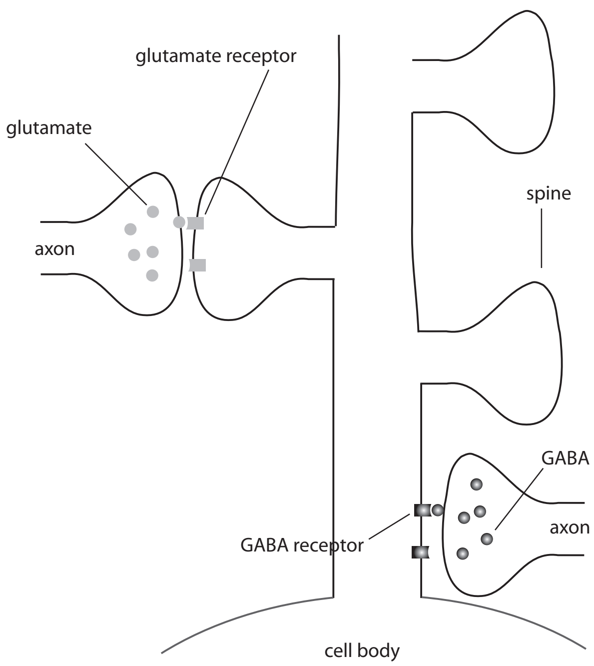

Fig. 3.8 Excitatory effects are transmitted to the postsynaptic cell mainly through glutamate, while inhibitory ones mainly through GABA.

[47][*PANKSEPP* 1998, p. 82]. [48][*ROEDERER* 2005, p. 136]. [49][HEIDELBERGER 2007].

Actually, when glutamate binds to the outside part of a postsynaptic cell, it allows a passage to be opened through which positively charged ions enter the cell, and this explains the change in charge when an action potential occurs (depolarization). The reverse is caused by inhibitory neurotransmitters (hyperpolarization). The effects of neurotransmitters are often not a property of the chemicals themselves as such but a function of the receptors to which they bind: The same neurotransmitter may be excitatory or inhibitory. This is further evidence that a message is selected only at the end of the process and not at the start [Subsec. 2.3.2]. Moreover, as we shall see, any reaction of a cell to an external stimulus can lead either to a sensitization or to a habituation. *Sensitization* operates as an alert signal warning that what will follow the stimulation is not trivial. *Habituation*, on the contrary, operates as if the cell has learned that there is nothing relevant occurring or no relevant consequences will follow.

There are not only excitatory or inhibitory actions but, as mentioned, some neurons are also modulatory (this again shows the importance of regulatory aspects in information processing). The function of the latter is not that of transmission but rather to cause longer-lasting changes in the way in which postsynaptic cells respond to excitation and inhibition. The transmitters that act at synapses of this kind are called neuromodulators and begin a cascade of enzymic reactions. These effects (as the action of glutamate on NMDA cells) are involved in memory and learning.

3.3.3 Brain Waves

It is useful to briefly consider the different brain waves, determined by populations of neurons and denoting different brain activities. It is a fact that inputs to different brain areas are due to local field potentials and are smooth.[50] Thanks to EEG, we are able to distinguish five different rhythms[51] that I shall order from the slowest to the quickest:

- *Delta rhythm* (0.5–3 Hz, that is, 1/2–3 cycles per second): Is characteristic of a sleeping subject, the so-called slow-wave sleep (SWS). We probably have no (or very low) energy expenditure during this phase [Subsec. 2.2.1].
- *Theta rhythm* (4–7 Hz): It is ~~also~~ typical of deep meditation, unconscious processing, negative emotional experiences such as frustration. It is also typical of REM (rapid eye movement) sleep (during dreaming). REM is an ancient mammal function. This reflects an active information integration. It is interesting to observe that the electrical activity of the brain stem during dreaming is the mirror image of waking (most neurons firing during waking are silent during REM). This is a very important point, as we shall see. REM can be seen as a waking state with behavior paralysis. During REM the amygdala is very active and the hippocampus exhibits highly synchronic activity (probably reflecting memory and learning consolidation). It is the same rhythm when animals explore their environment. This means that this activity is integrative, and it costs energy.[52]
- *Alpha rhythm* (8–12 Hz): It is a rhythm that is typical of rest. It reflects a complete synchronized activity of neurons and neuron populations, and is a sort of standby of the brain, that is, it expresses its autonomous information-processing activity.
- *Beta rhythm* (13–30 Hz): It is a faster rhythm and reflects the sudden desynchronization of the brain activity (the neurons no longer fire in phase) due to external stimuli impinging on our senses or in general to our interaction with the world. It is probably the "affective" rhythm.
- *Gamma rhythm* (more than 30 Hz): It probably reflects high cognitive processes, perhaps binding and decisional ones.

[50] [LOGOTHETIS *et al.* 2001]. [51] [*PANKSEPP* 1998, pp. 87–90, 125–42]. [52] [VINCENT *et al.* 2007].

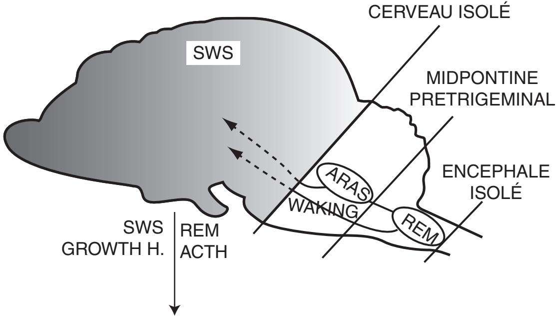

Fig. 3.9 Brain cuts (see text for explanation). Adapted from [*PANKSEPP* 1998, p. 132].

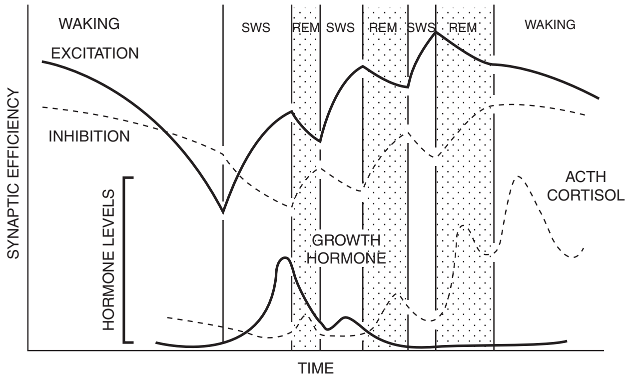

Fig. 3.10 Activation-state cycles of the brain and hormones. Adapted from [*PANKSEPP* 1998, p. 137].

The cycle delta–theta–delta rhythm (BRAC: basic rest–activity cycle) during sleeping is dependent on the metabolic rate of animals (20 minutes for cats). It is interesting to observe that if the higher brain is completely disconnected from the sensory inputs coming from the peripheral nervous system (the so-called *encephale isolé*), it maintains the normal BRAC [Fig. 3.9]. If, on the contrary, the cut is at the high midbrain level (*cerveau isolé*), animals become comatose and remain in their slow-wave (delta) sleeping activity. Only after many weeks do modest returns of desynchronization occur. If the cut is a bit postcipated (the midpontine pretrigeminal cut), there is a great deal of waking EEG activity, even if the sensory information is still interrupted by this cut. This suggests that waking is independent from sensory inputs. This is the ascending reticular activating system (ARAS).[53] Another interesting point is that REM is conserved between the *encephale isolé* and the midpontine cut. It seems that REM deprivation compromises a mammal's ability to learn complex emotional tasks that are not prefigured in the animal's evolutionary history, such as two-way avoidance, which requires back and forth movements between safe and danger zones of a test chamber.

[53][MORUZZI 1949]. See also [*PANKSEPP* 1998, p. 132].

Finally, it is interesting to observe that hormone secretion and neurochemical changes in the organism are a function of cycling brain states [Fig. 3.10].

3.4 A Short Look at the Brain's Structure

3.4.1 Structural Considerations

Phylogenetically and developmentally, the vertebrate brain can be divided into a *hindbrain* (cerebellum, pons, and medulla oblungata), which controls basic functions of life (if destroyed, then life ceases), a *midbrain*, which controls basic functions of awareness and behavior (as we have seen, if damaged the subject will enter into a comatose state), and a *forebrain* (diencephalon, cerebral hemispheres, and corpus callosum), which, in higher mammals, controls the high mental and behavioral processes (for instance, if damaged the subject can show impairment in problem-solving). In primates, the forebrain contains the neocortex. Different cortical lobes can be distinguished: An occipital lobe (the posterior part of the brain), a parietal lobe (lateral on the top), a temporal lobe (lateral below), and a frontal lobe (on the anterior part of the brain).

More specifically, the spinal cord is the place where the information coming from the sensory organs is gathered and eventually directed to the brain. The spinal cord is actually the interface between the CNS and body (it transmits inputs in both directions). Let me remind the reader here how the CNS is constituted: The informed reader may skip these pages. The CNS is made up of[54] [Figs. 3.11–3.12]:

(1) The cerebellum, which is essential for maintaining posture, for coordinating hand and eye movement, in fine-tuning the movements of muscles and learning skills.

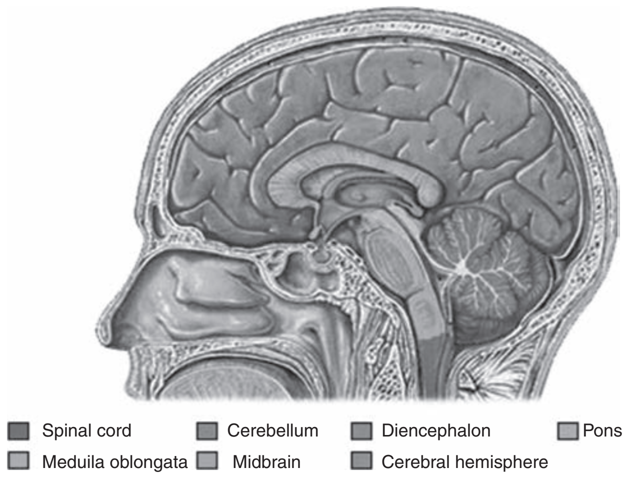

Fig. 3.11 Structure of the brain. Adapted from http://www.lifespan.org/adam/graphics/images/en/19236.jpg.

[54][AMARAL 2000].

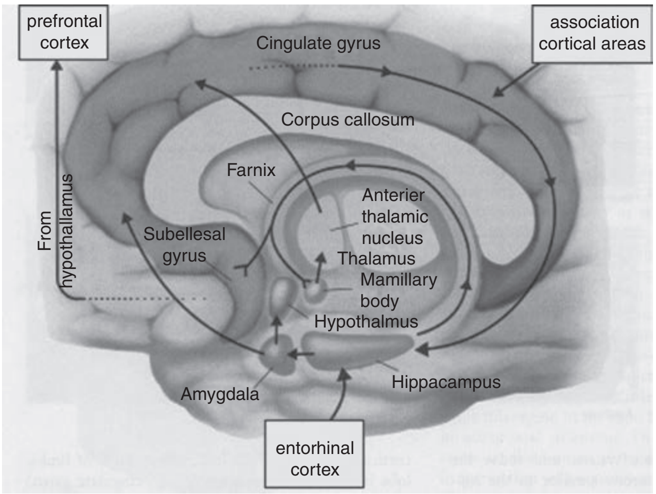

Fig. 3.12 Magnification of the limbic systems. Adapted from [*GAZZANIGA et al.* 1998, pp. 81–2], a very good textbook for this section.

(2) The brain stem, which is essentially the interface between the metabolic system and the brain, and covers the hindbrain (apart from the cerebellum), the midbrain, and a part of the forebrain (the diencephalon). The brain stem is then subdivided into
- Pons. The dorsal portion is involved in respiration, taste, and sleep. It also receives and passes on information about motion.
- Medulla, an extension of the spinal cord controlling blood pressure and respiration.
- Midbrain, essentially providing linkages between components of the motor systems, though dealing with some sensory information.
- Diencephalon, which consists in the thalamus and hypothalamus. Hypothalamus is essential for regulation of vital functions, like homeostasis and reproduction. The thalamus is the gateway to the cortex for sensory information and for motor information coming from the cerebellum and basal ganglia. Its function is probably much more dynamic than it was traditionally thought and probably influences the way in which information is dealt with in the cortex.[55]

(3) Cerebral hemispheres, connected by the corpus callosum, with which they constitute the cerebrum, consisting in
- White matter. It is functionally relevant for assuring quick and reliable connections between neurons.[56]
- Basal ganglia and striatum, which regulate movements and are fundamental for learning new skills.
- Amygdala, which is fundamental for associating emotional responses to world events.
- Hippocampus and hippocampal formation. The hippocampus is relevant for the storage of new memories.

[55] [SHERMAN 2006]. [56] [WEN/CHKLOVSKII 2005].

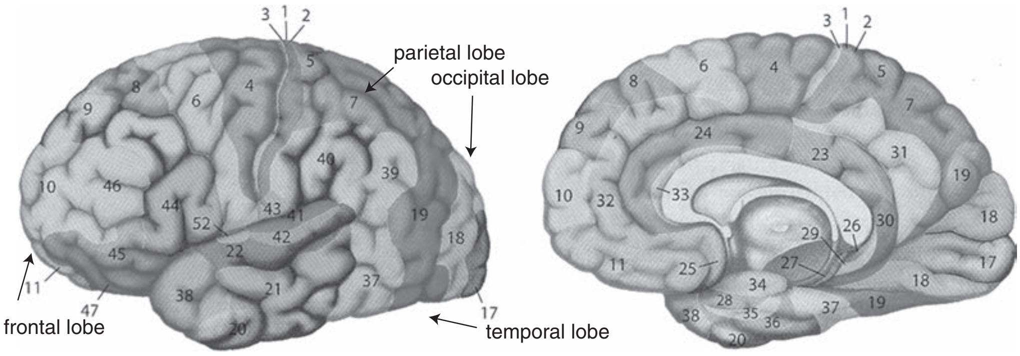

Fig. 3.13 The traditional 52 brain areas individuated by Brodmann.

- The cerebral cortex, the evolutionarily most recent part of the brain, as mentioned, has four major lobes. Two additional regions are the cingulate cortex, which surrounds the dorsal surface of the corpus callosum, and the insular cortex.

In the mammalian brain, the very deep folds of the cortex are called fissures and the shallower folds are called sulci.[57] The folded cortex between two adjacent sulci is called a gyrus. The cortex shows several layers (they are conventionally stipulated as six in number).

The Brodmann model of the human brain consists in a parcelization of 52 areas [Fig. 3.13]. Sometimes the parcelization of von Bonin is used. It is very important to consider that the human brain shows a spectacular number of different areas even relative to other primates.[58] Despite several efforts to modularize the brain, which we shall consider below, it must be said from the start that the common properties of all cortical areas are overwhelming. Neurons are organized into radial columns. Columns do not represent pieces of a mosaic, but rather a flexible, adaptive field of cooperative activity.[59] Interactions are mainly local, otherwise they pass through the underlying white matter. The cells must interact at high speeds in large numbers, without giving up their local autonomy. As we shall see, this is one of the most fundamental characters of self-organization. In other words, the microscopic, local activity of single neurons (it is spatially and temporally localized, and stimulus-locked) and the macroscopic, global activity of neural populations (spatial-temporal patterns, distributed over the entire sensory cortex, directed to the meaning of the stimulus for the organism) coexist. It is again a manifestation of the complementarity between point-like local activity and global structure [Subsec. 2.2.5].

3.4.2 Functional Considerations

In the cerebral cortex we functionally distinguish between the sensory areas, articulated into primary sensory areas, unimodal sensory areas, posterior association areas, and the motor areas, articulated into the anterior motor association area, premotor cortex, and primary motor cortex:

[57][*ABELES* 1991]. [58][*CHANGEUX* 2002, p. 30]. [59][*FREEMAN* 1995, pp. 48–53].

(a) Each primary sensory area conveys information to a secondary unimodal association area, which refines the information of a single sensory modality. Here the information is processed in parallel.
(b) These elaborated pieces of information are then sent to a multimodal sensory area that integrates information from several sources.
(c) From this area there are projections to the multimodal association motor area, and
(d) From here to the premotor (secondary motor) area.
(e) From the secondary motor area the information finally joins the primary motor area.

Therefore, while the primary sensory area receives the initial input, the primary motor area receives the final cerebral output. I stress that no prefrontal area receives direct input from primary sensory or primary motor cortices. As can be seen after this short examination, the brain, though showing global aspects, is not a single structure with a single function. It is also true that all pieces of cortex basically perform the same types of transactions on the information they receive [Subsec. 3.3.1], so that the differences among brains of different species mainly lie in the number of transactions that can be carried out concomitantly. One can distinguish[60] between levels of analysis of information (a higher-level question is largely independent of the levels below it) and levels of processing (although there are convergent pathways, the convergences are partial and occur in many places and times). Two features characterize this organization: (1) Feedback connections and (2) the fact that a large number of neurons is almost always involved in information-processing.

In this complex elaboration process, the single stimuli are smoothed into an integration process according to which it is the sensory area of the brain that distinguishes between signal (perceptual import) and noise in the stimulus.[61] The brain is engaged in constant, chaotic but stable background activity—we have already considered the activity of neurons between two successive action potentials [Subsec. 3.3.2]. The global origin of this activity is what enables a microscopic stimulus to send the entire system through a state transition into a unified jump from one pattern to another (often the brain is in a so-called metastable state). However, the action of this stimulus is not simultaneous but begins at one point and spreads radially in a wave-like form. A relevant point, as we shall see, is that this new pattern is only *triggered* by the stimulus but not selected automatically by it, as far as it is also determined by previous experience with this class of stimuli.

There are reasons to suppose that, between the activity of single neurons and that of the whole cortex, there are also neuron assemblies.[62] They can represent multiple dimensions of a single behavioral event.[63] Georgopoulos gave rise to an interpretation of an ensemble of neurons in terms of a population-vector theory[64] for explaining, say, the hand's movement: Each neuron is assumed to "vote" in its preferred direction—the direction in which the cell is maximally active—with a strength that depends on how much the activity of the neuron changes for the movement under consideration. Then a weighted distribution and a common direction are found. The direction of the hand movement was estimated to fall within a 95% confidence cone constructed around the direction of the population vector. This has also been analyzed in the case of rotations.[65] However, successive studies[66] have shown that this model cannot account for hand motion: The population

[60][*CHURCHLAND/SEJNOWSKI* 1992, pp. 18–37]. [61][*FREEMAN* 1995, pp. 58–67].
[62][*ROEDERER* 2005, pp. 137–8]. [63][DEADWYLER/HAMPSON 1995].
[64][GEORGOPOULOS *et al.* 1986].
[65][GEORGOPOULOS *et al.* 1989]. See also [GEORGOPOULOS *et al.* 1992]. [66][SCOTT *et al.* 2001].

vectors should be the sum of all cell vectors and should be congruent with the direction of hand movement, and instead it is not.

3.4.3 Neural Plasticity

Until the 1980s there was a controversy between brain scientists and behavioral scientists, because:

- According to the first group, the brain should already be fixed after childhood, while
- For the second group it was evident that humans and animals also show plastic behavior in a ripe age.

In the 1960s, the first neural evidence was already available that in animals like cats there could be some recovering of visual ability after the removal of parts like the entire occipito–temporal neocortex and this was judged to be critical for vision.[67] This recovery process is due to the already observed circumstance [Subsec. 3.3.1] that the brain treats any input information in the same way. It has also been suggested that sensory stimulation of animals previously deprived of suitable environmental conditions (characterized by some anatomical changes of the neural connections) will simply restore the permissive conditions of normal genetic and epigenetic development.[68] There is also evidence that shows that active zones and neurons in rats "overrun" zones and neurons which correspond to functions that are not active.[69] Moreover, the connection along the retinogeniculocortical pathway in cats shows a precision that goes beyond simple retinotopy to include many other response properties (receptive-field sign, timing, subregion strength, and size). This complexity in wiring suggests the presence of a developmental mechanism that can select among afferents that differ only slightly in their response properties.[70]

Relatively recently, it has been shown that all parts of the brain can be modified through experience. Michael Merzenich and coworkers proved that an alteration of sensorial input in monkeys through training in discrimination of a tone at a certain frequency causes a modification of the anatomical-physiological structure of the brain.[71] Elbert and coworkers[72] showed that learning can also modify the neural structure (for instance, the brain of a violinist is differently mapped in respect to a non-violinist). This is in perfect accordance with the result that memories are continuously generated and replaced across the whole life of a mammal[73] [Subsec. 3.3.2].

Other evidence of neuronal plasticity is that in neonatal primates visual acuity is six times and contrast sensitivity at least ten times worse than in adults. Even if at birth most neurons are already responsive to visual stimuli, the spatial structure of their receptive fields is extremely immature.[74] The contrast sensitivity of individual neurons in the LGN and the cortex also matures in a way that reproduces the improvement in behavioral performance (activity). Therefore, development of spatial vision depends on visual experience. Although plasticity decreases with age, the brain retains the ability to rewire itself throughout its whole life, so that all the maps of the brain shift in response to their activity.[75]

[67][SPRAGUE 1966]. [68][*FUSTER* 2003, p. 40].
[69]For a summary of the effects of sensory deprivation on animals and humans see [*WEXLER* 2006, pp. 39–83].
[70][ALONSO *et al.* 2001]. See also [KATZ/SHATZ 1996]. [71][RECANZONE *et al.* 1993].
[72][ELBERT *et al.* 1995]. [73][GROSS 2000]. [74][*BLAKEMORE* 1990b].
[75]An impressive example is represented by Bach y Rita's experiments with teaching blind patients to see with their tongue [*HAWKINS/BLAKESLEE* 2004, p. 61].

Finally, I stress that transplanted cells in the brain usually transform to the type required for the new location, but the older the organism is, the more cells become committed to their identities.[76] This shows that it is the context that determines the function of a neuron.

3.4.4 Selectionism versus Constructionism

Another important scientific controversy has developed between:

- A *selectionist* party, which supported the idea that, with experience, no new connections are created in the brain, but some are only selected and reinforced from a pool of initially very weak connections. According to selectionists,[77] experience reinforces certain connections and eliminates those which are not used (as we shall see this has a central importance for memory).
- An *instructional* party, according to which new connections are created.[78]

In a study supportive of the latter point of view, M. Stryker and A. Antonini[79] showed that during development there is a certain increase in complexity in the axons. As a matter of fact, these structural changes are not completely new but are added to preexisting connections, so that it is probably a phenomenon of quantitative growth rather than a creation of new structures. It may be interesting here to mention that this growth occurs under the stimulation of synaptic signals even if in a direction that is opposite to the direction of the stimulus.[80] Further evidence favorable to the selectionist point of view comes from studies[81] showing that in the first six months after birth all senses are connected in children, and only thereafter are some connections interrupted.

Selectionism and constructionism could also be evaluated in relation to three parameters[82]: (a) number of synapses; (b) axonal arborization, and (c) dendritic arborization. (a) has been found to be very significant; (b) and (c) show selection but also refinement and specification of the structures. This could be the consequence of the fact that axons and dendrites do not cover a specific location but rather a *whole* area. It is indeed important to consider that information in the brain is mapped with a high sensitivity according to similarity and connectivity of the stimuli. It is a form of contextual sensitivity and perhaps a general representational principle. The order (permutation) of afferent connections onto an excitable dendritic arbor is indeed crucial for determining the cell's response to different patterns of synaptic input. It is this spatial ordering that allows for pattern discrimination. A true representation must preserve some "spatial" ordering of the represented object, but what features are relevant here is dependent on the selective way we deal with our environment. This shows that selection is not in contradiction to growth and construction, but only that one should consider a larger region of cerebral "space" that will become more and more determined during development. As I shall show in the next part of the book, growth processes are fully compatible with selection processes (for instance, during epigeny).

3.4.5 Brain Size

Here, I shall make some general and final considerations about the significance of a large brain. The problem of brain size has traditionally been overestimated. Rats have a brain that is relatively bigger than ours and elephants have brains that are absolutely bigger.[83] The apparent increase

[76][*GAZZANIGA/ HEATHERTON* 2003, pp. 117–19].
[77][CHANGEUX/DANCHIN 1976] [*EDELMAN* 1987]. See also [*WEISMANN* 1889, I, pp. 85–8].
[78][*LEDOUX* 2002, pp. 70–96]. [79][ANTONINI/STRYKER 1993]. [80]See also [*CHANGEUX* 2002, p. 197].
[81][*HARRISON* 2001]. [82][QUARTZ/SEJNOWSKI 1997]. [83][*DEACON* 1997, pp. 160–3].

of the brain in primates is in reality a decrease in somatization. However, a distinction must be made here. Dwarf animals (like the chihuahua) exhibit slowed body growth only in the late fetal and postnatal phases, whereas primates start out with small bodies and their growth is regular. Humans are more similar to dwarves (there is a truncation in the growth curve). This does not mean that we are dwarves, because our body mass has actually increased in the course of our evolution. The fact is that the growth of our body and the growth of our brain seem to be *distinct*, and the growth of the brain, in particular, is prolonged. Three points seem relevant here:

- By increasing size, there will be some fragmentation of functions, according to the rules of system theory [Subsec. 2.4.4].
- Larger brains will be less able to rely on genetic mechanisms to determine structural differences and more prone to proportional perturbation.
- The brain's plasticity also means that the brain adapts to the body: There is no preestablished harmony between brain and body.

3.5 The Stages of Information-Processing

In humans, there are four neural and psychological stages in the response to a sensory stimulus, as shown by EEG[84] [see Fig. 3.14]:

- Response to the physical qualities of a stimulus. This happens in sensory areas and the two waveforms are called P1 and N1 (where P and N indicate the polarity—whether positive or negative—of the waveform). It is the stage in which *sensation* occurs.
- The waveforms of the second stage (P2 and N2) occur in the temporal cortex and may reflect the brain's detection of the general category of an event. The detection of a deviant event enhances the amplitudes of the N2 and is called *mismatch negativity* (MMN or N2a): It is independent from the level of attentiveness. It rather represents an initial *orienting* of the organism and the first stages of perception.

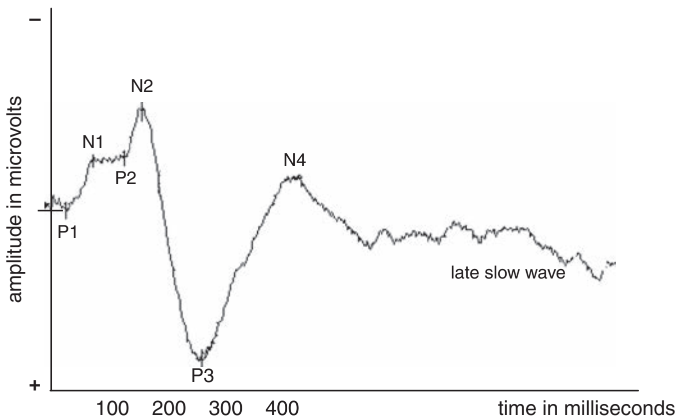

Fig. 3.14 The four steps of information-processing.

[84] [*KAGAN* 2002, pp. 96–108].

- The next waveforms are called P3 and N4, and may reflect the reaction to a particular event and its initial evaluation. There is a big individual variation in this stage. A wave called P3a reflects in particular the initial discrimination of an unexpected or unfamiliar event. Another wave called P3b deals with its implications. This is the phase of the *attention* to the *external event*, and also represents a highly developed perceptual processing.
- The late slow wave reflects the evaluation of the event and its proper semantic categorizing. This is the cognitive work of concept formation, which is typically human and involves higher symbolic and interpretative functions; an issue that will occupy much of the investigation later.

The whole process is ruled by information-selection mechanisms. Selective listening experiments have taught us much about the way humans select and pay attention to sensorial data. We shall come back to the problem of attention in the next part of the book. For now, let me simply point out some fundamental aspects related to the perception and treatment of information. These experiments have led to three main conclusions[85]:

(1) A central, as opposed to a mere sensory, factor is involved.
(2) The effects vary with the number of possible messages that might have arrived and from which pool the actual ones have been chosen (the rate at which the information is arriving at is important).
(3) The capacity of the brain will limit the number of tasks that can be performed simultaneously; therefore a part of the presented information must necessarily be discarded. However, information is not discarded (selected) at random.

Even if we shall recognize some of the limitations of this model in due course, these conclusions are nevertheless for the most part correct. The relevant fact for information theory is that the reaction time to a stimulus varies with the information conveyed by that stimulus. In other words, the larger the ensemble from which the stimulus is drawn, the longer the reaction time. Instead, if the ensemble is kept constant in size, the reaction time increases as the different possible stimuli are made more similar, that is, when their position is very near in an abstract information space in which the different stimuli are codified: It is easy to see that in Fig. 3.2 the unit 128/255/0 is less distinguishable from the unit 255/255/0 (pure yellow) than the unit 0/255/255 (pure green) is, and that the unit 192/255/0 is still less distinguishable from pure yellow. Note that both any new or intense stimulus and high-pitched noise are very likely to be perceived (the so-called governing selection).

3.6 Modularity of the Brain

After the first developments in the 1940s in cybernetics thanks to the work of Wiener and Ashby and the early stages of computationalism thanks to the efforts of von Neumann, Simon, and others, the cognitivist school was born in the 1950s [Sec. 3.2]. It was characterized by several theses:

- *Modularity* of different information-processing brain areas, against the idea of a central processor.
- *Nativism*, according to which most brain activities follow prewired structures and rules that are inborn.

[85][*BROADBENT* 1958].

- *Strict separation* between syntax and logical rules on the one hand, and between syntax and semantic contents on the other: Syntax was discovered as an intermediate level between pure physics and semantics, and this seemed to provide a solid basis for cognitive sciences. This assumption was rather common also among the computationalists.

I limit the scope of the present examination to the first thesis, leaving the other two assumptions for a much later discussion. Modules have been defined as domain specific, mandatory, generally not conscious (especially in their lowest levels of computation), fast, partially informationally shielded systems against both background and external feedback information, having swallow outputs, associated with a characteristic neural architecture, exhibiting specific and characteristic breakdown patterns, and finally as having a characteristic sequencing in information-processing.[86] All modules operate in the same way and essentially protect from disturbances that may arise from other areas and computational tasks.

As we have seen, the differentiation and localization of specific activities is, according to system theory [Subsec. 2.4.4], a general feature of any system that is sufficiently large and complex. Modularity is an aspect of the general hierarchization of biological systems (as we shall see in the second part of the book), and it may be considered as due to the tendency to autarchy of parts of the system as well as to a parallel integration of the subcomponents of each subsystem (increased information sharing among those subcomponents).

The assumption that brain activity, which is complex, is modular was quite common in the early days of cognitivism. Later on, some criticism of this assumption was developed. It can indeed be shown that the same frontal regions are recruited for a broad range of cognitive demands.[87] Moreover, Farah[88] has pointed out that there is no evidence supporting the assertion that damage to one module does not affect other modules (the locality assumption), which deprives modularity of its appeal. Finally, the big individual variability in the size and exact location of areas throws uncertainty on modularization.[89] Again, this is in accordance with the examination of Subsec. 2.4.4.

Perhaps a useful compromise could be the following:

(1) Modules (at least in mature brains) are mostly of a *functional* type, rather than of a structural type (in the next chapter we shall see some evidence to support this). This means that the modularity of the brain may well apply not to global systems (for language, vision, and so on) but to functions (playing, exploring) and specific tasks (a modular circuit for riding a bicycle, for driving a car), i.e. either on a smaller scale or more distributed than according to Chomsky and Fodor's initial proposal. In those cases, different specific, say, visual and motor competencies are integrated in a whole that, as time progresses and an increasing canalization of the activity takes place, can even function almost independently from any other activity.

(2) There are also modules that are anatomically distinct and well localized, but this much more for *initial and mid-step* information-processing. In particular, it concerns primary sensory and premotor areas.[90]

(3) Modularity is a matter of *degrees.*[91] If we assume that such a degree is also temporally *variable* and depends on the task at hand, then different modules may plastically cooperate in specific tasks. This approach is also to a certain extent consonant with Simon's idea of the near decomposability of systems[92] which means that interactions between elements pertaining to different subsystems or quasimodules are weaker than interactions among elements of the

[86][*FODOR* 1983]. [87][DUNCAN *et al.* 2000]. [88][FARAH 1994]. [89][*MOUNTCASTLE* 1998, p. 83].
[90][*MOUNTCASTLE* 1998, p. 87]. [91][SCHLOSSER/WAGNER 2004b]. [92][*SIMON* 1969, pp. 197–200].

same subsystem, but are not totally absent. Let me provide evidence coming from studies on attention, which proves that auditory tasks strongly interfere with auditory tasks as well as visual tasks with visual tasks, but also that auditory tasks tend to interfere with visual tasks and vice versa.[93] These results show that a certain modularity is present but also that pools of attention are not completely separated.

As we shall see, this compromise seems to be the way in which many neurologists consider modularity today.

3.7 Mapping

The way the brain produces maps is a very interesting subject. Let me first say that there is no mapping of one neuron to one external item, so that the so-called grandmother neuron (a single neuron that should be associated with your grandmother) does not exist.[94] In general terms, the problem is the following: As proved by Edelman and coworkers,[95] there are not sufficient neurons available for producing the astronomic number of representations any one of us can have. In fact, as I have stressed, neurons associate in functional groups in order to perform computational tasks[96] [Subsec. 3.4.2].

3.7.1 Maps

There are two types of maps in central neurons[97]:

(1) Maps that reproduce (wholly or partially) the spatial relationships present in the peripheral sensory epithelium or retina and these are called *topographic* or projection maps. The functional role of these maps is difficult to establish because the coding of spatial organization might not be the factor determining their topographic organization. Nevertheless, as we shall see, they play an important role in the first steps of processing visual information. Moreover, topographic maps of the surface of the body in the brain are formed during development by projections that become ordered, in part, through competitive interactions.

(2) The other type of map is called *centrally synthesized*, because they are the result of higher cognitive integration activities. Here, primary sensory cells neither register nor extract the location, delay, or orientation of a stimulus. The selectivity for these cues is created by neuronal circuits in which the neurons forming the map are sort of nodal points. These types of maps play an important role in representational processes.

Here, we shall deal only with topographic maps. We can take it as a fact that the primate's brain has a topographic map of the body[98] [Figs. 3.15 and 3.16]. It is also true that there is no one-to-one correspondence between, say, fingers and cortical-motor neural cells, so that individual movements are shaped from more rudimentary synergies, such as those used to open and close the whole hand.[99] Moreover, neural territories controlling different fingers overlap.[100] Single primary motor (M1) neurons are active with movements of different fingers. The control of any finger movement appears to utilize a population of neurons distributed throughout the M1 hand area.

[93][DRIVER/SPENCE 1994, MACALUSO *et al.* 2002]. See [*WILLINGHAM* 2001, pp. 73–75].
[94][*CHANGEUX* 2002, p. 51]. [95][TONONI *et al.* 1992]. [96][VAADIA *et al.* 1995].
[97][KONISHI 1986]. [98][*SHERRINGTON* 1906, *SHERRINGTON* 1942] [GENTILUCCI *et al.* 1988].
[99][SCHIEBER 1990]. [100][SCHIEBER/HIBBARD 1993].

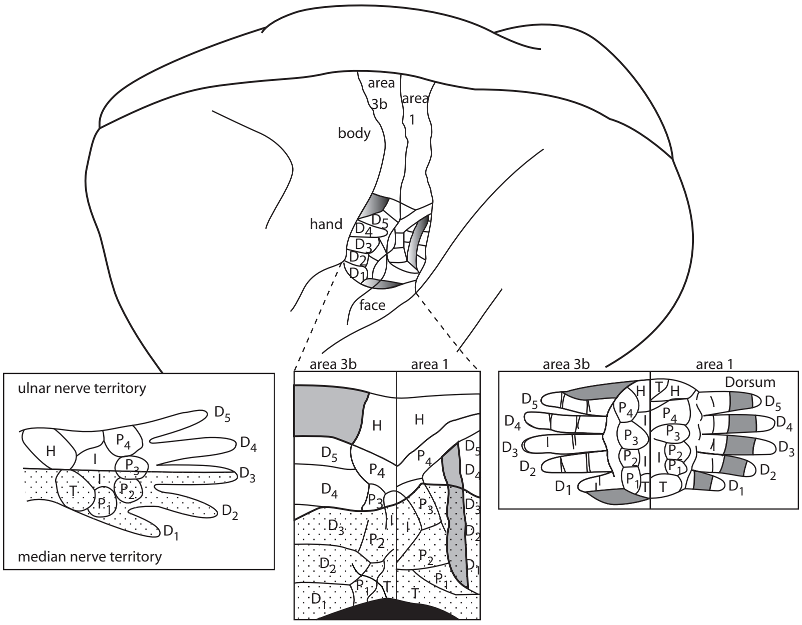

Fig. 3.15 Schematic drawing of the representation of the hand in the brain of owl monkeys: it is a map for motor purposes. Adapted from [*CHURCHLAND/SEJNOWSKI* 1992, p. 33].

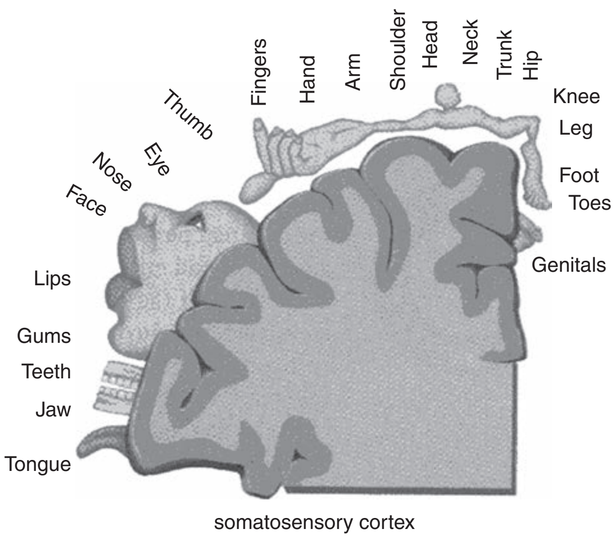

Fig. 3.16 How the human brain maps the body. Adapted from [*GAZZANIGA/ HEATHERTON* 2003, p. 111].

The topographic (or retinotopic) maps should also be understood in dynamic terms.[101] Indeed, there is a certain evidence that the reorganizations of sensory and motor maps occur in all major systems at subcortical as well as primary cortical levels[102] [Subsec. 3.4.3]. For instance, the frequency representation in the cortex of adult animals is able to reorganize in response to partial deafness.[103]

3.7.2 Images?

Let us now discuss an important topic. The fact that the brain produces or shows maps almost naturally implies the question of whether or not images in the mind are analogues of the perceived objects. Three types of imagery have been distinguished[104]:

(1) Topographic representations of spatial relations (located in the posterior parietal lobes, angular gyrus; they make use of motor processes),
(2) Figural imagery (located in the inferior temporal lobes: This type of image occurs when a low-resolution topographic image is generated from the activation of stored representations of shapes and other object properties),
(3) Depictive imagery (allows one to reorganize, reinterpret, or compare shapes).

Only the first type of image is relevant for the question above. Indeed, the existence of maps of the second or third type above, which are both kinds of centrally synthesized maps, does not prove that there is some analogy between representation and represented item, because ultimately these maps only concern some further elaboration of the spatial relationships between different stimulus sources. Let us consider the case of the lateral geniculate nucleus, which contains a topographic map of the retina and therefore of the external space. Now, if the neural map when perceiving a leopard reproduces the spot structure, this spot structure is in itself only a spatial relationship between "dark" and "bright" stimuli and reveals in no way that the object is a leopard, which is a construction that comes much later in our information-processing system (as we shall see in the next chapter). Moreover, topographic maps, like any spatial mapping, are subjected to typical distortions and selection processes. Therefore, it would be an error to mix the representations of spatial relations on the one hand, and the figural and depictive imagery on the other.

At an evolutionary level, there has clearly been a process of adaptation to our environment. Now, adaptation has determined a fit between the objects present in our environment and our brain representations. This does not imply, however, any direct relation between object and representation [Subsec. 2.2.3]. In other words, the argument assuming that maps are pure analogical reproductions of the external world is rendered circular. As a matter of fact, more elementary species do not perceive objects as we do.

On this important issue there was a debate between Pylyshyn and Kosslyn[105].

- According to Kosslyn, we produce images as analogues of things: Mental images are formed step by step with analogical but partial correspondence to external structures and the parts are only later arranged and integrated into the proper configuration.[106] Also recently, Kosslyn[107] pointed out that visual mental imagery and visual perception rely on common mechanisms. A certain

[101] [*GAZZANIGA et al.* 1998, pp. 110–11]. [102] [KAAS 2000]. [103] [KING/MOORE 1991].
[104] [*KOSSLYN* 1994].
[105] See [*BODEN* 1988, pp. 27–44] [ANDERSON 1978] for reconstruction and analysis. See also [FREEMAN 2000b].
[106] [*KOSSLYN* 1980, KOSSLYN 1988] [KOSSLYN *et al.* 1990]. See also [*KOSSLYN* 1994].
[107] [KOSSLYN/THOMPSON 2000].

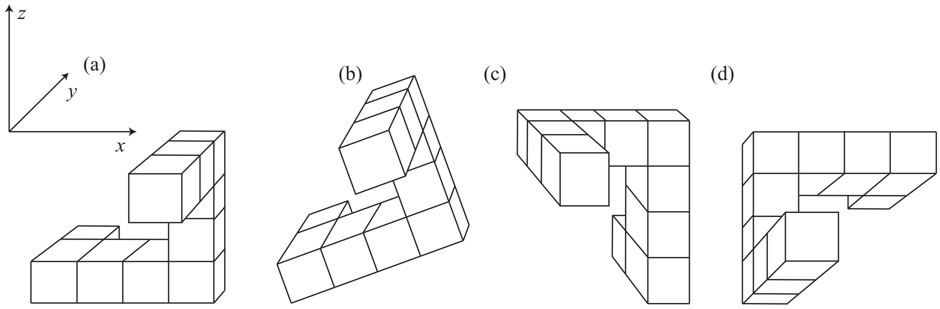

Fig. 3.17 Consider the object shown in (a). When it is rotated at an angle of 20 degrees around the y axis, as in (b), it is easily acknowledgeable as the same object. However, the task becomes increasingly difficult when the angle is at 90 degrees, as in (c). This shows that we somehow effect a mental rotation of a representational image of the object. However, when rotated at 180 degrees, as in (d), the task can become easier, because this configuration is apperceived as the result of a double reflection relative to the initial position (a), due to the cyclicality of rotation.

evidence supporting this view was provided by Shepard and Metzler.[108] They tested the capacity to judge whether solids, which were actually rotated relative to samples, were the same or not relative to the latter one, and found that the amount of time needed by human participants was proportional to the magnitude of the angle of rotation [Fig. 3.17].[109]

- Pylyshyn pointed out that images cannot be considered as internal pictures of things [see also Subsec. 3.3.1] because this would require a further interpreter, and so on ad infinitum.[110] In general, Pylyshyn showed that an isomorphism between an object and its representation is not enough: It does not suffice to specify a representation's form; one should also specify the *way* in which it can be used (a procedural or operational issue), due to the fact that the *same* form or neural pattern can represent very different things. As we shall see in the second part of the book, this is a central point to any theory of representation. Pylyshyn further showed that images are cognitively penetrable since they can be mapped to descriptions, whereas, if images are analogues of objects this could not be the case. However, from this fact he incorrectly inferred that imagery must be explained in terms of language-like representation. Nevertheless, Pylyshyn proved that subjects may mentally jump from one given location to another independently from the distance, giving evidence against a pictorial, figurative interpretation of representation and for a propositional interpretation of it[111], at least for humans. This shows that to a certain extent Pylyshyn was the representative of a traditional, classical computational approach to the brain [Sec. 3.2], while Kosslyn was somehow in the distributed-processing camp, as will be explained below.

Summing up, even if images cannot be understood as pure reproductions of external objects, they still cannot be considered irrelevant. If we abandon the strict mirroring theory, figural and depictive imagery, that is, images that are built before or after perception, may be relevant.[112] It has been shown that the most modality-specific cortical areas activated by stimuli in visual perception are also activated by higher-order brain regions in imagery.[113] Images share common features

[108] [SHEPARD/METZLER 1971]. [109] [*BADDELEY* 1990, pp. 71–4]. [110] [PYLYSHYN 2002].
[111] [*PYLYSHYN* 1984, pp. 243–4]. [112] [*JEANNEROD* 2009, pp. 103–14].
[113] [KREIMAN *et al.* 2000][FARAH 2000b].

with visual percepts. Mental images interact with other visual and verbal task components as if they were ordinary visual representations. At least some modality-specific cortical representations perform "double duty," supporting both imagery and perception, and these representations are functioning in an analogous way. Indeed, in a very interesting imaging experiment performed by Kosslyn and coworkers[114] it was shown that there is a meaningful overlap between cerebral areas activated during perception and by visual imagery tasks. This shows the insufficiency of a pure computational point of view, like that supported by Pylyshyn, and the important role of the iconic aspect of representation not only in perception itself but also in the so-called iconic memory, as we shall consider in the next part of the book.

During production of figural and depictive imagery, the direction of information flow is reversed, and some of the intermediate, spatially mapped, representations are reactivated by (top-down) higher-level mechanisms rather than by stimuli. An important difference between perception and imagery lies in the automaticity of the processes involved in the first. One cannot perceive a familiar object without simultaneously recognizing it as that familiar object (I cannot perceive a picture of my mother without being simultaneously aware that it is a picture of *my* mother), but one can think about familiar objects without calling to mind a visual mental image. This suggests that the activation of spatially mapped visual-cortical regions from memory requires the intervention of a separate, attention-demanding process, needed for image generation but not for visual perception.

Several studies have confirmed these results from another point of view[115]: Before executing a task, metabolic rates have already increased. This effect should be due to the central decisional system and in particular to an action simulation (where the motor output is inhibited). The function might be that of preparing the organism for the concrete task much better than a mere simulation of kinematic parameters would allow. Studies of Decety *et al.* [116] show that the brain areas activated during motor imagery are the frontal and parietal lobes. The most important finding of these studies is that to execute and to imagine an action are functionally equivalent (a concept that we shall develop below), even if the network activated during perceptual imagery is different from the one activated during motor imagery.

3.8 Neural Communication and Networks

3.8.1 Associationism and Connectionism

The idea that the brain is a network rather than a central processor [Sec. 3.2] has a long prehistory. I would like firstly to distinguish between three approaches: Associationism, connectionism, and parallel distributed processing (PDP).

- *Associationism* is very old (its roots can be traced back to the English empiricism of Locke and especially Hume) and it amounts to the idea that the results of cognitive activities like categorizing and conceptualizing are developed by empirical association between sensory stimuli.[117] A branch of this stream is represented by behaviorism,[118] with which we shall deal in the second part of the book. In fact, it seems that associations can be formed following a pure law of habit.[119] Association by simple contiguity of the stimuli in memory storage is here the

[114][KOSSLYN *et al.* 1997]. See also [*WILLINGHAM* 2001, pp. 286–7].
[115][*JEANNEROD* 2006, pp. 25–41]. [116][DECETY *et al.* 1989, DECETY *et al.* 1991].
[117][*LOCKE* 1689] [*HUME* 1739–40, *HUME* 1777]. [118][*WATSON* 1925] [*SKINNER* 1938].
[119][*SPENCER* 1860–62, pp. 211–12] [*JAMES* 1890, v. I, pp. 550–71].

most basilar aspect. Associationism is also the theoretical basis of connectionism and parallel distributed processing.

- *Connectionism* is an actualization of associationism and can be considered a derivation from Thorndike and Hebb's contributions[120]: It consists in the idea that the brain and the mind work as a network of units instead of being an analogue of a central-processor constructed computer [Sec. 3.2]. In particular, when two elementary brain subsystems have been active together or in immediate succession, they tend to recur together.
- Not all connectionist models are distributed[121]: They may use representations similar to those of the symbolic approach that are called *local* representations. *Distributed* representations, instead, account for the fact that there is content addressability: For instance, any part of a past occurrence or scene may lead to its later retrieval from memory.

Relative to cognitivism, connectionism represented an important novelty [Sec. 3.6] in five respects:

(1) It is opposed to the nativist positions first supported by Chomsky, advancing the idea that most or all of the brain's activities have an empirical, associationist source.
(2) It is opposed to the idea that in order to explain cognitive activities one needs to assume that the brain applies explicit principles and rules (like Chomsky's universal grammar, which shall be examined in the third part of the book).
(3) Any connectionist information-processing is context-sensitive, and for this reason it does not admit a strict separation, again supported by Chomsky, between syntax and semantics.
(4) Connectionist networks can be mapped in (reduced to) a dynamical physical system.
(5) Additionally, for the reason given in (2), supporters of distributed models, in particular, refuted the strict observance of information-codification models typical to any form of computationalism.

Exploration of the last three points must be postponed to the subsequent parts of the book, since they deals with the issue of language (Point (3)), the issue of complexity (Point (4)), and the nature of representations (Point (5)). Therefore, in the following subsection I shall focus mainly on Points (1)–(2).

3.8.2 Hebb's Rule

Hebb's assumption[122] asserts that when an axon of neuron A is near enough to excite neuron B and repeatedly takes part in firing it, some sort of growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased[123] [Subsec. 3.4.4]. This also means that there is a time window during which such development is reversible (the connection is not sufficiently strengthened).[124] This is a further form of neural selection. Another mechanism is when two (or more) afferent neurons of the same order participate in the firing of

[120][*THORNDIKE* 1931] [*HEBB* 1949]. [121]On this point see [*EYSENCK/KEANE* 2000, pp. 272–7].

[122][*HEBB* 1949, p. 62]. See also [*ROMANES* 1884, p. 35].

[123]Mountcastle has expressed his disagreement that this idea can be attributed to Hebb [*MOUNTCASTLE* 1998, p. 234], pointing out that it was a shared belief at that time—rather he let this chain of discoveries begin with T. Lomo's work at the beginning of the 1970s [*MOUNTCASTLE* 1998, pp. 141–3]. I would like to stress that it is quite common that a scientific idea is "in the air," so that the role of the famous discoverer is rather that of a catalyst. This, however, does not negate the fact that such a catalyst becomes the reference point for the future generations. This obviously does not diminish Lomo's role either.

[124][*HEBB* 1949, p. 229].

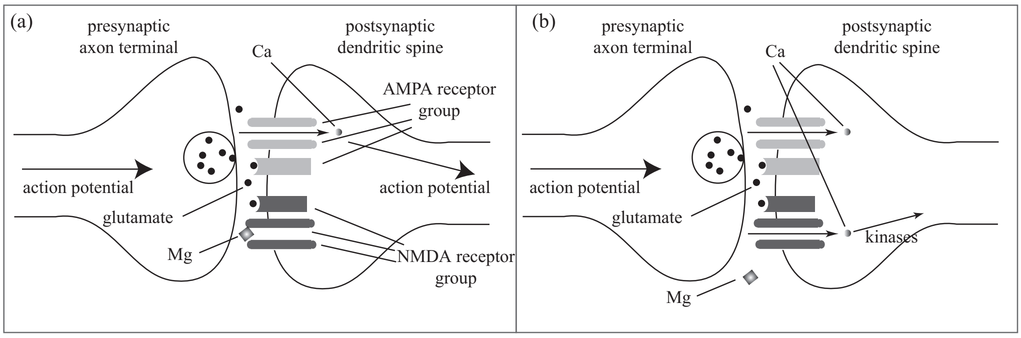

Fig. 3.18 Hebb's reinforcement mechanism. Glutamate (the excitatory neurotransmitter) binds to both NMDA (N-methyl-D-aspartic acid: an amino acid) and AMPA (Alpha–amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors. (a) However, it has no effect on the former because the receptor channel is blocked by magnesium (Mg) that does not allow the calcium molecules (Ca) to enter the cell. (b) When the postsynaptic cell fires an action potential, the magnesium blockade is removed, allowing the activation of kinase that travels up to the cell's nucleus, where several molecular processes occur, including gene activation, which leads to the synthesis of a new protein which contributes in strengthening the synaptic connection.

a third neuron so that they become finally interdependent.[125] This is very important for binding through *synchrony*.

Let us consider this mechanism a little, as it is known today[126] [Fig. 3.18]. Postsynaptic cells have two glutamate receptors [Subsec. 3.3.2], the AMPA receptor (AMPA actually mimics the effects of glutamate), involved in regular synaptic transmission, and a NMDA receptor, involved in synaptic plasticity (modulation). Presynaptically released glutamate goes to both receptors. Binding of glutamate to AMPA receptors is a major way to induce postsynaptic action potential. On the contrary, when glutamate reaches the NMDA receptor, it initially has no effect on the postsynaptic cell because part of the receptor is blocked. However, when the postsynaptic cell fires due to the reception of glutamate at the AMPA receptors, the block on the NMDA receptors is removed, allowing calcium to enter the cell, producing a long-term potentiation (LTP) as a result. Then, in order that NMDA receptors pass calcium molecules, both presynaptic and postsynaptic cells must be active (NMDA receptors are therefore coincidence detectors). Then, NMDA receptors allow the cell to record exactly which presynaptic inputs were active when the postsynaptic cell was firing and, in this way, when coincidence firing is repeated, a connection is strengthened. Less persistent and more persistent LTP is called early and late LTP, respectively. Several parallels have been found between early LTP and short-term memory, as well as between late LTP and long-term memory. While early LTP only activates preexistent proteins, through enzymes called protein kinases, late LTP involves the formation of new proteins.

[125] [*HEBB* 1949, pp. 70–1].
[126] [KANDEL/SIEGELBAUM 2000] [SIEGELBAUM *et al.* 2000] [*LEDOUX* 2002, pp. 144–51].

3.8.3 Neural Networks

On these general grounds, let us now introduce what is called the neural networks formalism.[127] A *neural network* is a computing system simulating the brain and consisting in a set of processing units, a state of activation, a pattern of connectivity among units, a propagation rule, an activation rule, an output function, and eventually a learning rule.

We must account for the fact that inputs into a given neuron are smoothly graded but the response is in a binary code: Either the neuron fires or not [Subsec. 3.3.2]. Recall that, when the traveling signal arrives at the endings of the axon, it causes the secretion of neurotransmitters into the synaptic cleft. The postsynaptic potential (PSP) diffuses in a graded manner (unlike the spikes in the axon) toward the soma where the inputs from all the presynaptic neurons connected to the postsynaptic neuron are summed. Information-processing in the network will therefore depend on two sets of variables only: The distribution of spikes among neurons and the list of synaptic efficacities (i.e. the connectivity between the neurons or the computational elements). The logical structure of a single neuron consists in: The processing unit (the soma), where the i-th unit is symbolized by σ_i, and a number of input lines going to the soma. To each input line a parameter w_{ij} is associated, where the subscript j refers to the input channel j (connected with the unity σ_j). The numerical value of w_{ij} is the synaptic efficacy which determines the amount of PSP that would be added to the i-th soma if the channel j were activated: The value 1 or 0, depending on whether or not it is active, can be assigned to each input box σ_j [Subsec. 3.3.1]). The grand PSP h_i for each neuron σ_i is the linear sum of the different components, i.e.

$$h_i = \sum_{j=1}^{N} w_{ij}\sigma_j, \tag{3.2}$$

where N is the number of presynaptic neurons σ_j. The unit is eventually activated and passes a signal to its neighbors. Recall that, in the case of real neurons, this signal is determined by both the level of activation of the sender unit and the nature of the connection involved, which can be either excitatory or inhibitory. However, this level of complication is unnecessary here.

Consider two input channels. If the output is a spike only when both channels spike, we can represent the logical AND; if only one is required, we have the inclusive OR. McCulloch and Pitts[128] introduced a time variable, so that the operation of this machine can be expressed as the logical truth function considering a successive time $t+1$

$$\sigma_i'(t+1) = \psi(h_i > T_i), \tag{3.3}$$

where T_i is the activation threshold and the function ψ can be 1 or 0 [Fig. 3.19].

Here, learning can be understood, in a first approximation, as a process in which the network adjusts its synaptic efficacies dynamically in order to accommodate to a certain pattern.[129] In particular, it would be suitable to introduce a Hebbian rule of learning at this level[130] [Subsec. 3.8.2]: The change of the weight

$$\Delta w_{kj} = \eta a_j \cdot a_k \tag{3.4}$$

[127] [*AMIT* 1989, pp. 1–96] [*ELLIS/HUMPHREYS* 1999, pp. 1–44], which are in general good introductions to all this subject matter. For early mathematical models of the way neurons integrate different dendritic inputs see [RALL 1962].

[128] [MCCULLOCH/PITTS 1943].

[129] See also [*AMIT* 1989, pp. 97–214].

[130] See [*ELLIS/HUMPHREYS* 1999, pp. 17–25].

978–0–19–960848–5 03-Auletta-03-drv Auletta (Typeset by SPi) 99 of 862 June 1, 2011 17:18
OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

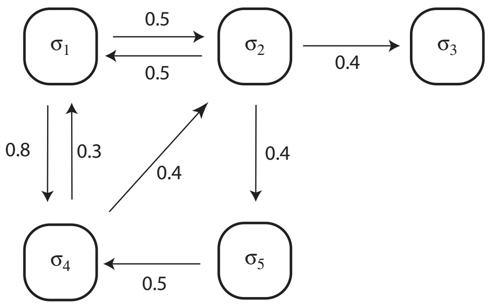

Fig. 3.19 Example of a generic and very elementary neural net. The following values have been assigned to the connections (the threshold is 0.5): $w_{41} = 0.8$, $w_{21} = w_{12} = w_{45} = 0.5$, $w_{24} = w_{52} = w_{32} = 0.4$, $w_{14} = 0.3$. All T_j are $= 0.8$. It is easy to verify that: σ_1 is active, since $w_{12} + w_{14} = 0.8$; σ_2 is active, since $w_{21} + w_{24} = 0.9$; σ_3 is not active, since $w_{32} = 0.4$; σ_4 is active, since $w_{41} = 0.8$. In fact, σ_4 receives nothing from σ_5, because the latter is not active, since $w_{52} = 0.4$. After one or more cycles, the network will eventually reduce to the connections between σ_1, σ_2, and σ_4 or be inactive. Suppose, for instance, that initially only σ_1 is active. It also makes σ_4 active; σ_1 and σ_4 together also activate σ_2.

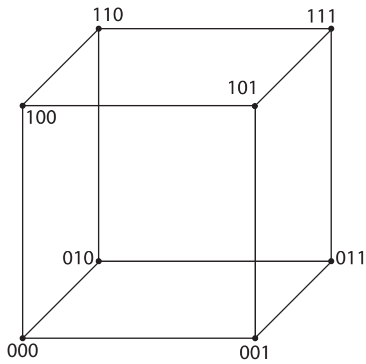

Fig. 3.20 Hamming binary distance in a three-dimensional space. The distance from 000 to 001 is 1, from 000 to 101 is 2 (000-001 + 001-101 or 000-100 + 100-101), from 000 to 111 is 3 (for instance 000-001 + 001-101 + 101-111: Actually there are 6 different paths connecting 000 and 111). See also Fig. 3.2.

of the connections is increased if and only if both units j and k are excited, where a_j, a_k are the activation values of the two units, respectively, and η is a learning parameter. As we have seen [Subsec. 3.3.2], neurons are essentially integrators of input information. Consequently, the quintessence of the function of the nervous system[131] is the ability to weigh the consequences of different types of information (and then to decide the appropriate responses). One can also measure the distance between different states of the network's parts, each coded by an N-bit word. A good measure is the Hamming distance, the number of bits at which two strings of informational elements differ. The space of all possible states can be represented as a hyper-cube [Fig. 3.20].

[131][*SHERRINGTON* 1906] [KANDEL 2000a, p. 29].

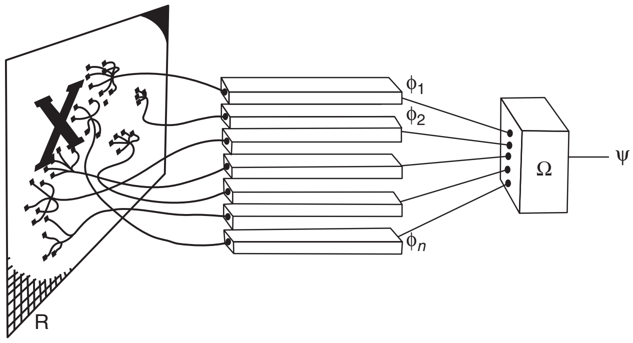

Fig. 3.21 Rosenblatt's perceptron. Adapted from [*MINSKY/PAPERT* 1969].

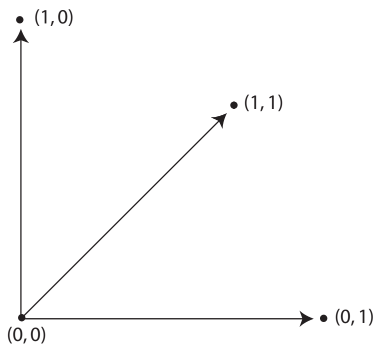

Fig. 3.22 Vectorial representation of linearly separable problems.

Rosenblatt[132] introduced the concept of a *perceptron* [see Fig. 3.21]. The idea is that the soma can receive inputs from much more than a few channels. In this manner the network can reproduce some simple predicates of the perceived world. The output (which is true or false) represents a classification of all possible inputs into two alternative classes. Indeed, it has been proved that perceptrons lead to a stable classification scheme. The problem is that perceptrons only deal with problems that are linearly separable, according to a pure classical computational point of view, and therefore are not very useful when dealing with the typical representational problems that are the object of the second part of the book. Consider the vectors shown in Fig. 3.22. It is impossible to divide the vectorial plane with a single straight line such that the points (0, 0) and (1, 1) are in one region and the points (1, 0) and (0, 1) are in the other. Instead, points (1, 0) and (1, 1) are linearly separable from (0, 0) and (0, 1). For this reason, perceptrons cannot perform the exclusive disjunction XOR or compute functions like commutativity or parity.

A further possibility is to avoid a single output and to have associated pattern pairs (or n-tuples) in a matricial way, as shown in Fig. 3.23. Therefore, by making use of the vectorial formalism [Subsec. 1.2.3], we have

[132] [*ROSENBLATT* 1962].

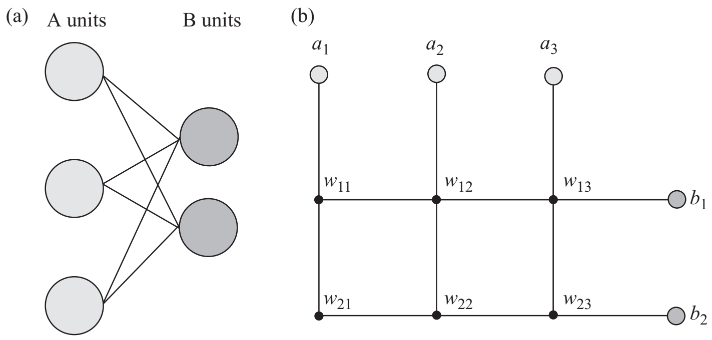

Fig. 3.23 A neural net with several outputs. (a) The connections between units A and units B. (b) A matricial representation of the same network showing the weights.

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \tag{3.5}$$

where

$$b_1 = w_{11}a_1 + w_{12}a_2 + w_{13}a_3, \tag{3.6a}$$

$$b_2 = w_{21}a_1 + w_{22}a_2 + w_{23}a_3. \tag{3.6b}$$

or, in more compact notation [see also Eq. (1.19)],

$$|b\rangle = \hat{W}\,|a\rangle\,. \tag{3.7}$$

Another improvement is when the multi-perceptron is closed into itself, forming a feedback mechanism.[133] This is called an Attractor Neural Network (ANN). Instead of Eq. (3.3) we write here

$$\sigma_i'(t+1) = \psi(h_i(t+1) - T_i), \tag{3.8}$$

where the value of ψ is either 1 or 0 depending on whether its argument is positive or negative, respectively, and

$$h_i(t+1) = \sum_{j=1}^{N} w_{ji}\sigma_j(t). \tag{3.9}$$

Once a determined configuration of firing neurons repeats itself indefinitely, we have an *attractor* (a stable configuration toward which the system tends spontaneously). A very simple example of a physical attractor is provided by a damped pendulum [Fig. 3.24]. For an ANN, the following assumptions have been done from a neurobiological perspective: The individual neurons have no memory, the network has full connectivity, the connectivity is symmetric, the dynamics of

[133] [HOPFIELD 1982].

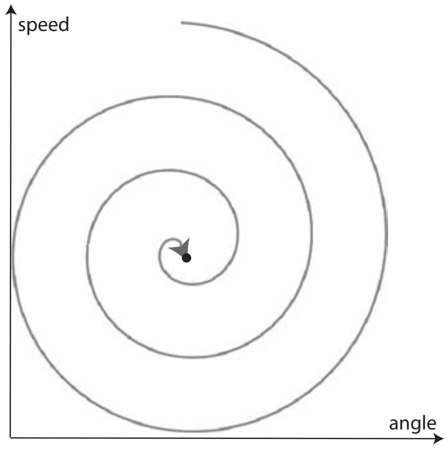

Fig. 3.24 The attractor for a damped pendulum: Starting with a big oscillation (high speed and large angle) it ends with a spiral trajectory in the dot representing the rest position (in general with the arm of the pendulum in a position parallel to the pull direction of the gravitational force).

the network are asynchronous. A network state is given by the list of the simultaneous axonal activities.

3.9 Concluding Remarks

Let me summarize the main results of this chapter:

- The brain shows both global wave-like behavior and local spikes.
- Also single neurons show a complementarity between continuity–discontinuity.
- We have considered that selectionism (the idea that specific brain connections are selected and not created *ex novo*) probably represents the right approach towards brain development, especially when considering larger areas of neuropil instead of single connections. We have also remarked that the brain shows a considerable plasticity even in the mature age.
- The brain shows a certain modularity (especially in the way different systems and subsystems are hierarchically nested) that cannot, however, be stressed too much at the expense of global behaviors. In particular, we have remarked the relevance of functional and specific-task micro-modularity.
- There is no isomorphism between the brain and external objects, even if, in a more narrow sense, we can speak of both topographic and central mapping.
- We have examined the associative Hebb's rule: When an axon of a neuron repeatedly takes part in firing another neuron, some growth process or metabolic change takes place in one or both cells such that they can fire together with higher probability. This is the theoretical basis of the connectionist theory of neural networks, which tries to reproduce the multipolar way in which the brain works.

We have also considered several schools, like classical computation theory (a central-processor model of the brain, symbolic mediation, the brain as an information processor and codifier); cybernetics (the relevance of irreversibility, feedback, and information control, block of environ-

978–0–19–960848–5 03-Auletta-03-drv Auletta (Typeset by SPi) 103 of 862 June 1, 2011 17:18
OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

mental noise and purposive behavior); the cognitivist school (modularity, nativism, and strict separation between syntax and semantics); connectionism (antinativist point of view, explicit rules as unecessary, relevance of context as well as the importance of having dynamical models; no centrality of having information codification).

The peripheral sensory system codifies information and both the sensory system and the CNS display information-processing activity. However, we shall show that the brain, in its representational function, does not codify information and essentially acts according to the rules pointed out by connectionism. We shall also see that the brain is a complex system in which a correctly understood modularity plays a central role. Finally, we shall discover that the fundamental function of the brain is to control environmental information and to display an analogical representational activity.

4 Vision

In the previous chapter we considered the brain as an information-acquisition device. In this and the next chapter I shall examine two specific and very important aspects of the brain's information acquisition: Vision and how the brain deals with motion. After having redefined the difference between sensation and perception, I shall discuss these two stages of vision in detail. Starting from the first steps of perception, we shall examine both how an object is perceived and the specificity of face recognition. Finally, some of vision's impairments are considered.

4.1 More on Sensation and Perception

Up to this point, we have examined some general aspects and functions of the brain. Let us now consider the specific ways sensory modalities process information. All perceptions can be classified as (1) somatosensory perceptions, including smell, taste, touch [Fig. 4.1], (2) hearing, or (3) seeing. The first modality is local, while the latter two are distal. Let us consider, in particular, the visual system, though a few words will also be said about the other sensory modalities. This is not by chance, since vision is probably the most studied sensory modality and surely the most important one in primates. This will confirm in part what has already been said about information-processing but will also open new perspectives that are worth further investigation.

It is important to underscore the distinction between sensation and perception:

- *Sensation*, as I explained above [Subsec. 3.3.1], is the pure process of information acquisition from the external environment and therefore consists mainly of stimulus transduction. Let us consider how visual transduction works. Since the eye is able to catch single photons, it is evident that quantum phenomena must *directly* play a role at the beginning of visual sensation. There are also other examples of the fact that biological systems, in dealing with light, are subjected to specific quantum-mechanical effects [Ch. 1], for instance in photosynthesis.[1] The nanoscale dimension of the photosynthetic complex is indeed critical for light harvesting. Chromophores in light-harvesting systems are densely packed, and the distance between different molecules is smaller than their overall size. At this scale the fine differences between the state vectors of both donor and acceptor ground and excited states are crucial. Both are strongly influenced by the environment. There has also been a discussion about the possibility that enzymes use quantum-mechanical effects, but the issue is still under scrutiny.[2]

[1] [FLEMING/SCHOLES 2004] [JANG *et al.* 2004] [SENSION 2007] [ENGEL *et al.* 2007].
[2] [KOHEN *et al.* 1999].

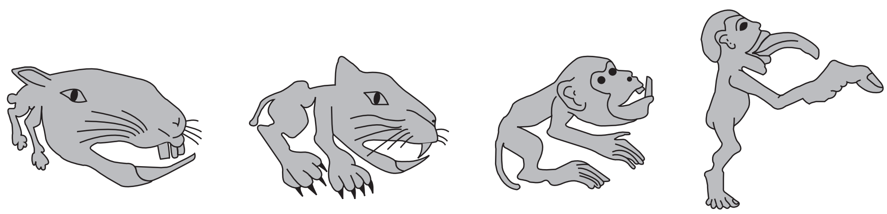

Fig. 4.1 The somatosensory cortex in some mammals. Adapted from [KANDEL 2000b].

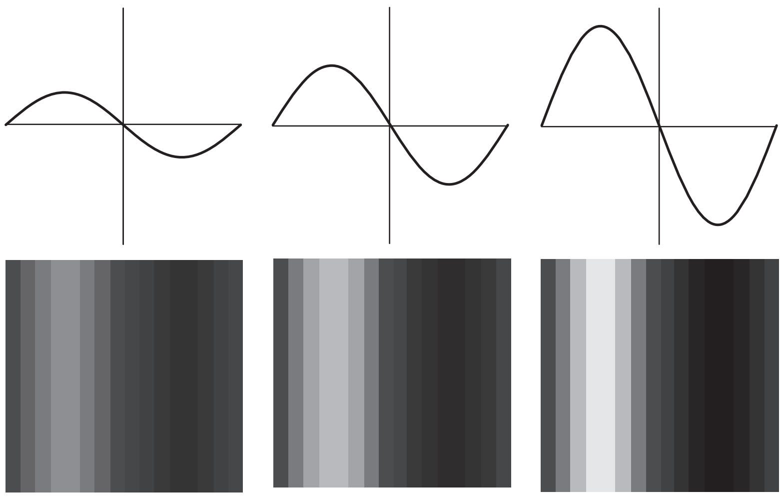

Fig. 4.2 Luminance as a function of the amplitude of the wave. The wave on the right is a superposition of the two waves on its left (the three blocks below represent a geometrical projection of the waves above). If the waves are relatively out of phase, then they can display more complex combinations with several peaks and valleys. See also Fig. 1.3.

As a matter of fact, as we shall see, visual sensation can be seen as dealing with two complementary aspects [Subsecs. 1.2.6 and 1.3.3], though it is not clear how the relevant effects can arise in this context. The two complementary ways to specify a pattern are either through measuring the luminance, i.e. the intensity of the light stimulus at many locations in which waves of different frequencies superpose [Fig. 4.2], or to discriminate between different values of the spatial frequencies.

- *Perception*, at the very least in birds and mammals, is instead accompanied by an expectancy.[3] Moreover, perception is characterized by the awareness of further facts associated with the object of perception—in humans, in the fourth state of stimulus processing [Sec. 3.5]. As an organized reconstruction, perception tends to build consistent systems. As we shall see, any perception in a human goes together with at least one implicit inferential process, in which the general

[3]As already understood by James [*JAMES* 1890, v. I, pp. 251–3; v. II, pp. 1–3 and 72–84].

characters of an object or event are reconstructed. Any perception of the relations constituting an object is due to such an inferential process. This does not mean that this cluster of relations is purely fictional.

Following W. James, I assume that there are two main principles of perception (as well as of higher cognitive processes):[4]

Any perception is maintained and becomes a habit until it is contradicted by some stimulus

and therefore[5]

All that can be merged is merged and nothing is separated except for what must be.

We can further distinguish between the *segregation* of the object from the background and the *discrimination* between several (at least two) objects. Subsequently, the second of the above principles states that the level of discrimination is always imposed by some practical need, and percepts remain as indistinct as they were, so long as no constraints force them to be otherwise. The pressures that force to introduced further levels of discrimination can be of phylogenetic or ontogenetic kind, but finally what does matter here is the ability of the organism to maintain an adaptive level of control on the environmental information.

At a more general level (involving not only vision), the difference between segregation and discrimination can be expressed as a duality between individuation and identification, and these two aspects are reminiscent of the localization of the system on the one hand, and the ability to obtain some structured percept on the other:

- *Individuation* means roughly to *pick up* a "black" dot in a uniform bright (noisy) background or *vice versa* [Subsec. 3.3.1]. Again, we see the importance of the environment for individuation. We shall also see that the environment is important for identification as well [Subsec. 1.2.8].
- *Identification* means to ascertain *what* this dot can represent or what survival value it has to the organism, for instance whether noxious or not. Any identification is actually a *recognition*, because, in order to identify an object, we must find out a perceptual schema among those that have been stored in the past able to fit the percept.

Often these two aspects have been confused. They are, however, completely different.[6] In many circumstances organisms only need to individuate a target and not to discriminate it. Take an easy example: A killer who must shoot a given person in a crowd is given only the information useful for individuating the target, for instance: The person on the left corner of the square wearing a red coat (actually, a coat is not a property of the person as such, even if it is somehow connected with her). Now, any other mark would suffice for the killer if it were suitable for target individuation. This shows that it is not the perceptual content of the mark which matters but only *its use for tracking* the target. These considerations can be easily extended to many predation contexts for most animals. Exactly the opposite is true when we need to identify something: We can do it only by stressing the different perceptual features that can specify the object. Therefore, perception, when identification is involved, always implies a general aspect: The involved properties.[7] Also in signal detection theory we find a similar distinction between detection and recognition,[8] which can be reduced to the above duality individuation/identification. Detection or individuation is

[4][*JAMES* 1890, v. II, pp. 288–9]. [5][*JAMES* 1890, v. I, p. 483]. [6][*HARRÉ* 1986, pp. 97–107]. [7][*MEAD* 1934]. [8][SWETS 1998].

sometimes strictly connected with a decision procedure, that is, with the procedure necessary to give rise to some successive operation relative to the target.

We can resume the previous discussion on the most general grounds by saying that there are two aspects involved here: Proximity (individuation, locality) and similarity (recognition, wholeness). This duality ultimately brings us back to the quantum-mechanical complementarity between locality and globality [Subsecs. 1.2.6 and 2.2.5].

4.2 Visual Sensation

Vision evolved providing organisms with distal control of their environments and relative actions. The protein mechanism involved in vision may also be found in a kind of mobile algae. Euglena, a single-cell organism that uses light as a source of energy, alters its pattern of swimming as a function of the ambient light levels in different parts of the pond or puddle in which it lives.[9] Subsequently, light has been used more and more by animals as an informational source about the environment in order to undertake the correct reactions and actions. With the further emergence of cognitive systems and complex social behavior, a good deal of motor output has become quite arbitrary with respect to the visual sensory input.

The eye was essentially a photodetector in the first stages of evolution. The most important step was the detection of light direction (thus segregation comes first). This was accomplished by an eye with the shape of a cup where light is projected on the back so that, by asymmetric perception (by activation of different neurons), the organism has a hint about the direction of light.[10] Vertebrates and insects have then developed independent systems for vision. The cells constituting the eye do not vary with the size of the animal so that eyes can never be very small or very big. However, there are anatomic constraints: In the case of insects, the pupil would be so small that diffraction would render a clear image impossible, and for this reason they have developed compound eyes. The single units are called *ommatidia*, each of which functions as a separate visual receptor. It consists of a lens (the front surface which makes up a single facet), a transparent crystalline cone of light-sensitive visual cells arranged in a radial pattern, and pigment cells which separate one ommatidium from its neighbors. The pigment cells ensure that only light entering the ommatidium parallel (or almost so) to its long axis reaches the visual cells and triggers nerve impulses. Thus each ommatidium is pointed at just a single area in space and contributes information about only one small area in the view field. In the following pages I shall discuss the mechanism of vertebrate vision. In the case of vertebrates, the smallness of the entering hole of the eye makes the detection of shapes possible by using their reflection properties, thereby allowing true discrimination.

Before explaining the route of visual information, it is important firstly to make some short historical remark on color theory. T. Young and H. von Helmholtz established the trichromatic theory of color[11]: The three fundamental colors are blue, green, and red. In 1878 E. Hering developed the opponent-process theory: One type of process produces the perception of green versus red, a second type of perception blue versus yellow, and a third type of perception white versus black. There is a recent synthesis of the two theories[12]: If the strength of the excitatory signal is greater than that of the inhibitory one, blue is seen; however if the opposite is the case,

[9] [GOODALE 2000]. [10] [*LLINÁS* 2001, pp. 100–1]

[11] [*VON HELMHOLTZ* 1883, pp. 3–23] [*VON HELMHOLTZ* 1867, pp. 275–384]. See also [*EYSENCK/KEANE* 2000, pp. 38–43].

[12] [*ATKINSON et al.* 1993].

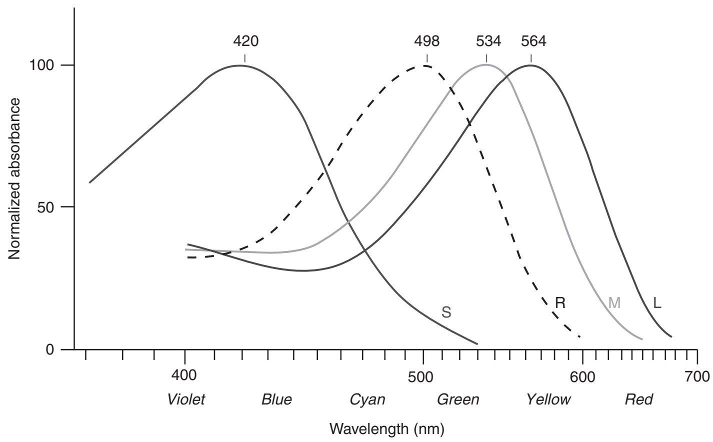

Fig. 4.3 Absorption spectra of the rods (R) and three types of cones: S stands for short waves (mainly blue), M for middle waves (mainly green), and L for long waves (mainly red). Adapted from http://en.wikipedia.org/wiki/Trichromacy. See also Fig. 3.2.

then yellow is seen. Color constancy is the tendency to see an object as being of the same color even if the color of light and all circumstances change. In general we decide on the color of a surface by comparing its capacity to reflect different waves as against adjacent surfaces. The presence of diverse pigmentation in our natural world creates an environment in which the spectral energy distribution reaching the eye of a viewer varies from location to location in space. Color vision is the capacity to extract information about the differences in these distributions, irrespective of their absolute energies. The perceptual result is a multihued world in which objects appear to merge and to contrast by virtue of their differences in color.[13]

Vision is carried out by multiple specialized systems that operate in parallel. In the retina there is a trade-off between[14]

- *Spatial resolution*: Restriction to those points that are stimulated and, as a consequence, poor sensitivity.
- *Sensitivity to light*: In order to detect lower levels of light the output of many receptors is pooled, implying loss of local information concerning which points were stimulated and which were not.

We already see here a complementarity between localization of light sources and what, at this level, can perhaps be called global patterns [Sec. 4.1]. The visual system partitions the image in two ways: The one favors local resolution (through cone cells), the other sensitivity (through rod cells):[15]

1. Cones have three different photopigments which absorb different wavelengths of light [Fig. 4.3]. The three fundamental colors define a color space that is the proper space of visual information-processing [Fig. 4.4]. Therefore, cones require ample (diurnal) light.

[13][*JACOBS* 1981]. See also [HILBERT 1992a].

[14]For the whole matter of this section I sharply recommend the textbooks of Farah [*FARAH* 2000a] and of Zeki [*ZEKI* 1993].

[15][*MCILWAIN* 1996, pp. 11–74].

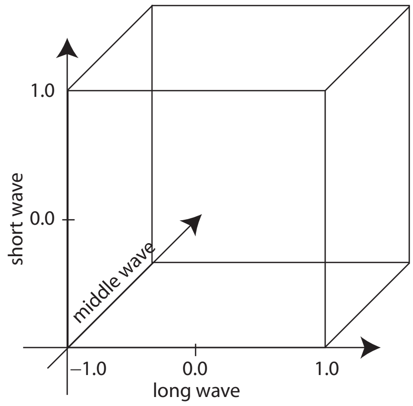

Fig. 4.4 The visual information-processing space. See also Fig. 3.2. Inspired by [LAND 1983].

2. Rods contain the photopigment rhodopsin and are capable neither of high resolution (the discrimination of two close spatial points), nor of discrimination of the spectral composition of light (at least two different receptor types are needed). Nevertheless, they are able to detect photons with low energy thanks to their mechanisms of reception and the fact that many of them, in the subsequent stage of information-processing, converge into single cells: The (on-center and off-center) bipolar cells. Indeed, nocturnal creatures have an eye anatomy that favors rods.

The eye mechanism shows therefore a complementarity between individuating specific frequencies and catching light in a superposition of any frequency [Subsecs. 1.2.6 and 2.2.5]. Since light is mostly in a superposition of different frequencies [Subsec. 1.2.2] but different objects emit preferred frequencies or specific combinations of frequencies, individuating a specific frequency is the surest mark of a specific, local source. The fact that large sets of entities such as wavelengths are monitored by a smaller number of basic features is an example of distributed representation [Sec. 3.8], as we shall see.

The distinction between sensitivity and individuation of frequency has also been tested by studying impaired patients. Patients with visual field defects in intensity perception were presented with tests concerning radiation intensity or wavelength[16]: The test was about either (a) achromatic target detection with differences in intensity only, or (b) both with differences in intensity and red/green discrimination on a low photopic achromatic background, so that here both intensity and wavelength were involved. While most patients succeeded in the second test, nobody did in the first one. This confirms that wavelength and intensity are treated differentially in visual information-processing.

An additional information acquisition is guaranteed by the retinal on-center and off-center cells, which are very good for contrast and not associated with any particular level of brightness [Fig. 4.5] but rather with differences in brightness between the item they individuate and the background.[17] In order to maintain a relative stability of brightness against different environmental conditions, the eye is probably able to compute three different components of brightness[18]: The illumination of the perceived object, the reflectance of the object surfaces, and the transmittance of the space

[16] [STOERIG 1987]. Note that wavelength is inversely proportional to frequency. [17] [*MCILWAIN* 1996, pp. 75–99]. [18] [*PURVES/LOTTO* 2002, pp. 42–64].

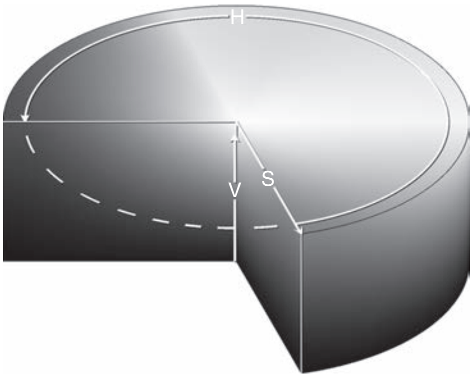

Fig. 4.5 Colors according to the three fundamental dimensions: (1) Brightness (V), the sensation elicited by different luminances (the local intensity of a visual stimulus), (2) hue (H), the degree of purity of each of the 4 fundamental colors relative to the other ones, and (3) saturation (S), the degree of approximation of each color to the central gray zone. Adapted from http://processing.org/learning/color/. Note that this representation fits that of Fig. 3.2. (The figure is reproduced in color in the color plate section.)

between the object and the observer. Relative brightness is also important, since we need to react to sudden light and to sudden shadow.

Apart from receptors and collectors, the retina also has ganglion cells [Fig. 4.6], which are divided[19] into [Fig. 4.7]:

- M cells that receive the input from a large number of photoreceptors (they are good for temporal resolution). M cells are suitable for rapidly changing stimuli, and therefore for perception of motion and for detection of sudden stimulus onsets.
- P cells that receive inputs from a smaller number of photoreceptors (they are good for spatial resolution). P cells, being sensitive to the onset or offset of light, play a role in object recognition.

This means that vision, in its first steps, is also a system for detecting dynamic and kinematic observables along two separate pathways [Subsec. 1.2.4]. I wish also to add that later shape formation (global features) can be obtained by integrating many different spatial stimuli over the P cells, while M cells are good for segregation of items.

Summing up, the retina already performs a significant computation before transmitting information.[20] This is due to the fact that the eye is directly connected to, and therefore can even be considered a part of, the cortex (this is a sort of peripheral system included in the CNS, which also explains its special relevance for the higher cognitive performances of primates) [Fig. 4.8].

4.3 First Steps of Perception

After initial information-processing in the retina, the signals are passed from the retina to the lateral geniculate nucleus (LGN), a six-layered structure of the thalamus [Fig. 4.8], where M and P cells project separately (actually, as we shall see, the projections are threefold).[21] Retinotopy

[19][LIVINGSTONE/HUBEL 1988]. [20][*GLYNN* 1999]. [21][*MCILWAIN* 1996, pp. 100–14].

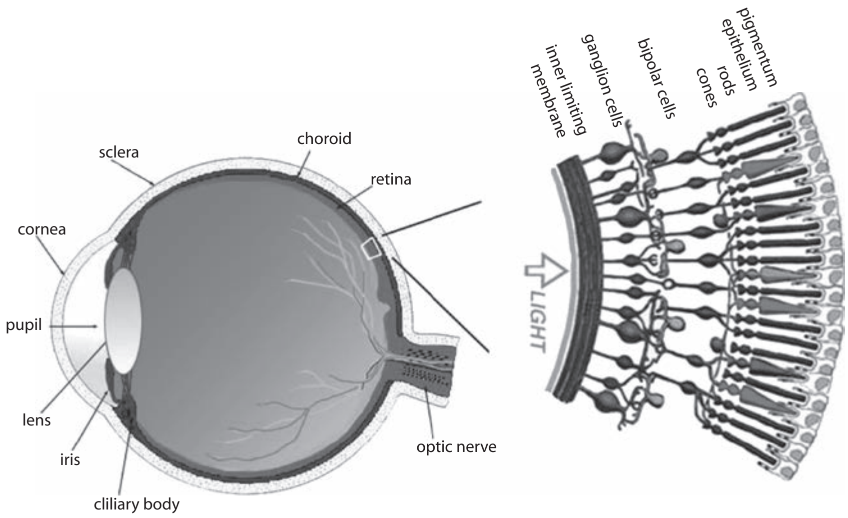

Fig. 4.6 The structure of the eye on the left and a cross-section through the retina on the right, showing three functionally distinct cells: Receptors on the right, collectors (middle), and ganglion cells (on the left). The receptors are cones and rods (the different cones catching different colors), while collectors are bipolar cells. Adapted from http://webvision.med.utah.edu/imageswv/Sagschem.jpeg. (The figure is reproduced in color in the color plate section.)

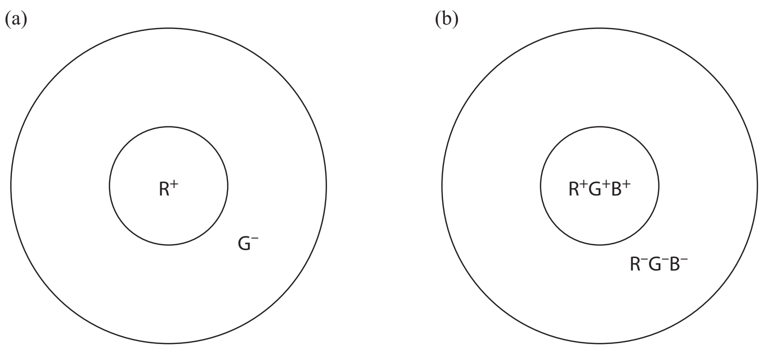

Fig. 4.7 P neurons (a) are excited over a small region by a single color (here red light), and inhibited over a larger region (here, by green light). M neurons (b) are instead excited by all wavelengths in the central region and inhibited by all wavelengths in the surrounding region.

is conserved [Sec. 3.7], and, among many other advantages, this provides a common framework within which the representation of an object at one location can be coindexed among disparate brain areas. It is a true labeling, whose importance will be understood later.

Therefore, at least two pathways may be distinguished in visual information-processing[22] [Fig. 4.9]:

[22][VAN ESSEN/GALLANT 1994] [*EYSENCK/KEANE* 2000, pp. 69–79] [*GOODALE/MILNER* 2004].

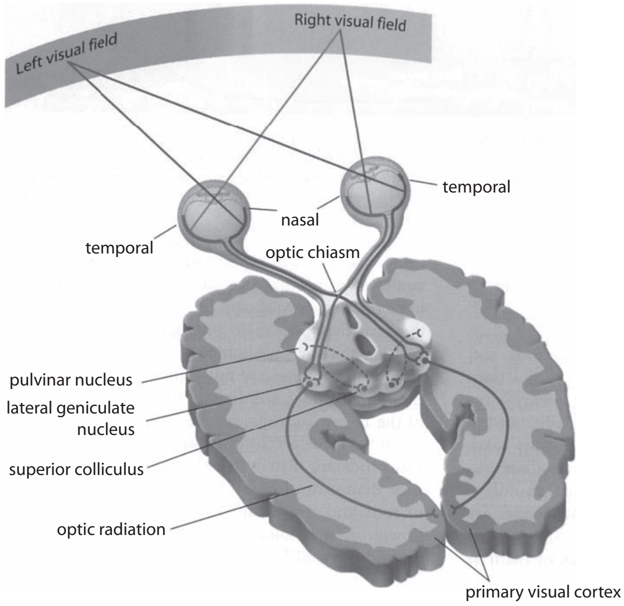

Fig. 4.8 The first steps in the treatment of visual information. Adapted from [*GAZZANIGA et al.* 1998, p. 152]. It is interesting to observe that, while vertebrates rely on accommodation, binocular convergence, or stereoscopic vision for obtaining three-dimensional vision, bees use the apparent size of familiar objects and objects' apparent motion [LEHRER *et al.* 1988]. (The figure is reproduced in color in the color plate section.)

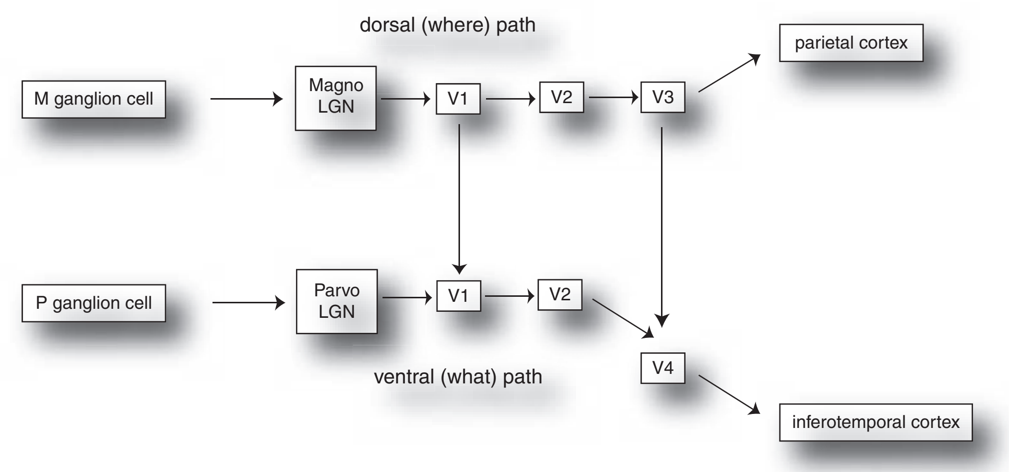

Fig. 4.9 A very schematic drawing of the two distinct pathways in the visual information-processing in a primate's brain; the one processing the "where" information, the other the "what". The parietal cortex is the upper part of the lateral cortex while the temporal cortex represents its lower part.

- The dorsal or magno LGN pathway, which is mainly concerned with motion and spatial processing, that is, with the *where is it?* question.
- The ventral or parvo LGN pathway, which is concerned with color and form processing, that is, with the *what is it?* question.[23]

Trevarthen, Schneider, Mishkin, and Ungerleider[24] discovered that corticocortical inputs from the striate Cortex are crucial for the visuospatial functions of the parieto-preoccipital cortex. Relative to the inferior temporal cortex, however, the parieto-preoccipital cortex was found to be especially dependent on ipsilateral striate inputs. Therefore, while pattern discrimination of the inferior temporal cortex is dependent on inputs from lateral striate cortex, visuospatial functions of the parieto-preoccipital cortex are equally dependent on inputs from the lateral and medial striate cortex.[25]

The dorsal system turns out not to be deceived by optical illusions, while the ventral may be, because the shape (the what) of the object must necessarily be independent of its distance from the perceiving subject. Exactly the opposite is true for where-perception: In this case, a strong perspective-like perception is necessary, in order to ascertain the true position of the object relative to the body. In other words, vision for localization and action is viewpoint-dependent and eventually uses short-living representation (due to the rapidity of moving objects or of the agent), while vision for shape perception is the opposite. The ventral stream is therefore more connected with memory and other long–term associations. Anesthesia has little effect on the visual selectivity of cells in the ventral stream, suggesting that these cells are not involved in the online control of the behavior of the animal. It is very important to realize that the dorsal stream also takes advantage of the contribution of subcortical information.[26] Moreover, reception of where-neurons (parietal lobe) is not selective while reception of what-neurons (inferior temporal lobe) is highly selective. Although there are mediators between these two functions and both channels contribute each to both ventral and dorsal streams[27] (but, obviously, most of the inputs to the dorsal stream are magno in origin), it remains true that the transformations carried out in the ventral (or occipitotemporal) stream allow the formation of cognitive representations, while those in the dorsal stream allow for the formation of goal-directed actions.[28] The fact that here we have two independent systems is also supported by the study of some vision impairments: Akinetopsia is an impairment of motion perception, while visual agnosia is an impairment of shape recognition. Specifically, since the ventral stream is associated with awareness (while the dorsal is not), this also explains a specific type of visual agnosia: Apperceptive agnosia.[29] Below, we shall consider several forms of visual impairment.

This organization of vision is due to the specific nature of light. However, a similar segregation in information-acquiring and processing is performed by the auditory apparatus, only that the localization is obtained here by comparing the sound perceived by the two ears. This is mainly studied for animals like owls and is not very well understood in the case of humans.

The major cortical path of visual information from the LGN is to the primary visual cortex (area 17 or V1). The world is topographically mapped in V1 in a smooth and continuous manner at the macroscopic level [Subsec. 3.7.1] while it is jittery and occasionally discontinuous on the microscopic scale.[30] Sometimes it is assumed that no significant transformation of the retinal input

[23] See also [BLYTHE *et al.* 1986] [WILSON *et al.* 1993].
[24] [TREVARTHEN 1968] [SCHNEIDER 1969] [MISHKIN/UNGERLEIDER 1982, UNGERLEIDER/MISHKIN 1982].
[25] This model is rather an oversimplification [DEYOE/VAN ESSEN 1988], that we shall partly correct below.
[26] [*JACOB/JEANNEROD* 2003, p. 3].
[27] [*MILNER/GOODALE* 1995, pp. 25–66].
[28] [GOODALE/HUMPHREY 1998, p. 186].
[29] [GOODALE 1995].
[30] [*KOCH* 2004, p. 78].

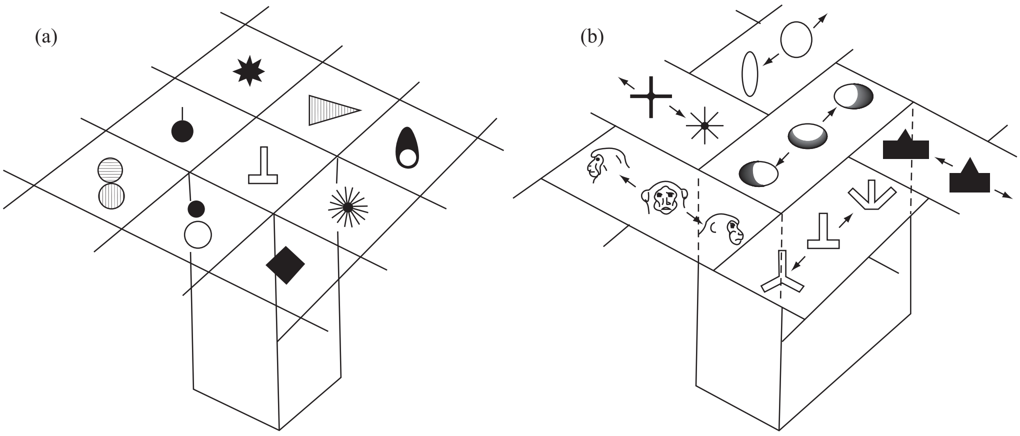

Fig. 4.10 (a) Columnar organization in the TE area of the temporal neocortex (the modules' size is about 400 μm in the horizontal dimension). (b) Neurons in adjacent modules of the same area respond to different but related stimulus features, and on this basis two or three may be linked into larger modules. Adapted from [TANAKA 1996].

occurs in the LGN. However, the forward projection from the LGN to the primary visual cortex is paralleled by a massive cortical feedback (in cats, ten times more fibers project back from V1 to the LGN than forward).[31]

In the passage from the LGN to the cortex,[32] two further categories of receptive structures add to the center-surround cells: (1) *Simple* cells, which respond to edges at particular locations and orientations (they are very important for motion perception), and (2) the so-called *complex* cells, which respond to information that is partly independent from location. This is a process of increasing specificity with respect to the form and increasing generality with respect to the viewing condition. Let us consider these two aspects [Fig. 4.10]:

(1) Hubel and Wiesel documented a highly systematic anatomical organization in the primary cortex:[33] There are columns with the same selective orientation[34] [Fig. 4.11]. Superimposed on this organization is a preference for ocular dominance (left or right eye). It takes about 18 or 20 columns to represent all orientations for both eyes, and this aggregate is called a hypercolumn. Each hypercolumn encodes information about a small sector of the retina, and adjacent hypercolumns represent neighboring sectors. Then, three different stimulus dimensions are simultaneously encoded here: Which eye sent the input, orientation, and retinotopic localization. However, the sequential information-processing hypothesis of Hubel and Wiesel,[35] according to which simple cells detecting edges give the input to complex cells that detect lines, and these to hypercomplex cells that are line segment detectors, was nullified when it was shown

[31] [*KOCH* 2004, p. 59]. [32] [*MCILWAIN* 1996, pp. 115–38].
[33] [HUBEL/WIESEL 1962]. Actually, the first hypothesis of a columnar organization of the cortex was formulated in [MOUNTCASTLE 1957].
[34] [GRAY/SINGER 1989]. [35] [HUBEL/WIESEL 1977].

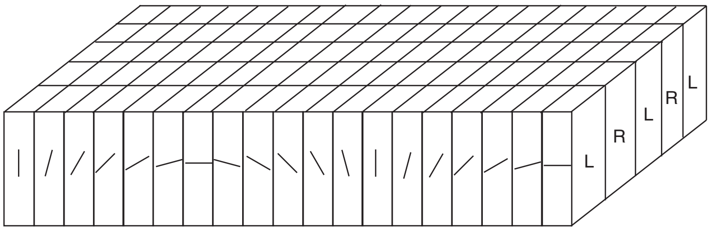

Fig. 4.11 Depiction of the organization of orientation selectivity and ocular dominance in the primary visual cortex. Inspired by [*FARAH* 2000a, p. 22].

that the predicted order of latency (the time in which the information is handed over) was not in accordance with their model.[36] Thus our knowledge of the brain's anatomy is then essentially clueless about the *functionality*, and this is the reason why the strict reductionist program is unlikely to work.

(2) In layers 2 and 3 there are other cells called *blobs*, which are located in the centers of local dominance columns. Blobs are suitable for color perception. Interblobs are good at tuning for high spatial frequencies and binocularity. They are good for shape perception. Both blobs and interblobs are the continuation of P processes. Blobs project to the visual area V2 (secondary visual cortex, area 18, also called prestriate cortex), which also has color-selective responses, which in turn project to the area V4. Objects' color depends on both the spectral reflectance of the surfaces and the composition of the incident light. The visual system is able to attain color constancy in V4 with the use of so-called Mondrian patterns, whose perceived color appears constant with varying illumination as long as the illumination is homogeneous. The visual activity is no longer dependent on wavelength. Instead, many V4 neurons respond to the color of the Mondrian patch and are not influenced by changes of illumination.[37] Summing up, the whole network can be cast as in Fig. 4.12.

4.4 Higher-Level Visual Perception

4.4.1 Loss of Retinotopy

In passing from the retina to higher visual areas which provide visual representations of objects, there is an increase in receptive field size. Both bottom-up and top-down mechanisms are at play[38] in a reversed Bayesian inference that will occupy us later on. In particular, when arriving at the visual association cortex (extrastriate visual cortex) the previous topographic organization of information (maintained up to the primary visual cortex, V1) breaks down [Subsec. 3.4.2]:

- The receptive fields of cortical neurons of V1 are less than a degree of visual angle (as are the receptive fields of the corresponding retinal ganglion cells and LGN neurons), while neurons in the extrastriate cortical areas have receptive fields often covering a substantial fraction of the entire visual field (up to 180° horizontally and 130° vertically) [Fig. 4.13].[39] The point is that up to the primary visual cortex the information is hierarchically organized and therefore also

[36] [*FREEMAN* 1995, p. 54]. [37] [NEWSOME/PARÉ 1988].
[38] [LEE/MUMFORD 2003] [YUILLE/KERSTEN 2006]. [39] [*PURVES/LOTTO* 2002, pp. 33–7].

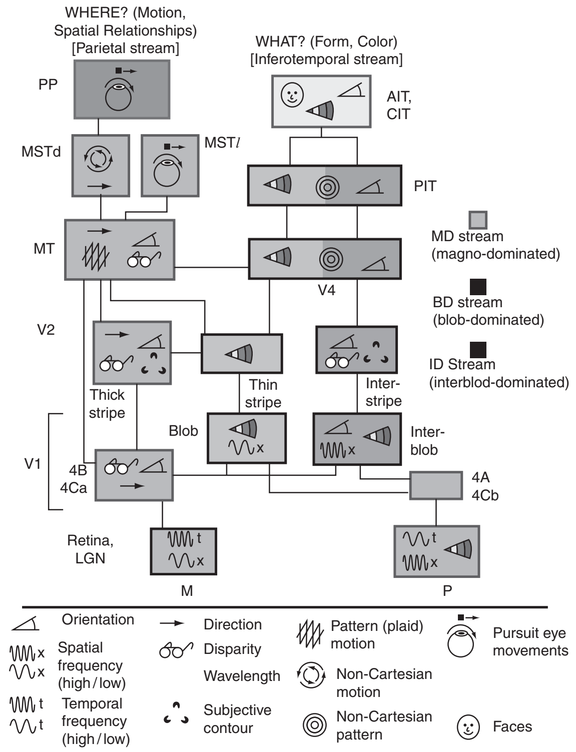

Fig. 4.12 Hierarchical organization of concurrent processing streams in the macaque monkey. Boxes represent visual areas, compartments within an area, and subcortical centers; solid lines represent major connections between structures (usually reciprocal pathways); and icons represent characteristic neurophysiological properties. Subcortical streams in the retina and lateral geniculate nucleus (LCN) include the magnocellular (M) and parvocellular (P) streams (gray and pink, respectively: see the color plate; the koniocellular stream, present in the source paper but poorly understood, is not shown here). Cortical streams at early and intermediate stages include the magno-dominated (MD), blob-dominated (BD), and the interblob-dominated (ID) streams (red, green, and blue, respectively). The PP complex is shown in orange. The IT complex includes posterior inferotemporal areas (PIT), which are components of the BD and ID streams, and central and anterior areas (CIT and AIT). Adapted from [VAN ESSEN/GALLANT 1994]. (This figure is reproduced in color in the color plate section.)

serially, while, in the passage from the LGN to the cortex, we observe a progressive fragmentation of the information (recall the organization in columns and hypercolumns), which is a clear index of the beginning of *parallel* information-processing [Sec. 3.8]. Moreover, it is interesting to note that the neglected region of space is not coded in purely retinal terms (for instance, by neglecting everything that is on the left of the center of gaze), but depends on the direction of the head and body or on the focus of attention.[40]

- Another important aspect to consider here is the feedback from these higher cortical areas to the areas V1 and V2. I have already remarked upon a similar feedback from V1 to the LGN

[40][*KOCH* 2004, p. 182].

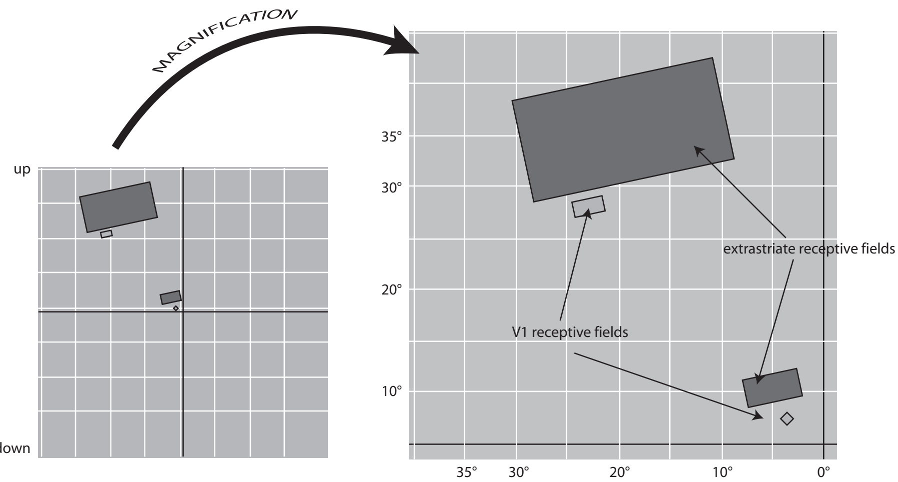

Fig. 4.13 Visual receptive fields in the primary visual cortex and in the extrastriate cortex. Adapted from [*PURVES/LOTTO* 2002, p. 36].

[Sec. 4.3]. This feedback as a whole will act in particular on the slow P neurons and will already have carved out the local analysis by means of a more global interpretation.[41]

- Finally, information of different types and coming from different sources is integrated into the higher areas. It has indeed been proved[42] that in the inferotemporal cortex (where form and color are processed) the analysis of the visual stimuli is not complex enough to specify objects on the basis of a single neuron, but groups of cells bringing different visual features are necessary in order to perform such a specification. Here, we find again a modular, columnar organization as above. However, the combination of information coming from different sources confers a new meaning on this organization. Here, we have the beginning of what we shall call, in the next part of the book, *representation*, which is no longer organized according to the linear combinatorics of information-processing but to new forms of nonlinear integration.

Therefore, when perceiving an object we must bind several stimuli referred to it. There are two methods for binding: Convergence of axonal projections (into a neuron) and dynamic selection due to temporal cues, in particular to synchrony [Subsec. 3.8.2]. The latter may be provided by a sequence of stimuli but is also endogenous.[43] The role of synchronization for binding has not been fully proven (it is indeed difficult to infer causality from correlation, and sometimes it is impossible, as in quantum mechanics [Subsec. 1.3.2]). Recently, the idea has been supported that, for binding form and shape, awareness is a necessary condition.[44] However, as I shall show in the next chapters, this is not necessarily the case.

The main modifications in object perception are due to[45]: Viewing positions, photometric effects, object settings (in different backgrounds), and shape changes. One of the most important features for distinguishing between different objects is depth. The cues for depth are occlusion, relative size, familiar size, and linear perspective. Another important feature is relative motion:

[41] [BULLIER 2001]. See also [*JACOB/JEANNEROD* 2003, pp. 57–61]. [42] [TANAKA *et al.* 1991].
[43] [SINGER 1999] [VON DER MALSBURG 1985]. [44] [MATTINGLEY *et al.* 2001] [ROBERTSON 2001].
[45] [*ULLMAN* 1996].

We tend to infer that two aggregates are different objects when they are in relative motion to one another. However, the form (shape) seems to be the most salient cue for object recognition. Object perception is also characterized by constancy: The apparent size and shape of an object do not change when the object is in motion.

The difficulty in perceiving objects is that there is an infinite number of three-dimensional possible configurations in the world that would give rise to the same pattern of stimulation on the retina.[46] Moreover, there is an infinite number of three-dimensional possible configurations of a single object: This can be seen by performing rotations of a given object about different angles along any of the three Cartesian axes [cf. Fig. 3.17]. It has been proved that, under normal circumstances, the changes induced by variations in viewing conditions of the same individual can be larger than the differences between distinct individuals. Therefore it is impossible to rely on simple image comparison and on pure association.

4.4.2 Elements of Objects

Visual recognition of objects therefore presents a really big explanatory problem, and many solutions have been proposed.[47] At one extreme we have the idea that we associate images directly with identified objects (strong dependence on viewing conditions) [see also Subsec. 3.7.2], at the other, the opinion that vision is completely independent of viewing conditions. The direct approach presents the problem that it relies on a simple notion of similarity—for instance, the Hamming distance between vectors [Fig. 3.20]. Now, the space of all perspectives on all objects would be prohibitively large, and in general the image to be recognized will never be sufficiently similar to images seen in the past.[48] Moreover, absolutizing the view-dependent aspect would run against experimental evidence. For instance, the *Drosophila* perceives some environmental details independently of the solid angle under which they have been perceived.[49] A mix of both approaches seems reasonable: We are dependent on viewing conditions but are able to transform our images in order to catch some invariants.

Essentially, in view-independent explanations there are two methods for facing the problem of the variety of viewing conditions in object representation[50]:

- The search for *invariant properties*. This will not work in many cases. For instance, in order to distinguish between a fox and a dog one needs a more precise description of the shapes rather than a restricted set of basic invariant properties. Moreover, there is no reason to assume the existence of relatively simple properties that are preserved across possible transformations. It is also impossible to find a set of invariant measurements that are independent of the viewing position: They must be tailored to the set of objects that need to be recognized.
- The *part description* introduced by Marr. Marr has built a computational theory of vision centered on the concept of visual description.[51] According to him, there are three major kinds of representation:
  (1) Primal sketch, a two-dimensional description of the main light-intensity changes in the visual input, including information about edges, contours, and blobs (essentially the content of Secs. 4.2–4.3). A first raw primal sketch consists of light intensities at the different pixels. Since this continuously fluctuates, several descriptions are needed. A full primal sketch is constituted when a label is given to any set of grouped elements (Gestalt principles are used

[46][*GAZZANIGA/ HEATHERTON* 2003, pp. 143–59]. [47][*FARAH* 2000a] [*EYSENCK/KEANE* 2000, pp. 83–96]. [48][*ULLMAN* 1996]. [49][LIU *et al.* 2006]. [50][*ULLMAN* 1996]. [51][*MARR* 1982].

here), and ambiguities are resolved only when there is convincing evidence as to what the appropriate solution is.

(2) 2 and 1/2-dimensional sketch: This incorporates a description of the depth and orientation of visible surfaces. Both sketch (1) and (2) are observer-dependent. A range map with local point-by-point depth information is constructed. Then the information of several maps is combined.

(3) 3D model description: Three-dimensional and viewpoint-invariant descriptions of shapes and relative positions of objects.

There are five criteria for this description:[52] (1) Accessibility, (2) scope, (3) uniqueness, (4) stability, and (5) sensitivity to different shapes. The primitive units that Marr proposed for describing objects are cylinders having a major axis [Fig. 4.14(a)]. The concavities are used to divide the visual image into segments.

There is some experimental evidence for Marr's sketches.[53] For instance, the anterior part of the macaque inferior temporal cortex, area TE, shows that neurons selective for 3D shapes are concentrated in the lower bank of the superior temporal sulcus, whereas neurons in the lateral TE are generally unselective for 3D shapes but selective for 2D shapes.[54] However, there are also some difficulties:

- Often objects are distinguishable not because of a different arrangement of the parts but because of detailed differences at specific locations. As we shall see, this is the issue of the distinctive mark.
- Not all objects are easily decomposable into parts (a sleeping cat, for example).
- Perception can be seen as a process in which sensory information, context, and expectation are combined to create the analogue of a hypothesis relative to an object. For this reason, a step-by-step process as envisaged by Marr does not accurately reflect the whole process of visual perception (as already remarked in Sec. 4.3), even if it remains true that the sketch describes the first steps or later particular substeps in a relatively good way. For instance, faces can be recognized without information about a three-dimensional layout, implying that the analysis of the three-dimensional form is not a necessary step for discriminating faces.[55]
- The visual system probably makes use of different types of shape units in order to construct complex structures. If this is true, we do not need a single microshape (like cylinders). In fact, Portland has extended this theory with the superquadrix proposal, and Biederman[56] showed that there are at least 36 basic geometric shapes (called geons) [Fig. 4.14(b)–(c)].
- The same geon should be used, for instance, in perceiving all cups, whereas we are indeed able to identify THE cup we normally use. Furthermore, perception consists in identifying the general configuration of objects, which is not the mere assemblage of parts, as we shall see.

Explanations of this kind have perhaps deemphasized the perception of *singular* objects and the *environmental context* of perception a little too much. In other words, they should be combined and corrected with viewpoint-dependent theories.[57] It seems that viewpoint-invariant mechanisms (like those envisaged by Marr and Biederman) are used in easy categorical discrimination tasks, whereas viewpoint-dependent mechanisms are used when the task requires hard, within-category discrimination, especially when we are confronted with new objects. There is then a

[52] [MARR/NISHIHARA 1978]. [53] [BOOTH/ROLLS 1998]. [54] [JANSSEN *et al.* 2000].
[55] [*PASHLER* 1998, p. 78]. [56] [BIEDERMAN 1987, BIEDERMAN 1990]. [57] [TARR/BÜLTHOFF 1995].

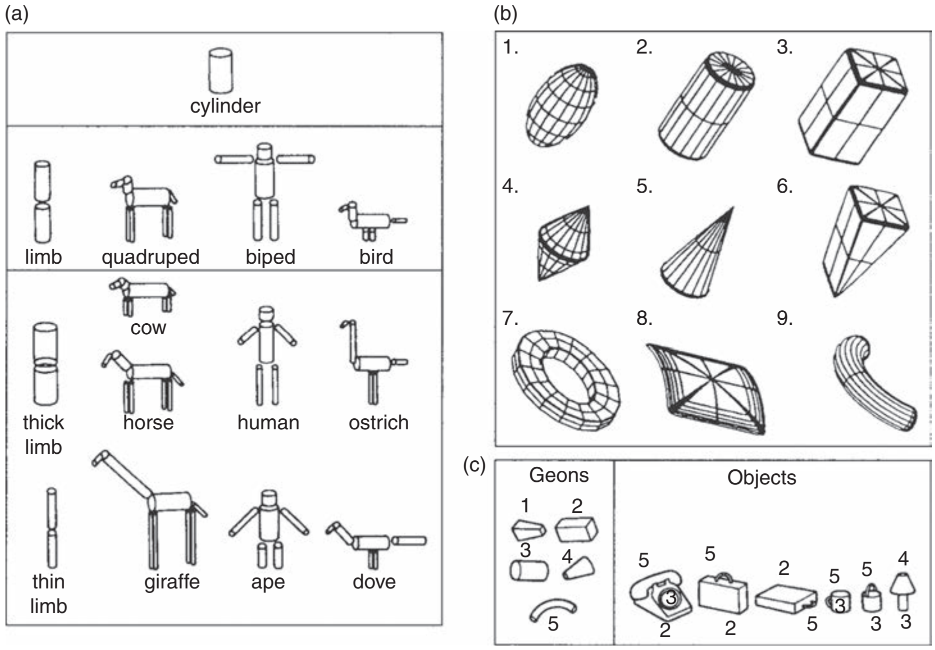

Fig. 4.14 Examples of volume-based primitives: (a) Marr's generalized cylinders, (b) Pentland's superquadrix, (c) Biedermann's geons. Adapted from [*FARAH* 2000a, p. 77].

sort of *a posteriori* automatization, as the perceived object is integrated into the experience of the perceiving agent, and it is here that viewpoint-invariant cues and processes can become subsequently dominant.

4.4.3 The Ecological Theory of Vision

A viewpoint-dependent explanation of vision was introduced by Gibson. Not by chance, it deals especially with motor aspects. He stressed that the information about a world that surrounds a point of observation is not just information about the world but also about the point of observation itself.[58] Therefore, exterospecific information and propriospecific (viewpoint-dependent) information are strictly connected. However, the latter type of information cannot be shared by other observers; this is the most fundamental perceptual ground for the distinction, at a visual level, between self and others. Gibson proposed a theory of direct perception without the involvement of internal information-processing or representation at all.[59] According to Gibson, perception invokes picking up the rich information directly via resonance, like a radio that is resonating with the signal contained in the electromagnetic radiation. The core of the theory is the movement of the individual (not necessarily a human subject) in its environment. The pattern of light reaching the eye is an optic array that contains all information about the environment and provides unambiguous and invariant information about the layout of objects in space [Figs. 4.15–4.16]. Some invariants in visual perception are the global focus of expansion (the motionless point toward which we are

[58] [*GIBSON* 1979, pp. 75 and 111]. [59] [*GIBSON* 1950, *GIBSON* 1966].

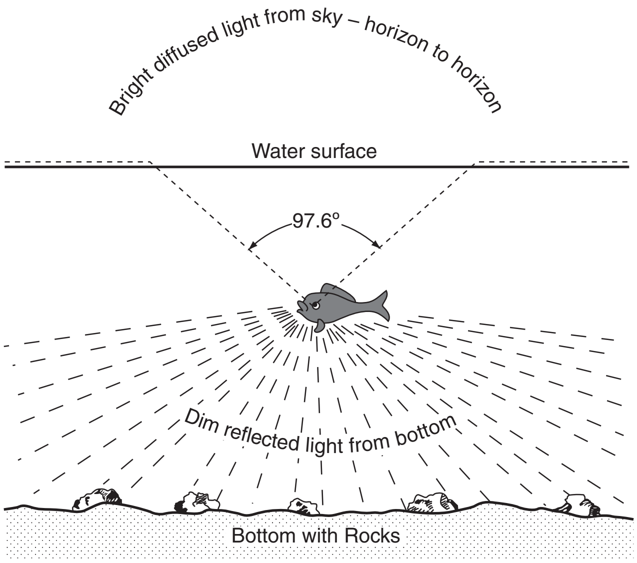

Fig. 4.15 Gibson's model of underwater vision. Adapted from [*GIBSON* 1966, p. 157].

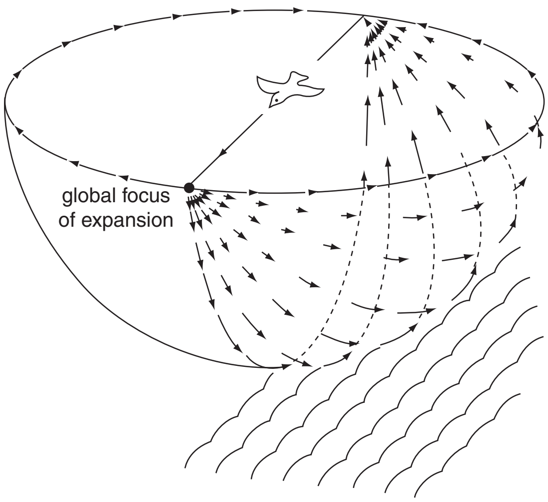

Fig. 4.16 Gibson's model of aerial vision. Adapted from [*GIBSON* 1966, p. 161].

directed) or the ratio of an object's height to the distance between its base and the horizon. All the potential uses of objects, i.e. their affordances, are also directly perceivable: Gibson called *affordances* the sensory constants of a particular biological species, which are the measure of the feasibility of carrying out certain actions. This is a concept that we shall develop in the second part of the book. The main idea of Gibson, therefore, is that vision displays hereditary structures.

Instead of Gibson's global radial outflow, according to which the overall outflow pattern specifies the direction, a "local focus of outflow" theory has been also proposed.[60] According to this theory, the direction of heading is determined by locating a stationary reference point (the local focus of expansion) in the environment. This account is especially interesting for subjects moving toward a fixed point.[61] Another very important approach[62] is focused on the time-to-contact, which is inversely proportional to the rate of expansion of the object's retinal image.

These kinds of approaches can explain very well how size constancy and the sense of the third dimension are innate visual abilities, and are also very fruitful for an ecological and dynamic understanding of object perception.[63] However, even if these viewpoint-dependent theories can probably account for many results, they cannot be considered as *general* explanations of vision. Experiences with chimpanzees show that they clearly have a mental (internal) image (map) of the location where the food is (food that they previously saw being put in that location).[64] This means that there is some information-processing at work [Sec. 3.7]. Moreover, affordances cannot explain the distinction, introduced by Fodor and Pylyshyn,[65] between seeing and *seeing as* (one can see a thing but without acknowledging it *as* the food one searched for). This is especially important for primates. Finally, the processes involved in identifying invariants not only deal with motion invariants like the global or the local focus-of-expansion but also with perceptual aspects. I have mentioned the possibility of integrating different theories. Constructivism (like that of Marr and Biederman) refers indeed to the ventral elaboration of information [Sec. 4.3], while the ecological approach to vision (like that of Gibson and Lee) refers to dorsal mechanism[66]: Therefore, we could establish an opposition between dorsal invariants versus ventral cues and variants.

4.4.4 Shape Perception

A third approach is to research "ventral" invariants but at a *global* and not at a local level (i.e. at the level of the global shape). This is the enterprise first endeavored by Gestalt psychology.[67] All laws of the Gestalt stem from the basic *law of Prägnanz*: Among the possible organizations in visual perception, the one possessing the best, simplest, and most stable shape will occur.[68] The most important aspect here is the figure–ground segregation [Sec. 4.1]. Proximity and similarity are in general the two criteria of grouping, which represents a more general aspect than binding, since it can also consist of the operation of collecting several objects [Fig. 4.17]. Similarity seems to have a temporal precedence on proximity[69] and both represent a higher perceptual manifestation of the where/what dichotomy. The *law of proximity* is grounded in the localization procedures of perception and, as we shall see, is the perceptual basis of metonymy, whereas the *law of similarity* is grounded in identification and is the basis of metaphoric transfer. To proximity and similarity Gestaltists also added (1) the law of good continuation, and (2) the law of closure. These two laws are actually not primary, since the law of good continuation is an application of the law of proximity (we suppose that a chain of proximal elements will follow even in zones that are not

[60][WARREN *et al.* 1988]. [61][*EYSENCK/KEANE* 2000, pp. 65–69]. [62][LEE 1976, LEE 1980].
[63][*KELLMAN/ARTERBERRY* 1998]. [64][*SAVAGE-R. et al.* 1998, pp. 37–39]. [65][FODOR/PYLYSHYN 1981].
[66][NORMAN 2002]. [67][WERTHEIMER 1923].
[68][*KOFFKA* 1935, p. 138]. See also [*EYSENCK/KEANE* 2000, pp. 25–49]. [69][QUINLAN/WILSON 1998].

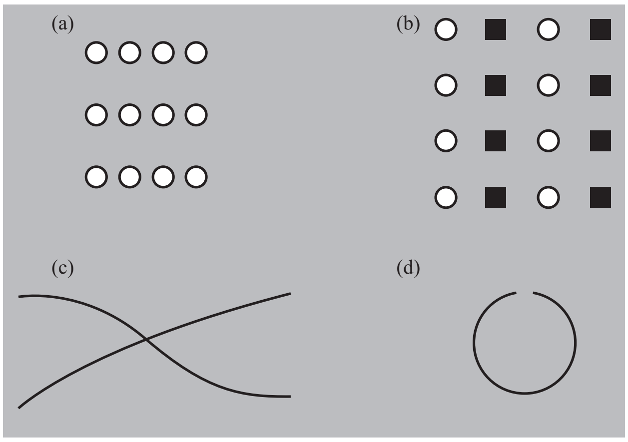

Fig. 4.17 Some of the Gestalt perception laws: (a) The law of proximity, (b) the law of similarity, (c) the law of good continuation, (d) the law of closure. The law of proximity is grounded in the localization procedures of perception and is the perceptual basis of metonymy, whereas the law of similarity is grounded in identification and is the basis of metaphor. Adapted from [*EYSENCK/KEANE* 2000, p. 26].

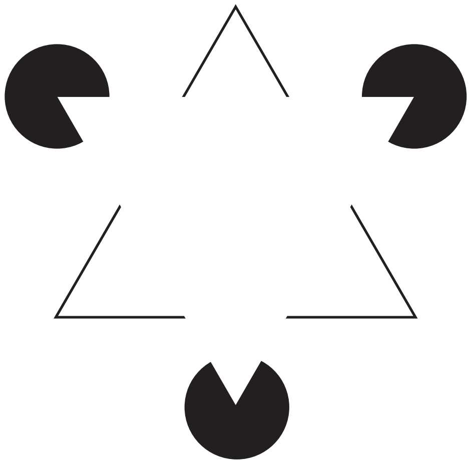

Fig. 4.18 Example of illusory contours giving rise to a ghostly triangle. It is an instance of both the law of good continuation and of proximity.

visible) and the law of closure is an application of the law of similarity (we associate the perceived shape with a memorized structure) [Fig. 4.18].

As we shall see, the two main ways for interpreting objects are exactly those proposed by the Gestalt theory, i.e. contiguity and similarity. Notwithstanding these important results found by the Gestalt theory, there are still several problems. The Gestalt theory was discredited among neurobiologists when Roger Sperry showed in 1958 that, by placing strips of mica and silver needles

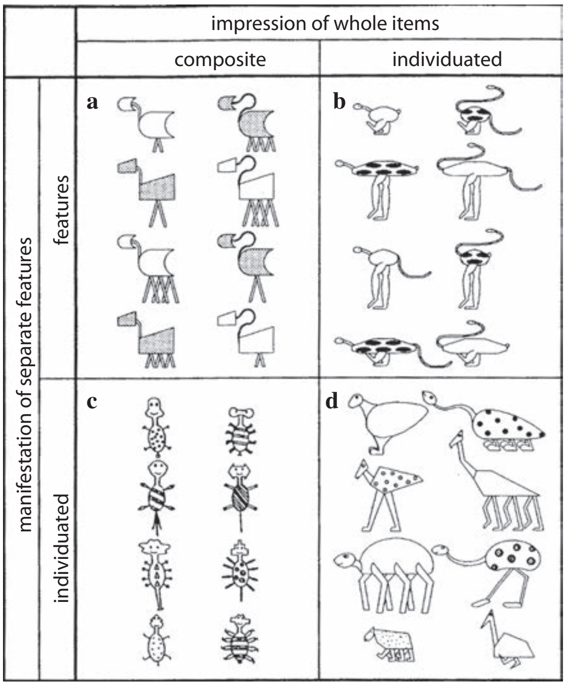

Fig. 4.19 Holism versus localism and invariant versus individuated aspects in vision. A: Marr-like composition of elementary units (e.g. a table). B: Gestalt-like individuated shapes (e.g. a pitcher). C: Functional tools (e.g. a hammer). D: Living beings (e.g. a cat). Adapted from [REGEHR/BROOKS 1993]. It is interesting to note that, in this paper, the shapes of functional tools are not well understood in their specificity.

into the visual cortex of cats and monkeys, distortion in the electrical fields had negligible effects on behaviors involving perception, in contrast to Gestalt's assumption.[70]

Moreover, for the Gestaltists, grouping is a bottom-up process where no object information enters, but contrary evidence was reported by Vecera and Farah,[71] who stressed that top-down activation partly guides the segregation process. This is particularly true when we perceive shapes in motion. Indeed, grouping must occur later than the Gestaltists supposed. For instance, proximity in three-dimensional space occurs with depth perception. It turned out, after studies by Restle,[72] that what is first perceived is a configuration of random points (in general in coordinated motion, that is, when points move together sharing a common fate [Subsec. 4.4.1]) that requires less calculation: If displays are viewed through a collimating lens, so that the points are at optical infinity, observers tend to give three-dimensional interpretations. When the points break away from a certain surface, then observers tend to see the distances between points as constant, reducing several motions to the motion of one rigid body, which is more "economic" than the previous complex of different motions. This is also evidence for information coding in the peripheral nervous system [Subsec. 3.3.1]. Therefore, the mechanism envisaged by Gestalt theory does not always seem appropriate.

[70][SPERRY 1958] [*FREEMAN* 1995, pp. 33–4]. [71][VECERA/FARAH 1997]. [72][RESTLE 1979].

Fig. 4.20 Some basic tools. Adapted from www.liquidlearning.com.

It is difficult to give answers to all these problems, but a synthesis of all the previous results can be helpful. An initial, very important distinction could be between holism and localism (or analyticity) in vision. These two aspects may be considered complementary [Subsecs. 1.2.6 and 1.3.3], as far as the former deals with the extent to which an item's features cohere into an individuated whole pattern, while the latter deals with whether a feature occurs with identical forms in different items.[73] However, this distinction does not cover a second one: That between invariant and individuated, so that we can distinguish between individuated composed objects and individuated wholes [Fig. 4.19]: In this case, we may think about objects whose global shape is not a cluster of parts (there are no distinguishable components) but is very much individuated. We know these sort of things very well: They are living beings. Indeed, the shape of a cat or dog responds to these criteria [Subsec. 4.4.2]. A different and somehow intermediate case is represented by our working tools, like hammers, scissors, pliers, and so on. In this case, they have clearly *distinguished* parts, but these parts are *specific* to the tool, since they must coalesce into a functional unity[74]: For instance, a hammer's handle is very different from pliers' [Fig. 4.20]. Therefore, Marr's theory (interchangeable composite objects) and the Gestalt theory (interchangeable whole shapes) correspond to very elementary processes of vision that apply very well to (less complex) inanimate objects, while a deeper distinction between specialized tools (individuated composite objects) and living beings (individuated wholes) can be drawn.

4.4.5 Object Perception

A common error is to suppose that later areas of vision reconstruct objects as they really are in the world.[75] This is illogical: Why should vision firstly reduce the perception of objects into small parts [Secs. 4.2–4.3] and later reconstruct them as they are? This is also phylogenetically counterintuitive, for then if our perception of objects were a pure template of them, we would expect

[73][REGEHR/BROOKS 1993]. See also [ZEKI 2001].
[74][MOSS *et al.* 1997, TYLER/MOSS 1997, MOSS/TYLER 2000]. [75][*FARAH* 2000a].

such a perception to be present not only in late evolved species, e.g. primates, but already in lower forms of vertebrates, especially considering its relevance for survival [Subsec. 3.7.2]. Consequently, either the more primitive animals should already have the ability to perceive complex objects, which does not seem to be the case, or object perception does not matter at all, which again seems hardly the case. Thus, the true solution could be the following: Objects *are constructions* of the brain, but they are constructed by using the material *already present* in the primary visual area (as well as in other primary sensory areas), and therefore their configuration is *inferred* according to certain regularities in the interaction between organism and environment[76] (the subject of the next chapter). Moreover, although objects are emergent constructs, delocalized aspects are already present in the first steps of the visual process [Sec. 4.2]. This means that objects' perception *does matter*, but does matter only *for advanced organisms* that can interact with them in an appropriate manner, and have coevolved with a niche environment in which such sophisticated perception is important,[77] according to the above two principles of perception [Sec. 4.1].

Evidence for this comes from the way the vision of objects is built. This growing complexity of vision does not consist in a linear process through successive steps from the retina to higher visual areas (as would be expected if object perception were based on templates). It is true that visual information is transferred in a feedforward fashion from low-level modules to higher-level ones, as we have seen in Sec. 4.3. However, we have also seen, in addition, that feedback connections transfer information in the reverse direction. Receptive field properties seem to reflect mostly the convergent–divergent feedforward cascade of information-processing, while, as already remarked, feedback connections outweigh the set of feedforward connections.[78] Strict feedforward models are indeed subject to the combinatorial explosion (cascade of increasingly complex information).

However, we should also avoid the opposite danger: To think that perception can be reduced to our own constructions and that the external world and its configurations are irrelevant. This is hardly the case, since we very often experience that the external stimulus *interrupts* our chain of thinking and representing.[79] This shows that we are not free in our perception of the external world.

Therefore, one is led to the conclusion that the higher visual areas extract information already implicit in inputs coming from the primary visual cortex. Although right, this must be properly understood: The correct point of view is to say that structures and relations are always implicit and potential [Subsec. 2.4.3], not only for the perceiving subject but for *any other* action that can be performed upon them, around them, or through them. In other words, we always perceive objects and properties of objects due to their interactions with us and even among them, since, in this way, certain configurations of things can be activated and realized.[80] This explains my previous guess that all properties are relational and interactional [Subsecs. 1.3.3 and 2.3.1]. As we shall see later, the feedback that the brain receives from the environment, especially when the high organism actively interacts with its environment, allows for the tuning of the represented structure in the brain with the structure in the world. However, this does not mean that the structure in the brain "mirrors" the structure in the world. They have a commonality because of the simple fact that they are (evolutionarily and developmentally) associated, but the important point is

[76][*VON HELMHOLTZ* 1867, pp. 586–9]. See also [GOPNIK 1993a] [CHURCHLAND *et al.* 1994].
[77]Similar considerations have been developed about color vision in [HILBERT 1992b].
[78][LAMME/SPEKREIJSE 2000].
[79]An issue pointed out in [*HERBART* 1824–5, Sec. 3], the relevance of which we shall see in the next part of the book.
[80][*VON HELMHOLTZ* 1867, p. 589].

that the structure in the brain is associated to the structure in the world since it has been *selected* as an appropriate response by the organism, and not because it is an iconic reproduction—again a form of generalized Darwinism [Subsecs. 2.2.3 and 2.2.6].

4.4.6 Face Recognition

A surprising result is that mid processing stages of perception are distributed [Sec. 3.8 and Subsec. 4.4.1], whereas late or higher ones are sparse, even if they are still population-based.[81] In the later processes, a single neuron can become associated with a given item acting as a mark, for instance individuating a certain person.[82] Since this association is independent from the cluster of particular properties we attribute to the person, which are always associated with the specific perspectives under which this person is perceived, it is unlikely that such an invariance could be explained by an ordinary perception process. As mentioned, it is a manifestation of a *marking* ability (individuating an item) that is not in contradiction with the distributed representation of the properties (identifying) [Sec. 4.1]. This also confirms that a single neuron or few neurons cannot be, in a proper sense of the word, a representation of that perceived item[83] [Subsec. 3.7.1]. This marking process, as we shall see in the next part of the book, can be generalized to the whole of brain activity.[84]

Let us consider this mechanism a little. The inferior-temporal cortex (IT) cells (the area where the "where" path culminates [Sec. 4.3]) are highly selective for faces, hands, and so on, indicating a certain marking process that is independent of perception of form, color, and so on. Indeed, whole faces undergo little or no decomposition at all. The reason is that distributed descriptions are not well-suited to represent a large number of items simultaneously (i.e. very complex objects like faces) because, with large numbers of items to be considered, it is much more possible that these elements can interfere.[85] Then, a marking system can be very helpful in solving such a problem. It is also important to understand that face recognition is very much context-sensitive. Indeed, face recognition is more orientation-sensitive than the recognition of other types of objects. Interestingly, adults and children older than 10 years cannot easily recognize faces that are upside down.[86] These different recognition processes belong to different systems, and are anatomically distinct.

To understand this point, let us now consider in more detail the distinction between whole and parts introduced in Subsec. 4.4.4. Objects (with the exclusion of living beings) are represented much more in terms of their components [Subsecs. 4.4.2–4.4.4]. In perceiving compound inanimate objects (categories A and C of Fig. 4.19), we use *first-order* relational information (spatial resolution of the parts of an object relative to another one), whereas in recognition of wholes (categories B and D of Fig. 4.19) like animate beings we make use of *second-order* relational information, which exists only for objects whose parts share an overall spatial *configuration* and consists of the spatial relations of the parts relative to their prototypical arrangement (e.g. the species-specific form, which is independent of the specific state or situation of the animate being).[87] Now, face representation is different from both first-order and configurational levels of visual analysis.

The reason is that at a configurational level of analysis, single properties are psychologically real or explicit, whereas in perception of faces they are not so. One could, of course, extract such properties from a holistic representation, and in this sense holistic representations implicitly contain

[81] [BRINCAT/CONNOR 2004]. [82] [QUIAN Q. *et al.* 2005]. [83] [CONNORS 2002].
[84] See also [CONNOR 2005]. [85] [*FARAH* 2000a, pp. 115–46]. But see [ABBOTT *et al.* 1996].
[86] [CAREY/DIAMOND 1977]. [87] [DIAMOND/CAREY 1986].

both first-order and configurational features. However, holistic face recognition is a level higher than pure Gestalt.[88] It is a special case of the perception of living beings. As we shall see, there are important differences between schemata and categories. A Gestalt is related to a perceptual schema, while animate beings and faces are related to categorization. However, the crucial point is that faces are not only individuated (like animate beings) but are also related to *individuals*: Face recognition involves within-category discrimination (sense of individuality) whereas perception of other items involves between-category perception.[89]

For this reason, an important issue here is whether face perception deals with known individuals or not. V. Bruce and A. Young[90] provided evidence that the recognition of familiar faces mainly depends on very specific recognition elements of the face, personal identity nodes, and name generation (all expressing marking actions), whereas processing of unfamiliar faces requires more structural encoding, expression analysis, and facial speech analysis (a true information processing of the overall Gestalt). There are patients[91] who match faces and names of famous people without recalling autobiographical information. By using a positron emission tomography (PET) technique, able to produce a three-dimensional map of functional processes in the brain, it has been shown[92] that a face-gender categorization resulted in activation changes in specific areas of the right extrastriate cortex. In particular, it is necessary to distinguish between two specific brain regions involved in face recognition[93]: We have a system for acknowledging invariant and universal aspects of faces located in the lateral fusiform gyrus and another for the recognition of faces of individuals located in the superior temporal sulcus.[94] In prosopagnosia (the impairment in face recognition), the impairment is limited to the invariant aspects, whereas in capgras syndrome the emotional acknowledgment of single faces fails. Cerebral activation during object recognition, instead, essentially occurs in the left occipito–temporal cortex and does not involve the right hemisphere regions specifically activated during the face-identity task.

It is interesting to observe that, also when identifying handwriting of a specific individual, perception becomes holistic and individuated.[95] Summarizing, it seems to me that the general lesson is that perception, when faces and handwriting are involved, concerns perception of *individuals*, while perception of (animate or inanimate) objects is more schematic and *general*. Therefore, there are reasons to believe that tools and living beings are not perceived as individuals (with the exceptions of pets and personal belongings of particular value, as we shall see). This is, however, a point to be reprised later on, as it has a general significance with regard to the way humans and primates categorize. Indeed, perception of individuals is a later product of evolution, in accordance with the analysis developed in the previous subsection. As a matter of fact, monkeys also show a certain sensibility to face recognition. Macaque monkeys[96] have been trained to look left if a face has been recognized and to turn right if a nonface has been shown. After training, several images have been shown, even of blurred faces. Cues for recognition were the profile and both the eye and mouth regions. Monkeys learn quickly to discriminate faces with different emotional expressions.[97] This ability is invariant with changes of color, brightness, size, and rotation.

[88] [FARAH *et al.* 1998]. [89] [*GAZZANIGA et al.* 1998, pp. 231–4].
[90] [BRUCE/YOUNG 1986]. See also [*EYSENCK/KEANE* 2000, pp. 106–16]. [91] [DE HAAN *et al.* 1991].
[92] [SERGENT *et al.* 1992]. [93] [*JEANNEROD* 2006, pp. 99–127]. [94] [HAXBY *et al.* 2000].
[95] [*GAZZANIGA et al.* 1998, pp. 235–7].
[96] [AFRAZ *et al.* 2006] [DICARLO 2006]. See also [KANWISHER *et al.* 1997]. [97] [DITTRICH 1990].

4.5 Some Impairments of Vision

4.5.1 Cortical Blindness

Cortical blindness is damage to visual perception of the *what* that does not affect motor reactions or perception of motion.[98] Indeed, Barbur *et al.* pointed out that blindsight patients show high sensitivity to fast-moving objects, but neither shape nor size is discriminated.[99] This field of study owes a lot to work of Weiskrantz,[100] who presented interesting evidence that cortical-blind subjects still show a pupillary response to light. Paradigmatically, a patient with a restricted lesion of the right occipital lobe (where the primary visual area, V1, is located) was investigated in order to assess the possible existence of some visual capacity in his hemianopic field which was conventionally considered blind. Though the patient had no awareness of seeing, he could move in the direction of visual stimuli with considerable accuracy, could distinguish between the orientation of a vertical and a diagonal line, and could discriminate the letter X from the letter O. The patient could also differentiate a grating of vertical bars from a homogeneous field. These findings show that a certain visual capacity can remain after damage to the striate cortex resulting in blindsight.

Perenin and Rossetti[101] have even shown that cortical-blind patients could "post" a card in an open slot but were unable to describe it. They confirmed the previous results in this way, since their patients could also guess (via eye movements) whether they were being shown a circle, a horizontal line, or a vertical line, but they could not see the shape, lacking any combinatorial or Gestalt processing.[102] They were also sensitive to a certain extent to colors—because of the survival of some "perception" cells,[103] but, as mentioned, they were especially sensitive to motion, which is strictly related to movements of head and eye.

A possible explanation of these abilities is that there are fast pathways directly to the prestriate cortex bypassing the (disrupted) V1 area. Ffytche, Guy, and Zeki found some evidence for this interpretation.[104] The studied visual field was $30^\circ \times 20^\circ$. The parallelism is dependent on the characteristics of the stimulus. Signals relating to fast visual motion (a speed of $22^\circ s^{-1}$) reach the prestriate cortex (located above the striate cortex) before they reach the striate cortex. Signals related to slow visual motion (speeds $< 6^\circ s^{-1}$) are traded to the prestriate cortex through V1. This means that the parallelism is not rigid but dynamically tuned to the stimulus.

Let us now take a short look at the opposite form of impairment: Patients that are disturbed in their visuomotor system and therefore show deficit in actual grasping, show good grip scaling when "pantomiming" a grasp for an object that was seen earlier and that is no longer present.[105] In the case of apraxia, there is an impairment of action representation, which implies that the patients affected by this disease cannot pantomime an action, even if they show no basic visuomotor impairment.[106] The superior parietal and intraparietal sulcus would monitor the action on objects, whereas the inferior parietal lobe would monitor the action with objects (for tool use and action programming[107]).

4.5.2 Visual Agnosia

Cortical blindness implies *visual agnosia*. Lissauer[108] distinguished two types of object agnosia:

[98] [*MILNER/GOODALE* 1995, pp. 67–86]. See also [FARAH 1991].
[99] [BARBUR *et al.* 1980]. See also [PTITO *et al.* 1991]. [100] [WEISKRANTZ 1990, WEISKRANTZ *et al.* 1974].
[101] [PERENIN/ROSSETTI 1993]. [102] See also [MILNER 1997]. [103] [STOERIG/COWEY 1989].
[104] [FFYTCHE *et al.* 1995].
[105] [GOODALE *et al.* 1994, MILNER *et al.* 2001]. See also [PERENIN/VIGHETTO 1988].
[106] [*JEANNEROD* 2006, pp. 13–15]. [107] [GLOVER 2004]. [108] [LISSAUER 1890].

- *Apperceptive agnosia*: The subject cannot achieve a coherent percept of the structure of an object, and
- *Associative agnosia*: A subject is able to achieve such a percept but still unable to recognize the object (it is a semantic disease).[109]

Therefore, disease in perception is typical only for apperceptive agnosia, while associative agnosia should consist in the connection failure between perception and reference to a certain object. Patients affected by apperceptive agnosia cannot copy simple objects, though they can draw objects on the basis of long-term memory. It is possible that in associative agnosia there is also some impairment of perception: Subjects need very long time intervals in order to draw an image of an object. Both types of agnosia are concerned with the ventral stream [Sec. 4.3]. Apperceptive agnosia is probably concerned with an impairment of the elementary ability to build perceptual schemata, while associative agnosia patients show a certain impairment in their concepts but not in understanding the functionality of the related objects,[110] which, as I shall show, is a typical mark of a level of information treatment that comes "before" the establishment of true concepts, namely categorization.

Apperceptive agnosia can be further divided into[111]

- Apperceptive agnosia in a narrow sense: Here the cause seems to be some disease in organizing and ordering, while local properties—related to sensation and to elementary perception—of the visual field are well perceived (color, contour elements, etc.).
- Dorsal simultanagnosia, which is a limitation of visual attention and of the perception of spatial relationship between objects (but not of their shape). Here a patient can generally perceive only one object, without awareness of the presence or absence of other stimuli.
- Ventral simultanagnosia: Patients can see multiple objects but generally need a lot of time for recognizing them (for example they read while spelling each word). Here, the recognition is piecemeal, it is limited to one object at a time, although—in contrast to dorsal simultanagnosia—other objects are seen.

When we speak of *associative agnosia*, we distinguish between

1. Associative agnosia in a narrow sense: Intact visual perception and normal recognition of objects through the other sensory channels, but difficulty in recognizing a variety of visually presented objects, for example in naming or grouping objects according to their semantic category.
2. Pure alexia: Patients cannot read normally despite visual capabilities and the ability to understand spoken language and to write.
3. Prosopagnosia: The inability to recognize faces. In prosopagnosia much of the processing of familiar faces can remain intact despite absence of awareness that recognition occurs (in other words, patients are able to perform same/different judgments about familiar faces faster than about unfamilair faces, even if not knowing that the former are in fact familiar faces).[112]

In an important study on prosopagnosia, A. Damasio and coworkers[113] stressed that patients can recognize other persons by identifying gestures and posture, voice, and also facial expressions. It is important to note that these patients also show problems in recognizing cars, pets, and personal effects. As I posited earlier [Subsec. 4.4.6], the problem here seems to be the failure to recognize the uniqueness or individuality of the perceived items. In fact, for Damasio an entity generates

[109] [*MILNER/GOODALE* 1995, pp. 120–55]. [110] [*JACOB/JEANNEROD* 2003, p. 81].
[111] [*FARAH* 1990, pp. 1–92]. [112] [DE HAAN *et al.* 1987]. [113] [DAMASIO *et al.* 1990b].

a multiplicity of representations within the same sense and across many sensory modalities. The important point is that, even when the entity remains the same, the context and the state of the perceiver are likely to vary. The number of combinatorial arrangements (very high for faces and less so for single expressions) can define the contextual complexity. Moreover, human faces almost constitute a continuum, which implies that faces are especially difficult to distinguish in their individuality but, conversely, are especially suitable for identifying individuals.

4.5.3 Problems with Specific Concepts or Categories

Although this is not the context to deal with the issue of conceptualization, it would be beneficial to consider some impairments in concept formation that are somehow related to vision.

A well-known case is that of a patient affected by global dysphasia showing impairment in object (but not living-being) categorization.[114] In the category of objects, he was more impaired in the comprehension of small manipulable objects than of large man-made ones. Moreover, he showed good comprehension of proper names having a unique and well-known referent (e.g. Churchill) and a worse one of common names (e.g. Jones). A very well-known case is that of a patient impaired in naming fruit and vegetables[115] but still able to categorize them.[116] This means that different weighting values from multiple sensory channels may be important for the acquisition of different categories, as well as the sensory/motor contribution.

Let us consider the impairments of the distinctions animate/inanimate and concrete/abstract in particular.[117] Humphreys and coworkers[118] found that normal subjects also named pictures of living things more slowly than pictures of nonliving things. Similar results are true with monkeys.[119] The fact that there is a double dissociation between the perception of objects and the perception of living beings can be explained by the fact that objects can be *manipulated* whereas living beings are mostly only *represented*. For organisms the problem is a high correlation of attributes and a high level of perceptual overlap [Fig. 4.19D]. Inanimate objects (especially the artificial ones) mostly have clearly defined functions. Therefore, there is a dissociation between *perceptual* features and *functional* significance. In the case of tools, structural properties are indeed strictly related to their function in terms of actions that can be undertaken with these objects [Fig. 4.19C]. A possible explanation of this dissociation is a diversification in semantics. T. Shallice cited three types of evidence supporting the multiple semantics hypothesis[120]: The existence of modality-specific aphasias, modality-specific semantic memory impairment effects, modality-specific priming effects. Caramazza *et al.*, instead, did not think that one can speak of multiple semantics.[121] Therefore, the impairment is not necessarily category-specific (and therefore semantic) but is probably a consequence of a *modality*-specific problem: (i) Living beings are more similar to each other than nonliving things are and (ii) tools are also able to be manipulated.[122]

This is confirmed by other studies. As already mentioned [Subsec. 4.4.6], it is important to consider that the information about characters of visual items is a distributed network of discrete cortical regions. Within this network, the features that define an object are stored close to the primary sensory and motor areas that were active when information about that object

[114] [WARRINGTON/MCCARTHY 1987]. [115] [HART *et al.* 1985]. [116] [FARAH/WALLACE 1992].

[117] [WARRINGTON/SHALLICE 1984] [DE RENZI/LUCCHELLI 1994]. [118] [HUMPHREYS *et al.* 1988].

[119] [GAFFAN/HEYWOOD 1993]. [120] [*SHALLICE* 1988, pp. 269–306].

[121] [CARAMAZZA *et al.* 1990]. More recently, Caramazza has strongly supported the thesis of a semantic (conceptual) organization of human perception [CARAMAZZA/SHELTON 1998], an important issue to which we shall come back in the third part of the book.

[122] [GAFFAN/HEYWOOD 1993].

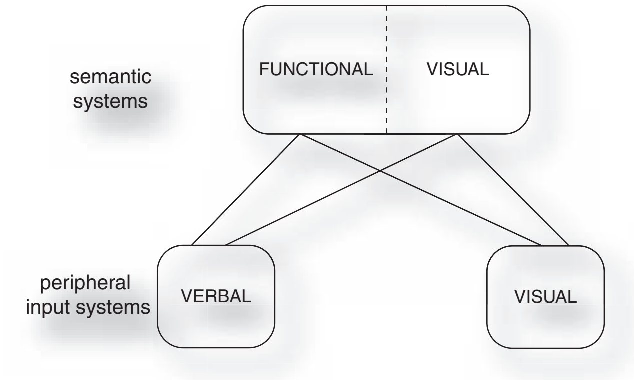

Fig. 4.21 Farah and McClelland's model. The differences are here in modality not in category.

was acquired.[123] Thus, the organization of semantic information parallels the organization of the sensory and motor systems (concepts must somehow be mapped into perceptual schemata and categories). An important finding is that naming tools is associated with activity in the left middle temporal gyrus in the same region that is active in verb generation and is also associated with a region in the left premotor cortex in the same region that is active when subjects imagined grasping objects with their dominant hand. In contrast, naming animals is associated with the activation of the medial occipital cortex, which is stronger left than right. It seems that this reflects top-down activation. Animals are then defined by their physical form and when the differences are subtle, the occipital cortex is brought into play. I also mention that people's names are restricted to the left temporal lobe.

Taking into account some of these aspects, a model proposed by Farah and McClelland[124] supposed a distinction between visual and functional units to which visual inputs enter (the units are also interconnected) [Fig. 4.21]. While the visual units possess information about the visual properties of objects, the functional units possess semantic information about the use of objects or about appropriate ways of interacting with them. It is a simple model that accounts for object recognition, the double dissociation between impairment in recognition of living and nonliving items, and also for the fact that impairments in recognizing living beings is more common. A limit of the model is to suppose that all units are interconnected with the consequence that patients should show impairment both in visual memory for objects and in understanding their functionality. On the contrary, there is evidence that one can be impaired in the perception of a tool but at the same time capable of understanding the action in which the same tool is used, like showing an inability to recognize a cup, but an ability to recognize the act of drinking with a cup.[125]

Humphreys *et al.* proposed a model with the stored structural description of objects, semantic (functional) representation, name representation, and categories.[126] It is cascade processing, such that, in naming pictures, there are three steps: Access to stored structural knowledge about objects, to semantic knowledge, and to stored names. Moreover, there is no sharp break but rather the effects of previous forms of procesing are passed to the subsequent ones. The three-step process postulates

[123] [MARTIN *et al.* 1996, MARTIN *et al.* 2000].
[124] [FARAH/MCCLELLAND 1991] [FARAH 1994]. See also [GONNERMAN *et al.* 1997].
[125] [*MCCARTHY/WARRINGTON* 1990, p. 38]. [126] [HUMPHREYS *et al.* 1988].

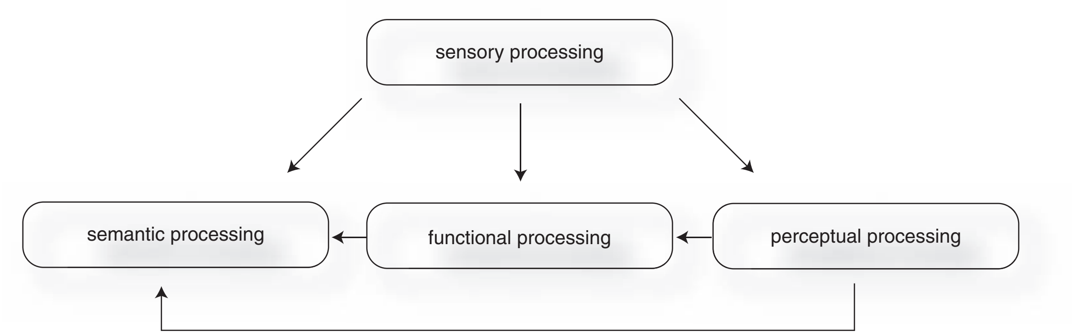

Fig. 4.22 Summary of the different levels and forms of information-processing.

that activation can be passed onto one stage before processing at an earlier stage is completed.[127] For structurally similar objects, functional and associative information is derived quickly, whereas there is an increased competition between category exemplars for individual identification. For structurally dissimilar objects, activation of functional and associative information will be slower but individual identification more efficient.

A way to summarize all the previous results is shown in Fig. 4.22, a reformulation of the scheme proposed by McCarthy and Warrington.[128] Perceptual processing has to do with structural properties (shape) of the objects. I also recall that the perception of functionality is also related to visuomotor representations.[129] This is the subject of the next chapter. The results that I shall present and the scheme shown above can be perfectly integrated, and this integration will be one of the tasks of the next part of the book.

4.6 Concluding Remarks

In this chapter several issues about vision, as an example of perception, have been raised:

- While sensation is the pure transduction of stimuli, any perception goes together with some expectancy.
- The first principle ruling perception is: Any perception is maintained and becomes a habit until it is not contradicted by some stimulus.
- The second principle ruling perception is: All that can be merged is merged and nothing is separated except what must be.
- Visual perception shows a characteristic complementarity between global sensitivity to light and local spatial resolution.
- Any perception is constituted by individuation and identification. In visual perception these aspects are called segregation and visual discrimination, respectively.
- This determines a fundamental complementarity between two paths in processing visual information: The dorsal where-path, leading to the parietal cortex, and the ventral what-path, leading to the inferotemporal cortex.

[127][HUMPHREYS *et al.* 1995, HUMPHREYS/FORDE 2001]. See also [*EYSENCK/KEANE* 2000, pp. 101–4].
[128][*MCCARTHY/WARRINGTON* 1990, p. 43]. See also [JOHNSON-FREY 2004]. [129][HODGES *et al.* 1999].

- Strong modularity has been shown that is not the correct explanation, especially on the higher levels of information integration where we no longer have a topographic organization of visual information. Here, there is no linear organization of information-processing.
- Object perception is actually a construction of the brain (and the mind). However, it is not arbitrary, since it finally represents a guess about external structures and configurations of events that are not immediately perceptible. In this sense, it is rather a reconstruction.
- When objects are perceived, we should distinguish a double complementarity: Between invariant holism and localism and between individuated composites and wholes.
- According to this distinction, tools (displaying different functionalities) are individuated composites while living beings are individuated wholes.
- Faces are a specific kind of individuated whole, namely not species-specific individuated but individual wholes. Here, both general and individual aspects are at play.
- Several impairments in vision suggest that visual perception is a three-level hierarchy, starting with perceptual (structural) processing, going on to functional processing (the object of the second part of the book), and ending with semantic processing (the object of the third part of the book).

The complexity of vision and the huge amount of problems related to issues like reference, categorization, and conceptualization show that a pure information-processing view is not completely satisfactory when dealing with such kinds of problems.

5

Dealing with Target Motion and Our Own Movement

One of the traditional limitations of both cognitivism and the neural-network approach [Sec. 3.8] is the absence of a link with motor features. In other words, these models are characterized by a strong representational and reproductive style and do not sufficiently take into account the fact that a central feature of cognition is not only the perception of motion but also an *action* on the environment. It is not by chance that the field of motor cognition is a relatively late development of cognitive neuroscience. This chapter is devoted to three main subjects: Motion perception, visually guided actions, and movement production. We shall see that many of the results of this chapter indicate the necessity of going beyond a pure information-acquisition model of the brain towards more sophisticated ways of dealing with information, which will be the object of the second part of the book.

5.1 Visual Motion Perception

5.1.1 To "See" as an Active Exploration of the World

The first aspect I would like to consider here is the individuation of a target's motion. There are different and parallel visuomotor pathways involved, with no unified representation of the external world [Secs. 4.3–4.4]. As we have seen, one of the first attempts at dealing with this problem was the distinction between where and what visual paths, i.e. between the spatial localization and the identification of objects (shape, color and so on).[1] Let me recall that this is supported by studies on vision impairments. For instance, a patient suffering bilateral posterior brain damage exhibited disturbance of motion perception in a rather pure form.[2] She could only discriminate between a stationary and a moving target in the periphery of her otherwise intact visual fields. She also had some perception in the central part of the field, provided that the target moved at a velocity not exceeding $10^{\circ}\mathrm{s}^{-1}$. Also her visually guided eye and finger movements were impaired, while the motion perception through acoustic and tactile cues was not impaired. The impairment was due to damage of the lateral temporal-occipital cortex and supports the view that motion's visual perception is a separated function.

Recent studies have enlarged this point of view and suggested that vision's main aim is not to statistically represent the world but rather to act on and interact with it. In animals like a water beetle (*Dytiscidæ*) or a frog it is evident that the behavior (output) is visually guided, which is far

[1] [TREVARTHEN 1968] [SCHNEIDER 1969]. [2] [ZIHL *et al.* 1983].

more important than the elaboration of inputs for representational purposes only, for instance when memorizing—recall that the behaviorist theory in particular was centered around input instead of output.[3] The same considerations remains true for mammals and even primates, although the latter have developed a more complex cognitive system for object identification [Subsec. 4.4.6].

We can go even further and say that an internal image of the world can only be carried out through movement,[4] so that motion is crucial even for identification and visual representation in general. As a matter of fact, plants by definition do not move or see, but rather use electromagnetic radiation for pure energetic–entropic goals and not for acquiring information from the external world. The whole of visual perception is influenced by motion perception and proprioception of a being's own motion: The gaze represents an active exploration of the world, a form of grasping.

5.1.2 Perception of Objects' Motion

One of the first findings of Goodale and coworkers[5] was that visual feedback about shape and the precise relative position of a target are not necessary for visually driven corrections to occur and that the mechanisms that maintain the apparent stability of a *moving* target in space are dissociable from those that mediate the visuomotor output directed at the representation of the target. Later on[6] this dichotomy was stated in terms of a dissociation between the perceptual report of a visual stimulus and the ability to direct spatially accurate movements toward that stimulus (patients with damage to one or the other systems were studied) [Subsec. 4.5.1].[7]

Saccadic eye movements are typically completed while the hand is still moving to grasp an object.[8] A saccade is a rapid movement—of the order of less than a tenth of a second—of both eyes occurring in coordination. When a second saccade, called a correction saccade—since it brings the target right onto the fovea—follows, the action can be corrected when the object is displaced (up to a certain range). No additional time is required for displaced-target trials. As such, the subject can correctly grasp the displaced object, though at no time is the subject able to perceptually realize that the object jumped to the new location. If the subject tries to deliberately follow the trajectory of a displaced object, the movements are slower and fall well outside of the amplitude–duration curve. In other words, adjustments in the trajectory are a fine-tuning of the visuomotor system independently of explicit representation. Representations have more to do with the identification of stable objects and their enduring characters so that they can be recognized when encountered again in different visual contexts, and therefore independently of the point of view. Instead, the visuomotor system deals with the exact *position* or displacement of the target (localization or individuation). This is confirmed by the use of a prismatic lens (shift of 3° of the visual array to the left or to the right). After a certain habituation time, the subject can perform the actions in the same way as there were the lens.[9]

An important aspect of motion perception is when we perceive the motion of other organisms. The perception of organisms' motion is subjected to specific and probably hard-wired constraints that constitute, as we shall see, a true biological module[10]: Trajectories must remain compatible with the biomechanic constraint of the organism; the velocity profile is very different relative to physical objects, showing a fast acceleration followed by a much longer deceleration; there are specific rules stemming either from involved goals or from some other requirements. It is possible,

[3] [*MILNER/GOODALE* 1995, pp. 1–24]. [4] [*LLINÁS* 2001, pp. 58–9]. [5] [GOODALE *et al.* 1986].
[6] [GOODALE *et al.* 1991]. [7] See also [AGLIOTI *et al.* 1995]. [8] [*MILNER/GOODALE* 1995, pp. 156–78].
[9] [JAKOBSON/GOODALE 1989]. [10] [*JEANNEROD* 2006, pp. 103–5].

as we shall see, that the pure perception of motion triggers a motor representation in the observer of how this movement must be performed.

5.1.3 Guessing Dynamics

We cannot directly perceive velocity since any image is fixed upon our retina and only our eyes move.[11] The velocity is reconstructed by comparing different positions of the object divided by time. In many cases, neurons maintain sensory elements for a certain time [Subsec. 3.7.2] in order to allow predictions for the future. In order to acknowledge absolute movements, the body uses inertial cues. To localize an object means to self-represent the movements that would be necessary to reach it. A motionless being (like most plants) would never have acquired the notion of space because, not being capable to correct, by its own movements, the effects of the change of the external objects' position, it would have had no reason to distinguish them from its own internal state changes.

An important problem arises due to nonlinear effects. According to Slotine and Li,[12] a mixture of variables (like speed) and their derivatives (like acceleration) are used by the nervous system in order to make nonlinear problems linear. As we have seen, there are different channels of object perception. The problem is how they are combined when the objects move very quickly. A study[13] suggests that the visual system extracts certain conspicuous image features based on luminance contrast[14] [Sec. 4.2], and that the signals derived from these are then attributed to other features of the object, a process that is called *motion capture*.[15] When either illusory contours or random-dot patterns are moved into the vicinity of a color border, the latter seems also to move in the same direction even though it is physically stationary. The perception of a moving complex object, such as a leopard, is not given by a complicated computational calculus in which each single spot is followed in its movement [see also Subsec. 4.4.2], but is more simply due to motion capture: If the shape of the leopard is perceived as moving, then so are all spots. In other words, vision discards the information about the motion of single spots [see also Subsec. 4.4.4]. Obviously, this implies that one cannot see particulars but this is a small price to pay if one needs to run away from a leopard as soon as possible.[16]

5.1.4 Spatial Perception and Motion

We have seen that there is a double visual dichotomy:

- Between perception of form and shape on the one hand and motion on the other [Subsec. 5.1.2],
- Between perception of form and shape on the one hand and spatial location on the other [Sec. 4.3].

We must account for this situation. We have distinguished between discrimination and segregation [Sec. 4.1]. Conceptually speaking, segregation implies two aspects: We need to locate an object and we also need to perceive its displacement if it is in motion. This is especially relevant when we wish to individuate or segregate animals. In other words, I am suggesting that localization and individuation (or segregation) are not exactly the same since the latter also involves *motion* perception. In light of this, there seems to be sufficient grounds for considering the visual system as having two large subsystems, one for object segregation, which is subdivided into perception of

[11] On this subject see the interesting monograph by Berthoz [*BERTHOZ* 2000].

[12] [*SLOTINE/LI* 1991]. [13] [RAMACHANDRAN 1987].

[14] Luminance is the photometric measure of the density of luminous intensity in a given direction that describes the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle.

[15] [CHURCHLAND *et al.* 1994]. [16] See also [RAMACHANDRAN 1990].

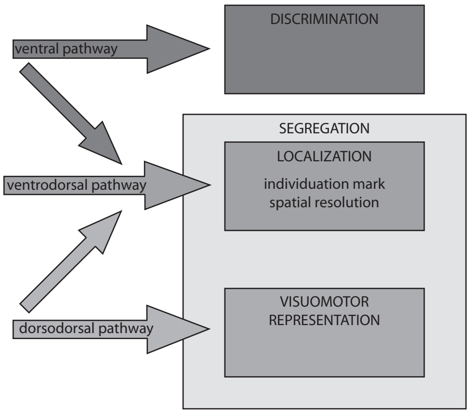

Fig. 5.1 The three vision pathways.

the object's motion and object localization, and the other for object discrimination. This double distinction to a certain extent is supported by studies[17] which indicate that we may probably distinguish between [Fig. 5.1]:

- A ventral stream or pathway (for visual perception of what),
- A dorsodorsal stream (for visuomotor transformations and representations), and
- A dorsoventral stream (for visual perception of where).

Therefore, the latter system also receives some relevant information from the ventral pathway.[18] Indeed, in order to spatially localize an object we also need to know something about it. However, in this case, the quality that we use for localizing the object is not used as a property (as a part of the form, for instance) but as a *mark* (for instance to wear a black hat in a crowded place) in order to individuate the object [Sec. 4.1]. As we shall see in the second part of the book, the concept of marking is very important when dealing with this class of problems.

5.2 Visuomotor Representations

We have seen that dealing with motion has not only a passive aspect (the perception of objects' motion) but it has also an active aspect, i.e. the ability to visually guide our own movement (for instance, prehension acts). We have already considered how important active motor features are when we deal with the functionality of objects [Subsecs. 4.4.4–4.4.6 and Sec. 4.5]. M. Jeannerod calls this aspect of vision *visuomotor representation* since it is somehow the bridge between the visual system and the motor system[19] [Subsec. 3.7.2]. When speaking of visuomotor representations, it is important to distinguish between *determining the position*, which is ultimately due to individuating objects in the environment that can act as reference marks (it is more allocentric), and *determining the direction* of our own motion, which makes use of gradients, far away objects helping for heading [Fig. 5.2].

[17][RIZZOLATTI/GALLESE 2006]. See also [CAREY *et al.* 1998].
[18]Recent studies seem to call such a classification into discussion [PISELLA ***et al.*** 2006].
[19][*JACOB/JEANNEROD* 2003, pp. 64–71].

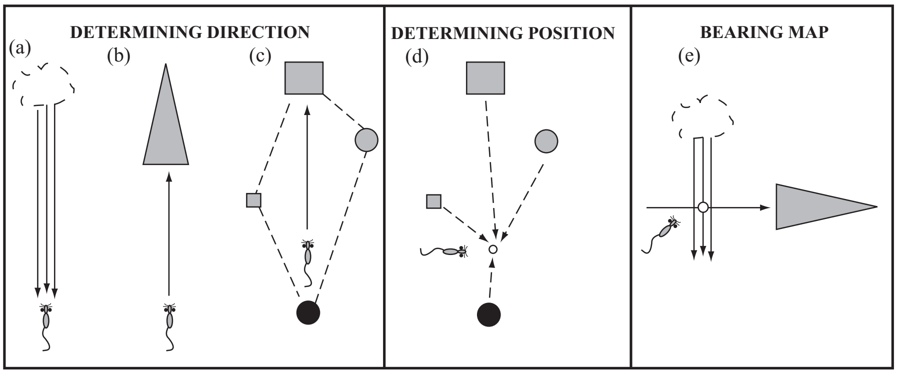

Fig. 5.2 Determining the direction (a–c) and determining the position (d–e). (a) a field of gradient intensity; (b) a distal landmark for heading (too far away to provide positional information); (c) direction deduced from the polarization of an array of positional landmarks; (d) a topographic map constructed from positional cues; (e) the building of a rudimentary bearing map by crossing (a) and (b). Obviously, we are dealing here with centrally synthesized map [Subsec. 3.7.1]. Adapted from [JACOBS/SCHENK 2003].

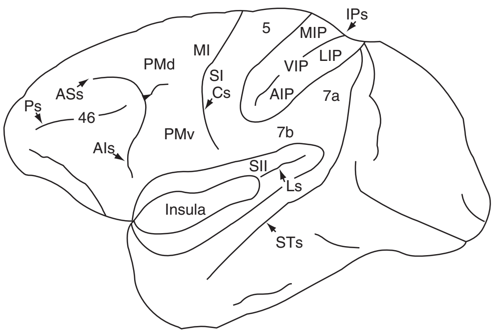

Fig. 5.3 The visual-motor organization of the monkey brain. I limit myself to pointing out some specific elements, the areas 5 and 7, the primary motor system (M1, area 4), the somatosensory area (S1), and the dorsal (PMd) and ventral (PMv) premotor cortex (area 6). A more extensive treatment will be the object of the second part of this book [see in particular Fig. 13.5]. Adapted from [*JACOB/JEANNEROD* 2003, p. 63].

One of the first discoveries in this field was made by Mountcastle *et al.*[20] who found that specific neurons in area 5 and in area 7 fire when an animal did goal-directed and not at-random movements [Fig. 5.3]. Subsequently, it was discovered that neurons in area 7a, in particular, encode visual stimuli for a specific position of the eye in its orbit, that is, they fire when an object stimulates a specific location of the retina and when the gaze is fixed in that direction. This means that information about the change of position of an object relative to the eye is available to the visuomotor component and not to the perceptual component of the visual system. Indeed, eye and hand movement exhibits automatic features and presents a specific motor representation of space

[20][MOUNTCASTLE *et al.* 1975].

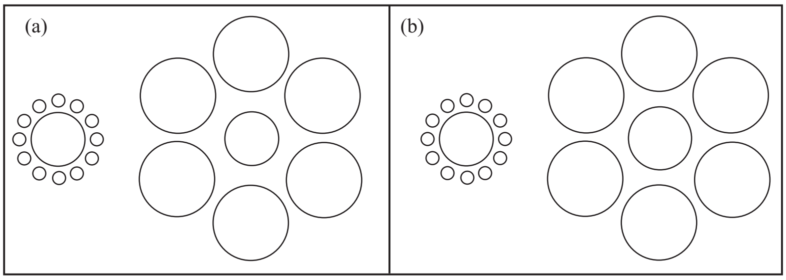

Fig. 5.4 Titchener's illusion: (a): The two circles in the middle are equal but seem to have different sizes according to the format of the annulus around it. (b): The two circles in the middle have different size (the right one is bigger) but seem to have the same format.

that is kept intact in patients with lesions producing a disturbance in the conscious perception of space.[21] Conscious and cognitive representation of motion can contaminate and override short-lived motor representations but not *vice versa*. A categorical representation of the action goal may indeed prevent the expression of a short-lived sensorimotor representation.

Bridgeman *et al.* found that the accuracy of pointing to a target was independent of the fact that its displacement was actually perceived, and nor did these failures in perceiving diminish the ability to point.[22] Moreover, Duhamel *et al.*[23] have shown that neurons in the area LIP fire as a response to a stimulus that is presented not in their receptive fields but within the area of space where their receptive field will project after the eye movement is made (80 msec before the movement starts). It is an anticipatory remapping of the representation of space for the guidance of the arm to the target. Parietal areas like that directly project onto the premotor cortex (area 6).

Experiments with illusions like the Titchener circles[24] [Fig. 5.4] and other ones with objects of the same weights but different sizes and *vice versa*[25] show that the subject often follows internal algorithms during image-distance or weight–size evaluation rather than relying on direct perceptual inputs. Indeed, these kinds of illusions have no effect when the subject performs unconscious actions.

Visuomotor representations are tightly connected with the dorsal treatment of visual information and much more loosely connected with ventral treatment [Subsec. 5.1.4]. Indeed, when a motor response (an arm movement, for instance) must follow a change of color of the target, this requires at least 80 msec more than a response due to a change in the location (happening in 100 msec or less).[26] This also shows a certain difference in the nature of visuomotor representations relative to perceptual ones: They serve as online motor corrections and therefore are particularly appropriate for learning a new skill, while the latter ones prevalently contribute to the memory buffer [Subsec. 5.1.2]. Although we cannot say that memory has no influence on motor behavior, it remains true that when we perform an action, we must compute the instantaneous egocentric coordinates and local orientation of the target object, which in turn implies that we cannot rely

[21] [ROSSETTI 1998].
[22] [BRIDGEMAN *et al.* 1979, BRIDGEMAN *et al.* 1981, BRIDGEMAN *et al.* 1997]. See also [*JACOB/JEANNEROD* 2003, p. 106].
[23] [DUHAMEL *et al.* 1992].
[24] [*TITCHENER* 1909]. Titchener was a supporter of the mentalist approach to psychology.
[25] [GORDON *et al.* 1991a, GORDON *et al.* 1991b].
[26] [*JACOB/JEANNEROD* 2003, p. 108].

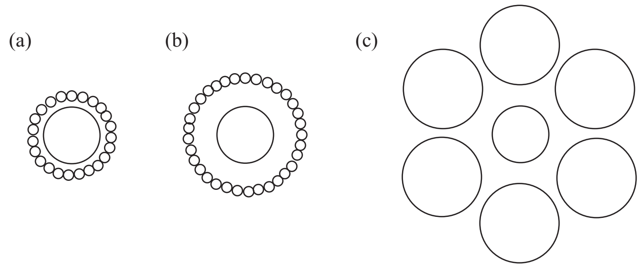

Fig. 5.5 Titchener's illusion revisited by recent experiments of Haffenden *et al.* (a) A disk surrounded by a small annulus at a small distance. (b) Same disk surrounded by a small annulus (of the same size as the former) at a larger distance. (c) Same disk surrounded by a large annulus at a large distance.

on memory because the precise position of the object with respect to our own body coordinates can vary from one occasion to the next.

To sum up, the existence of motor representations shows that we can perceive motion in two different ways[27]: As a pure perceptual event or as a motor execution. In the latter case, only pragmatic components count while in the former emotional elements are also involved.

Recent experiments performed by Haffenden *et al.*[28] seem to put these results into discussion. They used the so-called Titchener illusion but with some adjustments [Fig. 5.5]. The main result was that visuomotor representations can misrepresent aspects of the visual display. Indeed, while visual perception is more sensitive to the size of the middle circle and less able to estimate the distance, grasping is more sensitive to the distance between the disk and the annulus: Grasping is influenced by the size of the gap between disk and annulus whereas perception is not. These data, however, are not completely incompatible with a dualistic (or trialistic) model of vision. Moreover, it has been pointed out[29] that the comparison between size and distance in this experiment is not complete, since the fourth possibility has not been explored, namely a central disk of the same size as in the other three contexts and surrounded by a big annulus but located near the disk.

5.3 Movement Generation and Control

Later on we shall generalize the results of this section in a wider theory of action planning and execution. Here, I shall focus mainly on the typical features of brain activity when executing actions. The philosophers E. Anscombe and J. Searle have pointed out that, when dealing with human actions, the mind–world relation is reversed relative to ordinary representations,[30] in particular the direction of fit is mind-to-world while the direction of representation is world-to-mind. To a certain extent this is also true for other organisms and will bring us further than the problem of information acquiring (a process going always from the world to the perceiving agent).

5.3.1 Structural and Functional Considerations

In humans, we distinguish three types of movement[31]: Reflexive, rhythmic, and voluntary. In the case of animals as evolved as mammals we can speak of exploratory movements instead of voluntary

[27][*JEANNEROD* 2006, pp. 116–18]. [28][HAFFENDEN/GOODALE 2000][HAFFENDEN *et al.* 2001].
[29][*JACOB/JEANNEROD* 2003, pp. 125–9]. [30][*ANSCOMBE* 1957] [*SEARLE* 1983].
[31][GHEZ/KRAKAUER 2000].

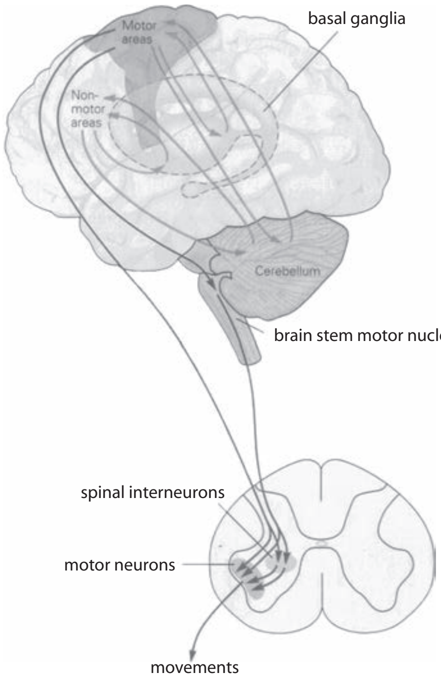

Fig. 5.6 The cerebral hierarchical organization of the motor system. The motor cerebral system may influence the spinal cord along two different pathways, either directly or through the brain stem. All these three subsystems receive sensory inputs and are under the influence of the basal ganglia and the cerebellum. Recall that these two latter systems act on the cerebral cortex through the thalamus. Adapted from [GHEZ/KRAKAUER 2000].

ones (as we shall see, in this case we have goal-directed movements that are not immediately induced by the environment and are neither reflexive nor rhythmic).

1. *Reflexive* movements are elicited by external stimuli. They are still executed even if the brain is cut from the spinal cord [Subsec. 3.3.3]. However, reflexive movements may also be modulated by higher functions. In general, any sufficiently coordinated reaction supposes the contribution of the brain.[32] Reaction time varies with the amount of information processed, for instance, a much longer time is needed when there are choice effects.
2. *Rhythmic* movements involve both the spinal cord and the brain stem. Though they may occur spontaneously, they are mainly activated by external stimuli.
3. *Exploratory* movements as well as voluntary ones can be partially triggered by external environmental conditions, but are initiated according to some internal goal.

The nervous system both in humans and in higher animals (at least vertebrates) learns how to anticipate obstacles and correct the movement. The anticipation part is the feedforward control; the error correction constitutes the servocontrol (a feedback loop) [Subsec. 3.2.2].

The motor system is then a three-level hierarchical system [Fig. 5.6; Subsec. 3.4.1]:

[32]As remarked already in [*ROMANES* 1884, pp. 27–8].

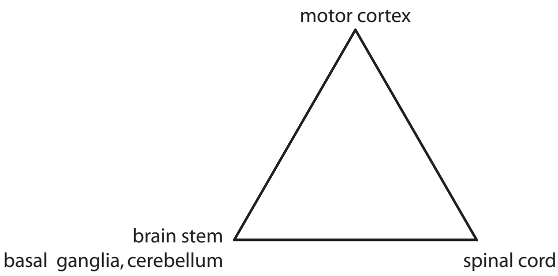

Fig. 5.7 The functional organization of the motor system. The motor cortex is the highest level of the command or decisional system for goal-directed behaviors and modulation of reflex and rhythmic movements. The brain stem, together with basal ganglia and the cerebellum, is responsible for feedback and modulation, while the spinal cord is the executive, being directly connected with the motor output.

- The reflex and rhythmic movements or locomotion generation are mostly dependent on the lowest level, the *spinal cord*. However, as I have stressed, these movements may also be modulated by higher-level subsystems. All motor commands, independently from their origin, converge into motor neurons [Subsec. 3.3.2] in order to be translated into actions through skeletal muscles. The motor neurons of the spinal cord are called motoneurons.
- The *brain stem* consists of two subsystems, the medial and the lateral, which receive inputs from the cerebral cortex and the subcortical nuclei, and project to the spinal cord. The medial descending system contributes to the control of posture, while the lateral descending system is important for goal-directed movement, especially of arm and hand.
- The *motor cortex* is the highest level, upon which exploratory behavior or fine motor skills like speech and hand–finger control depend.

The cerebellum and basal ganglia provide for feedback circuits by receiving input information from various areas of the cortex and projecting to the motor cortex via the thalamus, and in this way allow for comparing between command and afferent information as well as for error correction. They also act directly on projection neurons in the brain stem but do not significantly act on the spinal cord. We may functionally summarize this scheme as in Fig. 5.7.

Let us have a closer look at the cerebellum. The cerebellum represents 10% of the brain's volume but contains half of its neurons. It has three parts[33]: An external gray matter, an internal white matter, and three pairs of deep nuclei (the fastigial, the interposed, and the dentate). With one exception, the cerebellum's outputs originate in the nuclei. The neurons in the cerebellum cortex are organized in three layers. It also has three functionally distinct regions: (1) An internal part called the vermis, (2) two intermediate regions, and (3) two lateral regions. The vermis and intermediate regions correspond to motor execution and project to descending systems, while the lateral regions correspond to motor planning and therefore project to the motor and premotor cortices. From a functional point of view, Snider and Llinás have shown that the cerebellum is crucial for motion perception and execution. In particular it rules tactile activity (hands for humans and noses for rats). The maps [Subsec. 3.7.1 and Sec. 5.2] it produces and delivers to motor and premotor areas are patchworked, due to the need for rapid integration of information coming from different sense sources.[34]

[33][GHEZ/THACH 2000]. [HOUK/MUGNAINI 1999]. [34][BOWER/PARSONS 2003].

978–0–19–960848–5 05-Auletta-05-drv Auletta (Typeset by SPi) 144 of 862 June 1, 2011 17:19

OUP UNCORRECTED PROOF – REVISES, 1/6/2011, SPi

In the most elementary case,[35] motion is effected by a group of interneurons called the central pattern generator network. They are able to activate a set of motor neurons and receiving afferent information from sensory neurons. This ability is conditional on certain external or internal events, like perturbation, stimulus, goal-directed behavior, and so on.

5.3.2 Goal-Directed Movements

As we have seen, both motor and perceptual representation may either present or not present conceptual features. Any movement is a dynamic process such that there are several physical consequences, since inertial and centripetal forces are generated that contribute to movements while they are happening and constitute a continuously changing force field. Due to the complexity of these dynamic aspects, an organism has to decrease the degrees of freedom involved in order to control its movements.[36] This is obtained by means of the geometrical organization of the skeleton, by automatizing many movements and making many movements interdependent. However, the most important aspect is the ability of the organism to group several motor segments under a common denominator. Let us consider this point.

This field of study was very much improved by the insights of K. Lashley and N. Bernstein.[37] Lashley had observed that during a virtuous execution of music, finger alternation could attain a frequency of 16 strokes/sec, which exceeded the possibility of any sensory feedback influencing the decisional system. This means that the movements cannot depend on peripheral stimuli. In order to explain this, the notion of modular motor programs (or engrams) was later developed. However, the notion of fixed engrams seems to clash with the observed plasticity of movements. N. Bernstein[38] started by understanding that goal-directed movements cannot reflect a one-to-one relationship between the specific patterns of motor neuron firing, or indeed the forces generated by muscle contraction, and the actual movement produced. Also D. Hebb[39] pointed out that different motor actions share important general characteristics, even if performed in different ways.

Bernstein refined previous ideas by pointing out that single specific motor acts may be embedded in more complex ones in a serial and modular sense [Sec. 3.6]. In this way, the different parts can be assembled moment by moment when a certain need or stimulus would be present giving rise to a class of similar structural patterns that we call movements. In this way, the whole movement is plastic while the subcomponents can be hardwired or be ruled by engrams. As we shall see, in order to fully understand this structure we need to consider a movement as an equivalence class of chains of motor segments able to obtain the same result. In other words, several individual motor segments are (up to certain limits) appropriate to be combined in the pattern (for generating a certain result), therefore giving rise to different possible chains that can be considered as instances of the same movement, therefore constituting an equivalence class (individuating an abstract movement). For instance, there are several alternative motor paths for picking up a cup and they can be considered as equivalent if leading to the desired outcome, e.g. to drink some tea. Different motor performances able to lead to the same result are then said to be *motor equivalent*, which is therefore a functional equivalence. As we shall see in the next part of the book, the concept of functional equivalence is fundamental for biological systems. The issue of motor equivalence goes even further than that. Bernstein also acknowledged that the relationship between movement and the innervational impulse that evokes was not univocal. A given impulse can produce completely different effects under different conditions, since the response will depend

[35] [GRILLNER 1999]. [36] [*BERTHOZ* 2000]. [37] [*JEANNEROD* 2006, pp. 7–12].
[38] [*BERNSTEIN* 1967]. See also [*JEANNEROD* 1988, JEANNEROD 1999a, pp. 27–8].
[39] [*HEBB* 1949, pp. 145 and 153–7]. Actually he draws his ideas from K. Lashley.

on the initial position of the body and on the external force field in which the movement develops. This also establishes a further equivalence, that between action execution and action simulation or imagination [Subsec. 3.7.2].

The interesting point is how the organism reacts when there is an unexpected event. According to Bernstein,[40] when action encounters surprise, it is either impossible or irrelevant to reestablish the initial plan of action. In other words, action must be reorganized according to unforeseen events. In this way, a motor equivalence is also somehow established between the original movement and the corrected one. *Proprioception*, the sense of the different parts of the body, is involved in this kind of correction: For Bernstein, this is not reflex-triggering but rather contributed to the central representation of movement. Therefore, this conception stands half-way between the centralist and the peripheralist theories of motor programming.[41] For instance, reaching movements cannot be carried out as a series of sequential steps: They present a *hierarchical structure*,[42] which also means that the neural commands forwarded to muscle groups may be generated in parallel even though the overt movement appears to be sequential.

Summing up,

- Goal-directed movements are neither single motor performances (like a reflex), nor chains of details (that is, of individual performances), but hierarchical structures which are differentiated into schematic or prototypical details.
- Moreover, the goal-directed forms of movement are dynamic patterns: They develop and involute, and the interactions with the environment contribute to define the patterns.

5.3.3 The Motor System

When speaking of motion control we should also distinguish between informational control parameters and biomechanical variables.[43] *Control* parameters are specified by the CNS independently from current external conditions. The reason is the following: Although theoretically speaking it is always possible to rearrange variables, in the case of motor control this would mean a rearrangement of the causes of motion and their effects, in which case the controlling instance (the brain) would be controlled by the system that must be kept under control (the environment). This is why any control cannot be mechanical (which is always bottom-up) and must be performed in informational terms. Moreover, a pure mechanical explanation of motor control cannot work, since we cannot distinguish, in terms of pure muscle torque, between the two positions a and b shown in Fig. 5.8(a). The same is true for electromyographic models, as shown in Fig. 5.8(b). Therefore, a component of change in the membrane potential produced by descending systems (panels (c) and (d)) may be independent of muscle afferent feedbacks and be produced by central control.

Therefore, goal-directed movements presuppose processes, such as monitoring of the initial state and specifications of the commands that have to occur before the movement is generated, and some others that are a consequence of the reafference of the movement itself.[44] Reafference denotes then the consequences of self-produced movement on the sensory inflow. The problem is to build basic space structures that can apply equally well to both the body-centered frame of reference and the central representation of space, which assumes an allocentric, stable environmental frame of reference,[45] as displayed in Fig. 5.2. Let us consider this point a little.

[40][*BERNSTEIN* 1967]. [41][*JEANNEROD* 2009, pp. 48–67].
[42][JEANNEROD 1991]. See also [BIZZI/MUSSA-IVALDI 2000]. [43][FELDMAN/LEVIN 1995].
[44][*JEANNEROD* 1988]. [45]See also [BERTHOZ 1991].

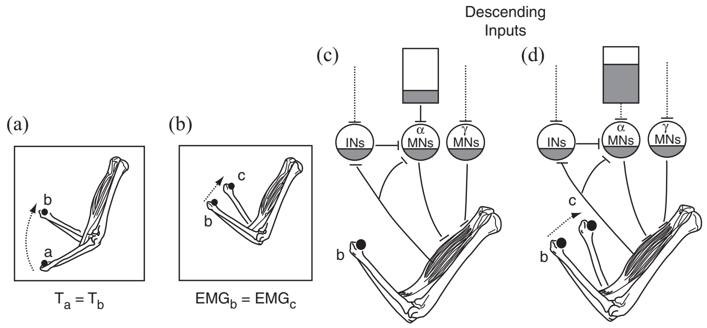

Fig. 5.8 Mechanical explanation of motor control and control variables. T stands for muscle torque, MN for motor neurons, IN for interneuron. Adapted from [FELDMAN/LEVIN 1995].

The interaction between subject and goal implies that detailed programs of any movements features (as, for example, trajectory shaping) are superfluous.[46] For a hand to be transported to a visual target located outside the body, its position with respect to the other body parts must be represented. This is the function of the *proprioceptive map* (the cerebellum is the center of the proprioceptive relation to the world), which is distinct from the *visuomotor map* or the representations. The visual and proprioceptive maps jointly project to another part of the hierarchical structure, where the *goal* of the action of reaching is defined by the decisional system. Here other forms of information are also integrated depending on the specific aims related to the manipulation of the object. The goal level is the area where the target loses its quality of mere visual stimulus and becomes represented as a goal *for* an action (we shall see in the second part of the book the saliency of this point). Several execution levels then depend on the goal level.

Bernstein developed a system-theory understanding of motion [Fig. 5.9], in which the brain is able to control muscle centers by comparing decisions taken by a decisional system using a sensory feedback, and therefore correcting the movement through a regulating system. This model can be reformulated as in Fig. 5.10, in which the brain's three facing an external environment systems are the sensory, regulatory–modulatory, and decisional systems. This is in accordance with the previous scheme shown in Fig. 5.7.

A target position coded in an allocentric frame of reference gradually influences visuomotor transformations when the efficiency of the egocentric frame of reference (necessary for movement execution) correspondingly decreases. To this purpose, a study[47] has made use of the Müller–Lyer illusion and has shown that test subjects pointed to a target whose position could be erroneously localized because of an illusionary effect [Fig. 5.11].

In schizophrenic patients the delusion of control consists in the lack of awareness of certain aspects of motor control, which in turn is a consequence of a failure in the mechanism by which the predicted consequences of an action are derived from a forward model based on the intended sequence of motor commands.[48] The reason is that the patient is unable to attenuate responses to sensations of limb movements (an overactivity in the parietal cortex is in fact observed) which arises from a long-range corticocortical disconnection which prevents inhibitory signals arising in the frontal areas that generate motor commands reaching the appropriate sensory areas.

[46][*JEANNEROD* 1988, pp. 1–40]. [47][GENTILUCCI *et al.* 1996]. [48][FRITH *et al.* 2000].

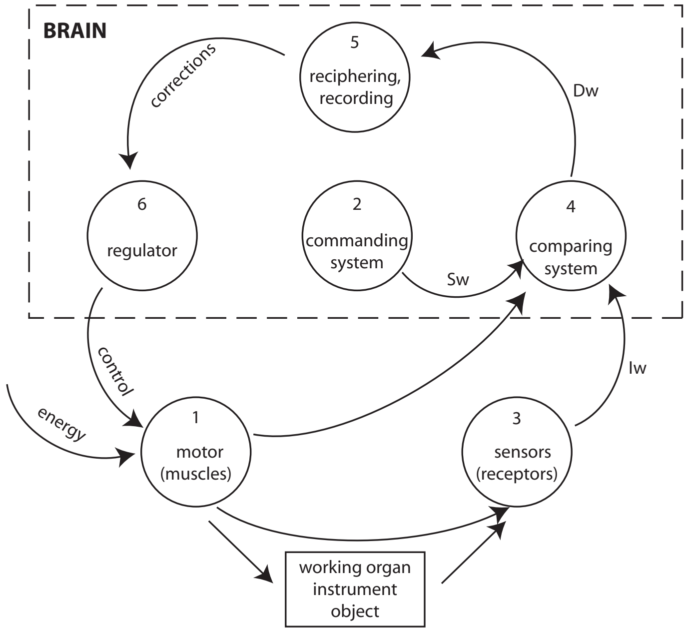

Fig. 5.9 Bernstein's comparator model. The information (Sw) generated by the commanding system is compared with reafferent input (Iw) resulting from movement execution. The output (Dw) of the comparator determines the global output of the system. Adapted from [JEANNEROD 1999a, p. 59].

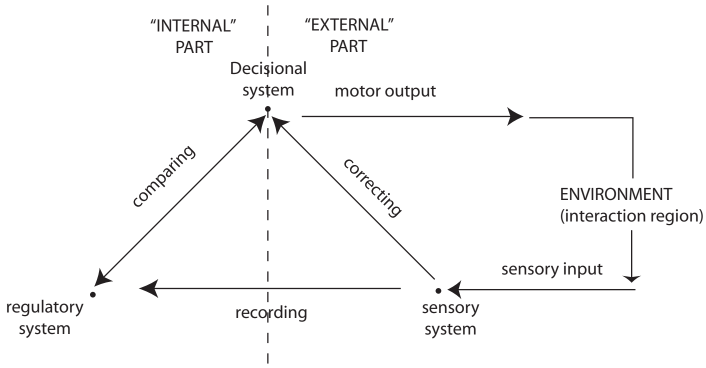

Fig. 5.10 Bernstein's model can be schematically simplified as follows: The sensory system produces allocentric maps. The reafferent information from the sensory system (visuomotor representations) goes to the modulatory (regulatory) system in order to be stored, and to the decisional system (the command system in Bernstein's terminology), where goals are established, for quick correction of motor outputs. Recall that the cerebellum is able to produce proprioceptive maps, which are also sent to the regulatory and decisional systems. This information will be compared with the previous outputs in a feedback loop between the decisional and regulatory system. Then, the decisional system finally sends a signal for new motor output (in view of the goal).

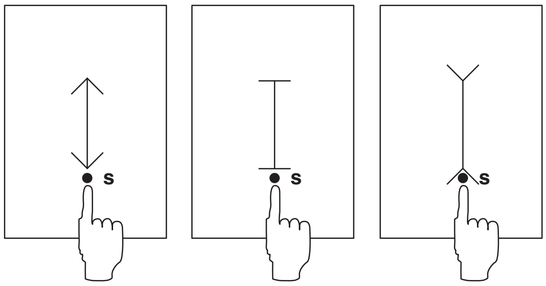

Fig. 5.11 The line between the arrowheads in the left diagram appears shorter than the line in the right diagram (Müller–Lyer illusion). In spite of this illusional difference, the subject can perform pointing correctly. S indicates the start position. Adapted from [GENTILUCCI *et al.* 1996].

5.3.4 Movement and Representationalism

In artificial intelligence and especially in robotics, there has been, in the last 20 years, a radical rejection of representationalism. Indeed, traditionally navigation in external space was thought to be performed by using an allocentric map. However, it was observed that going from one place to another frequently involves traversing a network of intermediate points, that is, following a route in which familiar landmarks indicate progressive success and dictate new courses of action.[49] Then, instead of building a system as a successive series of vertical slices representing functional units (from sensors to actuators), Brooks proposed a vertical (motor-like) decomposition of the problem (allowing multiple goals) by considering tasks achieving behaviors with a subsumption architecture and levels of competence. The robot is elastic and robust, and it is equipped with parallel processing. The central idea is to let robots act and interact in a real environment (a proprioceptive approach from the start). The main point here is that an intelligent system operates as such only in a given environment, since it is able to carve out environmental stimuli into useful information and therefore to treat the world as its best reservoir of information. No representations at all are involved here and such a system is intrinsically dynamic.[50]

Mataric[51] has produced robots acting following environmental landmarks and slipping from the behavior specified by a landmark into a behavior specified by the next, rather than travelling from one landmark to another. In this way, patterns of activity can interactively emerge. Here, Mataric follows Brooks' criticism of all-representational cognitivism.

Also, Maes had assumed that, instead of planning (i.e. internal representation of actions, goals, and events), a model of situated automata with emergent functionality is more suitable. This is a result of the interaction with a dynamic environment: Its properties are exploited to serve the functioning of the system. Here there is no predetermination of the action selection.[52] Maes' goal was to build adaptive autonomous agents, that is, systems inhabiting a dynamic and unpredictable environment in which they try to satisfy a set of time-dependent goals or motivations. They are adaptive if they are able to improve their competence in dealing with a task. These systems are built in general in an integrated way rather than in a modular one.

[49] [WILSON 2000]. [50] [BROOKS 1986, BROOKS 1990].
[51] [MATARIC 1992]. See also [*HENDRIKS-JANSEN* 1996, pp. 135–203].
[52] [MAES 1990a, MAES 1994]. See also [ALMÁSSY *et al.* 1998].

Brooks said[53] that cognition is nothing more than the point of view of the observer and that only such a point of view and related action exist. In reality, cognition is a mediation process between perception and action. Indeed. it is extremely problematic to completely eliminate the concept of representation or description, otherwise the organism will be unable to understand the effects of its actions on the environment.[54] Therefore, even if the previous approach has rightly pointed out the central role of the environment and of interaction with the environment in any goal-directed behavior, the fact remains that representational maps play an important role too [Sec. 3.7]. We have indeed seen that motion generation and perception are not completely separable. In fact, a *proprioceptive mode* of processing spatial information coexists with a *representational mode* and both modes generate and store their own mapping of space.

The boundaries of the sensorimotor space are specified by both the perimeter of the receptive field and the action radius of the motor apparatus.[55] The discriminative power of the sensory surface will determine the basic grain of such a spatial structure. On a mathematical plane, a collection of separated points is not sufficient for defining a structure in space.[56] One needs a *geometry* of description, i.e. a rule for describing the potential relationships between these elements (and here *mapping* is very important). Certain metric rules that define a structure in space, called a *path structure*, are of special interest here. These rules determine, in direction and distance, the trajectory to follow in order to move from one point to another. A path structure, superimposed on a collection of separated points, defines the locality of each of these points in a vectorial map. This is particularly suitable for describing sensorimotor space because motor commands are generally prescribed in terms of direction and distance. In other words, at least in the case of perception and generation of motion, localization and directional or vectorial structure, as displayed in Fig. 5.2, are inseparable aspects (but this is probably a general feature of the brain).

This integration of dynamic and representational aspects is also true for the body's maps [Subsec. 3.7.1]. The mapping in the brain of parts of the body is dynamically established and must be continuously maintained through activation.[57] Although we have seen that the brain has maps of the body [Figs. 3.15 and 3.16], it is also important to understand that such maps do not directly link to body parts; rather, such maps are the result of the acquisition of *sensorimotor schemata* during *execution* of temporally coordinated movements. Amputee patients progressively lose the brain areas devoted to the amputated limb, and this process can be slowed down if simulated actions with that limb are imagined. Singer's work has led to such a conclusion.[58]

5.4 Concluding Remarks

The main results of this chapter can be summarized as follows:

- We have refined our previous model and established that we have a ventral path leading to the discrimination of the object, a ventrodorsal path leading to localization, and a dorsodorsal path leading to motion perception. The latter two paths contribute to the segregation of the object.
- Visuomotor representations lead our cognitive interaction with the world.
- The motor system is a three-level hierarchical system consisting of the spinal cord for the purpose of ruling reflex movements; the complex brain stem, basal ganglia, and cerebellum for control of

[53][BROOKS 1991a, BROOKS 1999b].
[54]See also [O'REGAN/NOË 2001], and [REES/FRITH 2001] for criticism.
[55][PAILLARD 1991b]. [56][ARBIB 1981]. [57][*THELEN/SMITH* 1994, pp. 129–60].
[58][SINGER 1999, SINGER 2000].

posture and goal-directed movements; the motor cortex for exploratory behavior and fine motor skills.
- Goal-directed movements are global structures and dynamical patterns. Movements are hierarchically nested.
- Movements establish functional equivalence classes of physical and mechanical actions.
- Mechanical explanations of goal-directed and voluntary movements are wrong. Here, the concept of motor equivalence is very important.
- It is also important to distinguish between a sensory system, a decisional system, and a regulatory (comparing) system.
- The brain develops both proprioceptive and allocentric maps.

Let me now briefly elaborate on the point arrived at the end of this part of the book. We have seen that many aspects of vision and especially of the motor system raise important questions about referential and even semantic issues. There is no way to answer these questions in a pure information-processing framework, or in a pure information-acquiring context, of which information-processing, together with information-sharing and information-selecting, is only an aspect [Subsec. 2.3.2 and Sec. 3.2]. Indeed, information-acquiring is a pure physical process of dealing-with-information that does not presuppose the ability to be able to be referred to the input system or input information: It is a process in which a final selection of information, thanks to coupling, is able to tell (in principle) something about the input information (the initial source of variety). However, nothing here is said about the way in which this information can actually be accessed. For instance, this process can finally result in potential information that is stored in some laboratory device without anybody actually reading it (and without this "reading" it is destined to remain forever as potential information, like an inscription on a stone that is buried deep in desert sand). This somebody (the experimenter, for instance) is instead an agent that is able to understand that the output information is *about* a certain input system and therefore he or she can also undertake certain actions relative to the input system. This seems quite reasonable in the case of human agents. Now, the big hypothesis this book deals with is that prehuman organisms are also able to do this. They indeed permanently monitor, refer to, and deal with the external world through goal-directed actions. To understand this mystery, we need to switch to another perspective, which is the subject of the next part of the book. This perspective is biological and is centered on the concept of *function*, which requires an expansion of the concept of information to *information control*. The concept of functional equivalence that we have introduced here relative to movements will be generalized and shown to be a basic category of life.