SPRINGER BRIEFS IN PHILOSOPHY

Mario Graziano

Epistemology of Decision

Rational Choice, Neuroscience and Biological Approaches

Springer

SpringerBriefs in Philosophy

For further volumes:
http://www.springer.com/series/10082

Mario Graziano

Epistemology of Decision

Rational Choice, Neuroscience and Biological Approaches

Springer

Mario Graziano
Department of Cognitive Sciences,
Educational and Cultural Studies
University of Messina
Messina
Italy

ISSN 2211-4548 ISSN 2211-4556 (electronic)
ISBN 978-94-007-5427-0 ISBN 978-94-007-5428-7 (eBook)
DOI 10.1007/978-94-007-5428-7
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2012946189

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the memory of my father and Valentina

Preface

The story usually goes as follows. On one side there are the supporters of normative rationality, with their typical belief that to make a decision is basically a matter of ordering the preferences according to a rational framework; on the other side there are people who are enthusiastic about psychological findings showing that real individuals are not so "rational" as the opponents believed. In fact, human beings get continuously conditioned by biases and traps for the mind. It seems that human rationality is not so efficient as one can suppose. The controversy between normativists and descriptivists on decision theory is about the very nature of human understanding. How much is "free" or "bounded" the rationality to determine overt behavior? Is the maximization of personal gain the real "rationale" of human understanding? Immanuel Kant was irresistibly fascinated by the idea that rationality is able to autonomously determine the moral behavior. Indeed, in his opinion, a behaviour can be considered as "moral" only if this is the case. According to an old philosophical tradition, human behaviour is more or less "moral and efficient" insofar as it is more or less "rational", i.e., not conditioned by anything else. Emotions, of course, are especially to be avoided. We have to consider a choice as rational if the used means are appropriate to the given end; and, it seems that in balancing means and ends there is no room for emotions and any other non-rational elements. Nowadays, it seems that the scene is changed. We are finally aware of the ecological and embodied character of human understanding. Reason is no more regarded as an emotionally bland and cold thing. Moreover, reasoning itself—not emotion—is affected by a lot of bias which subconsciously drives our thought pathways to many kinds of mistakes.

This book argues that a third way between normative and descriptive accounts of rationality in decision theory is possible. It is matter of a sort of normative rationality with a human face, that is, a naturalistic account of rationality disciplined by the needs of the economic paradigm. This latter involves a certain interest in the way things ought to be. It is the economic perspective itself, even in the case of the "biological economy", which Mario Graziano—following Alfred Marshall—endorses, that implies a normative constraint. Economics, even if inspired by biology, cannot be a purely positive science. In this book the reader can appreciate a naturalistic account on decision theory. The word "naturalism" in philosophical discussions means many things. However, it is uncontroversial that if only natural science can tell us how the world really goes, then there is no space for any kind of normative facts—a crucial point for decision theory. In *this* book, naturalism is inspired by the biology of complex systems and neuroeconomics. One of the most intriguing facets of *this* book is the appeal to the social cognition in order to deal with the main problems of decision theory. This move depends on an epistemological worry, that is, to adopt a theoretical framework compatible with both the interpersonal dimension and a normative evaluation. Moreover, because of its dependency by neuroscience, social cognition provides the naturalistic compatibility which is a major tenet of the book.

The desire of a third way between normativism and descriptionism in decision theory is urged by the field of application of the analysis. Since economic behavior is the core *explanandum* of the book, some amount of normativity is requested. This attitude could be a good example in the usual debate. Should the decision theory be subjected to a normativist or a descriptive account? Perhaps the right answer can be: "It depends". In fact, it depends on the purposes of which our scientific enterprise is engaged. For example, if we are interested in improving certain budget standards in a health care system, then we ought to be also interested in some amount of normativiy. It could be useless to know only why the managers are conditioned by their mental biases. We would like to improve their behaviors. And, of course, "improvement" is a normative concept. Knowledge about the cognitive architecture and the neurophysiological basis of mental biases is a wonderful thing. We can deduce many significant consequences from that knowledge. The improvement of the behaviors in order to fulfill a certain purpose is another kind of question. In *this* book the "it depends-strategy" is guided by the theoretical needs of the economic perspective, but it can be proposed more in general as the right attitude towards the normativism/descriptionism debate in decision theory.

Anyway, in *this* book we can appreciate the possibilities of a naturalistic account on decision theory committed to the normative constraints involved in the economic behavior. A great part of this commitment depends on the appeal to the paradigm of social cognition. This way of reasoning is perhaps similar to what Steven Stich proposed in an essay on the Daniel Dennett's theory of intentional systems ("Dennett on Intentional Systems," *Philosophical Topics*, 12, 1, 1981, pp. 39–62; now in S. Stich, *Collected Papers*, vol. 1, p. 73). "So any object will count as an intentional system if we can usefully predict its behavior by assuming that it will behave rationally. And what is it to behave rationally? Here, Dennett suggests, the full answer must ultimately be provided by a new sort of theory, intentional-system theory, which will provide us with a normative account of rationality. This new theory "is envisaged as a close kin of—and overlapping with—such already existing disciplines as epistemic logic, decision theory and game theory, which are all similarly abstract, normative and couched in intentional language" (D. Dennett, "Three kinds of intentional psychology," in *Reduction, Time, and Reality*, Cambridge University Press, 1981, p. 19).

Pietro Perconti

Acknowledgments

I owe thanks to many people for making this book possible. I would like to thank all my colleagues at the Department of Cognitive Sciences, Educational and Cultural Studies. I wish to express my sincere gratitude to Director Prof. Ninni Pennisi for all the inspiring and motivating conversations. His constructive and thorough criticism is highly appreciated. I must also thank Vivian De La Cruz, Tricia Kress, Ignazio Licata, Franco Lo Piparo, Emanuela Mercuri, Simona Morini, Daniele Schilirò for reading, comments, and suggestions that have substantially improved the quality of the writing. A special word of acknowledgement is due to Lucy Fleet Editor of Springer for editorial comments and suggestions which contributed to the quality of the book.

Finally, I would like to thank Prof. Pietro Perconti, the supervisor of this research, for his support throughout the years, for his preface, as well as for his valuable remarks and comments. Pietro gave me an example of intellectual rigor that I have always tried to imitate and probably never attained.

Contents

Introduction

What does it mean to make a decision? One frequently cited answer is that a decision is the result of what we want and the chances we have of obtaining it: all the alternatives in this regard have true costs and consequences. As part of the social sciences, the scientific study of how we make our decisions, how we can optimize or at least render them satisfactory, and the factors that might influence them is known as decision theory. Diverse and heterogeneous areas of inquiry from various perspectives have been employed to study the decision-making process. Decision-making has been studied by philosophers, economists, psychologists, mathematicians, physicists, biologists, and sociologists, who have attempted for different purposes to shed light on the mechanisms and modalities that lead individuals to make a certain decision rather than another.

As part of the normative approach to the study of decision-making, a basic concept is "rational choice", conceived as the result of a decision-maker's calculation that employs perfectly logical processes of thought. Initially, this concept was the dominant paradigm in economics. In fact, a fundamental premise of neoclassical economics was that economic phenomena were essentially due to the action of fully rational agents, equal and therefore indistinguishable from one other, all of whom were individually pursuing only their own personal gain. This postulate was followed by another epistemological postulate that found in economic phenomena the possibility of applying certain general laws that could be expressed in mathematical terms. Undoubtedly, over the past two decades, this explanatory model has exercised considerable influence beyond the sphere of economics.

The extension of rational choice theory to all the social sciences was initially due to its universal aspirations. In fact, if we hypothesize the same operations and the same intent to all individuals and rationality as manifested in an identical way in all agents at all times, all social phenomena can be explained by the same model. Furthermore, a purely mathematical model allows both the construction of a personal process of deliberation and a description and analysis of social phenomena.

The adoption of this model of explanation is therefore motivated by the deductive and universal productivity that it applies to the simple and rigorous analysis of all social phenomena. In addition to traditional economic challenges, the propagation of this model primarily stems from the fact that rational choice theory has long been considered a scientific and not a purely speculative theory because it was possible to rigorously, precisely, and coherently derive from it many accurate predictions through mathematical description, and these predictions could also be compared with social facts observed.

Therefore, the normative models of choice proved to be excellent tools that were appropriate and useful for analyzing decision behavior in many situations. However, since the 1950s, extensive literature has highlighted numerous theoretical and empirical limitations regarding the analysis of a wide range of real decisions. One of the first scholars to explain the discrepancy between actual behavior and the standard theory of decisions was undoubtedly Simon (1972), who strongly criticized the notion of normative rationality centered on the concept of optimization with emphasis on cognitive and evaluative human limitations and the assumption that individuals act according to a bounded rationality. Beginning from Simon's considerations, many researchers have subsequently sought to emphasize decision-making agents by considering actual human cognition. One of the most famous and important decision-making models (if only because its creator won the Nobel Prize) is the Prospect Theory of Kahneman and Tversky (1979) that, moving from expected utility theory, proposed amendments to introduce explanations of decisions of the changes that can make decisions on behalf of a real individual. The subjects in the experiments tended to choose situations by their own classification (framing), i.e., according to how they perceived them, showed a marked aversion to risk and loss of a certain sum of money (a greater propensity to win the same amount of money) and were even more averse to showing ambiguity and lack of information. Research conducted as part of economic psychology, especially cognitive psychology (also known as Behavioral Decision Theory), has led many economists to note that decision-makers actually depart from the assumed model of rational choice theory. This result emphasizes that certain psychological phenomena (such as how one is mentally presented with a decision problem, how one presents that information, risk aversion, etc.) lead individuals in the act of making choices to commit many more "errors" than expected by normative theory, which suggests an extremely pessimistic view of the subjects' reasoning ability. In fact, regardless of the specific implications of the theories proposed, the profile of the man who emerges from these studies is one whose forms of reasoning are inevitably forced by their very nature to be limited and fallible and have little in common with the image of *Homo Economicus*, a proposal from neoclassical economic theory.

Experimental economists argue their position based on empirical data obtained through the use of laboratory data, which show that agents' reasoning deviates systematically from standard inferences defined by law. However, although we agree with these scholars that "perfect rationality" is an idealization, we attempt to resist the temptation to diminish the importance of our rational standards (including logical and probabilistic reasoning) to fit our limited cognitive abilities. Support shall understand that we humans are "rational animals" because of our ability to reason and certainly not because of the achievement of perfect rationality. In this perspective, then, our rationality is the right use of reason to make choices in the best possible way: to deliberately achieve our best with the available means. In fact, both design standards and the theories of experimental economists do not consider sufficiently worthy the reality that our judgments about a choice's rationality result from highly contextualized assessments, which involve myriad facets and regulatory standards pertaining to a situation, including the limits of cognitive agents. How is admitted even by some psychologists, although the reasoning of the subject is wrong on the basis of a normative theory, this does not prove that they have reasons for their evaluations. This finding implies that agents do not always deliberate intensively about their choices and the alternatives of each decision. The resolution was only a conscious way about the reason for their actions. Many other actions are not the result of a decision but of routine and habits. For example, we do not make specific decisions when leaving the house in the morning to go to work or while driving a car. However, as the vast majority of our habits, i.e., those who remain aware of what are the reasons why we are going to work or why we walk one way rather than another, and normally we would be able to motivate our choices only if we are required to do so.

In addition, it is sometimes entirely rational to act in a seemingly irrational manner. For example, in amorous situations, a suitor sometimes pretends to be indifferent to arouse curiosity, or in a game, a strategist occasionally enacts "stupid mistakes" to check an opponent via unpredictability. These types of behavior seem "irrational" but might be entirely rational in the long run. Therefore, there are good reasons for doing something one should not do, and in certain circumstances, it is "reasonably appropriate". In this broader sense, rationality is not opposed to irrationality and does not logically imply that a human being as a "rational animal" should respect most principles of our best normative theories of rationality. Rather, this sense of rationality highlights the difference between a "biological organism" that is considered plausibly able to meet the standards of rationality and is evaluated on this basis and organisms such as plants that are not. Therefore, based on these considerations, the orthodox approach of decision theory in terms of maximizing expected utility (which, as we will see, provokes particular problems) should not be construed as the arbitrator or the owner of rationality but as the servant of a more fundamental and deep-rooted concept of rational assessment. Similarly, there is no welcome at all costs, the results of psychological experiments only because there is little doubt that a normative theory can explain all spontaneous and naive intuitions that lead agents to make decisions. We attempt to emphasize that although agents in certain cases and contexts might systematically depart from standard models of rationality, this does not necessarily mean that humans cannot be considered rational animals.

Therefore, in this book, we simply reject the need for a principle of separation between descriptive and normative decision theories, and as we move into the economic paradigm, we simultaneously attempt to justify a conception of rationality according to which rational agents are biological organisms in an environment, without advancing a concept of "natural rationality" purged of any normative dimension. In this way, we advance a naturalist position that seeks to defend evaluative and adaptive rationality by rejecting the terms of the "dilemma" that impose a descriptivist rationality. To delineate this position that will not be produced in the course of the chapters some axiomatizations, and neither to defend models of psychological–emotional, but we will defend the supplementary explanatory models that have been advanced in the field of neuroeconomic, game theory, and the biology of complex systems.

The volume is organized as follows. In the first two chapters, I will emphasize how rational action is a central category for the assessment of human reasoning and decision-making. Specifically, in the first chapter, we initially provide a historical reconstruction of the theory of rational choice and describe through a discussion of certain concepts of "game theory", as the dynamics of strategic interaction in the form of cooperation and conflict can influence individual decision-making. We attempt to highlight how the fallacies of the theory of rational choice pursued by experimental economists who use laboratory data are not exempt, in turn, from criticism. At most, recent empirical studies indicate the difficulties that we encounter on a daily basis to achieve our aspirations of rationality. Moreover, it is paradoxical to infer from these limits that humans are not rational given that the empirical research underlying these conclusions assumes precisely the same standard conception for men as for rational animals. In the second chapter, the epistemological framework is outlined including neuroeconomics, a branch of behavioral economics that seeks to investigate the role of the psychological mechanisms of economic analysis without rejecting a neoclassical paradigm, which provides a theoretical framework based on utility maximization and energy balance. In this sense, neuroeconomics has been defined as the state that allows the use of brain processes to find new foundations for the economic theories (Camerer and Loewenstein 2002). To better understand the mechanisms by which the brain assesses and compares the alternatives there may be traceable to determinations of our choices and our behaviors. Within neuroeconomics, we will focus mainly on the studies that have shown certain parallels between the model's utility and dopamine (dopamine is a neurotransmitter, and its reduced presence in some nuclei of the brain is related to Parkinson's disease). We will also, as some models of game theory (Prisoner's Dilemma, Ultimatum Game) have been used by neuroscientists believe that such collaboration can come out and new results are useful to both economists, is the same brain science. However, the hypothesis to be advanced is that to truly understand what triggers the minds of players in a strategy game, rather than relying on laboratory experiments or on the findings of neuroimaging, we must rely on studies of "social cognition". According to these studies, many players' choices (within game theory) prove to be far from a presumed (at least in the opinion of certain experimental economists and neuroscientists) emotionality or irrationality.

In the final chapter, the second part will emphasize that it is inherently implausible that a realistic description of our expertise in psychological reasoning within decision theory cannot consider knowledge derived from biology and evolutionary theory. In this light, we will examine how tissue biology has strong ties with many disciplines, including economics. This type of dialogue has been possible based on the premises that the two disciplines share characteristics such as the concept of competition, scarcity of resources, maximization, etc. As several authors have referred to an economy as a biological system, Alfred Marshall notes that one can look for an enhanced evolutionary paradigm. In fact, Marshall is the first to argue in favor of "biological economy" based on the assumption that economic and biological phenomena share many affinities, a complex and organic nature, involvement in a world of continuous development, submission and influences both qualitative and quantitative that imply that future events do not ever reproduce the same conditions. The originality of his idea of economic dynamics (or rather of economic development) derived from a biological model was long ignored until evolutionary theories permeated many fields of knowledge. We will see how in biology, natural selection, not a law of nature but a general principle from which one can construct models of explanations that serve to forecast. Developments in the field of evolutionary economics have been paralleled by research activities and publications that have been categorized under the label of "complexity science". The science of complexity has arisen from the interaction of different scientific fields, including physics, mathematics, biology, economics, industrial engineering, and computer science and now covers numerous important fields of scientific research. We would therefore examine how complexity science can offer a new vision of decision-making. Beginning from the simplest biological systems (ants or bees), this type of analysis has shifted to social systems to study the action of economic agents and to use simulations to test consequences, considering actions and interactions among economic agents. The focus will be mainly on the unpredictable and ambiguous world, the importance of nonlinear relationships and the role of self-organization, emergence, and co-evolution in organizational dynamics; thus, the analysis will avoid a reductionist explanation, foresight and linearity.

These brief ideas, sufficient to reveal the intricate complexity of the problem, will all be tested in the following pages. Ultimately, we do not wish to argue that individuals always behave rationally, although we can agree that individuals do not always behave irrationally. However, as Gould (1980) has already written, although irrationality is the major source of evolution, evolution removes irrationality. My hope is that this book can prove that agents act, to use Kantian terminology, not according to reason but with reason as a purpose.

References

Camerer, C. F., & Loewenstein, G. (2002). *Behavioral Economics: Past, Present, Future*. Caltech, Pasadena: Mimeo, Division of Humanities and Social Science.

Gould, S. J. (1980). Is a new and general theory of evolution emerging? *Paleobiology*, *6*(1), 119–130.

Kahneman, D., & Tversky, A. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Simon, H. A. (1972). Theories of Bounded Rationality. In C. B. McGuire, C. Radner (Eds.), *Decision and Organization* (pp. 161–176). Amsterdam: North-Holland.

Part I
Rationality and Neuroeconomics

Chapter 1
Rationality and Experimental Economics

1.1 The Theory of Rational Choice

The theory of rational choice (TRC) is a model of explanation used by social science theorists to interpret behavior. Initially, the theory was the dominant paradigm of economics. A fundamental postulate of neoclassical economics was that economic phenomena primarily resulted from the action of agents who were fully rational, equal and therefore indistinguishable from each other and all agents pursuing their own personal and individual gain. This postulate was followed by another epistemological postulate that found in economic phenomena the possibility of applying general laws expressed in mathematical terms as was achieved in physics. These postulates have been translated into fine mathematical models, reaching essentially hypothetical-deductive conclusions obtained from a basic idea introduced during the classical era by John Stuart Mill in his Principles of Political Economy, according to which economic phenomena comprise individuals who are all indistinguishable from one other, acting as agent-atoms in economic processes (Bertuglia and Vaio 2011).

Therefore, the assumption of neoclassical economics to establish economics as a mathematical economics attempted to identify a suitable role to play, as energy plays in physics, i.e., to find a function able to locate the maximum and minimum to define states of equilibrium. This comparative measure is utility, introduced several decades earlier by Jeremy Bentham (1789) in response to the problem of the justification of moral law (in which the correct choice was the one from which the most positive consequences derived and that considered the remainder of the balance in terms of happiness, between an action and its consequences) and reinterpreted as a mathematical function (Ingrao and Istrael 1987). According to the dominant interpretation of neoclassical theory, rational choice consists of acting to maximize personal gain, i.e., the option that allows the realization of the highest level of satisfaction for the agent. The latter determines the action: choose the greatest profit or the lesser evil. Therefore, during the deliberations, the agent compares opportunities and chooses the alternative that is more advantageous

M. Graziano, *Epistemology of Decision*, SpringerBriefs in Philosophy,
DOI: 10.1007/978-94-007-5428-7_1, © The Author(s) 2013

for him according to his beliefs. In brief, this option maximizes the difference between its costs and its advantages. The theory is limited to this instrumental concept of rationality: A choice is made based on the expected results (as when, for example, we undertake a program of university study in view of the salary we expect from the labor market). A rational agent's process of deliberation invariably begins by examining every opportunity in relation to each other and only in relation to his preferences. An agent is driven on the one hand by his beliefs, that is, by his information and expectations of the possible consequences, and on the other hand, by his desire for them.

Rational analysis generally begins with the premise that an agent wishes to choose what he invariably prefers. However, this premise is only part of the equation. The other important element in decision-making is the presence of specific "constraints" that make the choice necessary, and one virtue of the rational approach is to explicitly illustrate "the pros and cons" of possible alternatives. Therefore, a choice is rational if it corresponds to the scale of the agent's preferences, which is obtained by comparing various opportunities possible and converting preferences into utility functions. However, this conversion is possible only if the preference structure complies with certain restrictions: the four axioms of rational choice, namely, reflexivity, completeness, transitivity and continuity. The axioms function within the rational model as follows. First, the choice of an agent is rational if it conforms to his scale of preferences. Associated with different possible options, the latter (preferences) should always possess a value equal to themselves. Preferences must therefore be reflexive: ($xi = xi$). This condition is purely a formal necessity and depends on common sense. The second axiom, completeness, is necessarily involved in the structural formation of the agent's preferences. In fact, it is agreed that preferences can be ordered: ($xi \geq xj$) or ($xj \geq xi$). An individual must be able to compare all options and then prefer one or be indifferent because of their equivalence. Third, the scale of preferences must be transitive, i.e., it must conform to the classic example showing that if a person prefers an orange to an apple and an apple to a pear, then he must also prefer an orange to a pear. The order of preference must simply reflect an internal coherence; there must be no ambiguity. Finally, an agent's preferences must be constant. To illustrate this last condition, let us again consider the example of an agent faced with a choice of various goods. Two goods can be contained within a single set, and their quantity in the same can be amended to allow a comparison of the utility of each choice. Therefore, this axiom stipulates that there is no good that is absolutely necessary to a set and that cannot be exchanged for another. Thus, reflexivity and transitivity determine the order of preferences, whereas completeness and continuity are the conditions that allow a representation of the utility function. Ultimately, the rational choice theory offers an instrumental conception of rationality conceived as a coherent relationship between preferences, information and action.

Considering the axioms of rationality presented above, we can understand why the theory enjoys a special status among the social sciences. The theory's ideas are parallel to the developments and needs of other disciplines; analytical strength and operational effectiveness. In addition to responding to the needs of the social sciences, i.e., to describe, predict and prescribe, the theory performs an analytical theorization of social phenomena and explains it using clear arguments and assumptions. Insistence on these factors grants the theory special status in the social sciences and a role that transcends the areas covered by the economic sciences. The theory's importance derives from the fact that it avoids categorizing the explanation of individual actions "as a question of fact," and it allows predictions from what is postulated. In fact, in addition to conceiving and describing behavior as a process of maximization, the theory is able to predict the latter, again according to the criterion of maximization.

In 1944, John von Neumann and Oskar Morgenstern proposed the rational choice model in relation to decision processes in which for every choice, more consequences could be related, which completed the entire picture of the dominant paradigm. Specifically, in the renowned work The Theory of Games and Economic Behavior, the authors formulate the so-called theory of expected utility in which the function of utility (expected) was determined from the totality of the associated utility and from the possible results multiplied by their probability of occurrence. In essence, when one is faced with two alternatives, X will be preferred to Y based solely on the expected utility of X is greater than the expected utility of Y. To understand how the theory of expected utility functions, it is best to offer an example: An employee must decide whether to accept an offer of employment from agency A rather than from agency B, in which both agencies initially offered an equivalent beginning salary. Moreover, if the employee accepts the offer of agency A, he or she has a 50 % chance of obtaining a salary increase of 20 % during the first year of employment. However, if the employee accepts the offer of agency B, there is a 90 % chance that the salary increases by 10 % in the first year. According to the theory of expected utility, the employee must multiply the utility of the result of every alternative by the probability of obtaining that result; hence, the expected utility associated with the first option will be $0.50 \times U(20)$, and the expected utility associated with the second option would be $0.90 \times U(10)$. Moreover, one must note that the theory of expected utility does not prescribe what one must choose in the case of the given example of the employee. In other words, the theory focuses on the structure of the preferences rather than on their content because the objective of both authors was not to describe individuals' real behavior but to indicate how they should have acted based on the criteria of the coherent logic of the preferences and the calculation of the probabilities that one can hypothesize in the elaboration of the information. Therefore, overall, economic theory has opted for a formal approach. The function of utility used in economics was nothing other than a way of mathematically representing the order of individual preferences.

As explained by Giulio Giorello and Simona Morini, "These functions assign values, precisely called utils, to the possible outcome of actions from which an individual can choose and the theory prescribes to choose the action whose outcome maximizes its utility (that is, the one whose outcome is preferred to all others)" (Giorello and Morini 2008, p. 56). The utility discussed by the authors is ordinal utility because it merely expresses information relative to the governing of preferences regardless of their intensity. In contrast, the cardinal utility function is different because it assigns numbers to individual preferences. Therefore, the latter function allows the rational choice models to be extended to decisions in a condition of risk and uncertainty: Because the function is a question of numbers, it can "multiply the usefulness of the outcomes of different actions by their likelihood (obtaining the predicted or expected utility of choice) and define the rational decision as planned or expected utility maximization" (Giorello and Morini 2008, p. 57). For the simple fact that a number is assigned to the preferences, this function is often confused with the classical utilitarianism of Jeremy Bentham and James Mill, reworked by John Stuart Mill. However, in the hedonistic psychology of Bentham and Mill, usefulness was considered to identify an individual's overall happiness, which included the measure of various pleasures and pains. In this way, social utility was an objective value and usually measured as the maximization of the average utility of the individuals who are a part of the social utility. However, social utility implied an uncertainty in its extent, which is why the initial step towards the formalization of rational choice was the abandonment of the hedonistic concept of utility due to the logic of ordered preferences. In the ordinal utility theory, an economic decision could in fact be constructed without the need for psychological hypotheses regarding the intensity or the content and/or perceptions of sensations. Thus, utility became a relational rather than an absolute concept, defining a relationship of preferences between two alternatives: "A is preferred to B" is equivalent to "A is more useful than B". Therefore, an economic agent should not attribute an absolute value to a choice A but had to simply determine if (1) A is preferred to B, (2) B is preferred to A, or (3) A and B are equivalent. In this way, utilities and preferences become two inseparable concepts: utility referred to the scale of relative and subjective preferences, and preferences defined a utility function.

Therefore, preference can be considered the notion preceding economic rationality: decision theory, game theory and general equilibrium theory are all theories and models in which the rational agent is identified as the one who makes decisions or chooses optimal strategies or possible actions among those that are apparently offered to him and suited to the maximization of his subjective usefulness. In this way, decision-making was then represented through a formal mechanism in which the value of the model was measured by the ability to describe the behavior of the system studied. The preferences, strategies and utility functions provide a mechanistic explanation of decision-making without commitment on an ontological level. As noted by Binmore (1994), the standard practice in economics is to consider theories of rationality as formal exercises: "Axioms are propounded and the properties of rational individuals are then deduced mathematically. If the necessary mathematics is sufficiently challenging, attention then concentrates on whether an author's theorems are true rather than the more fundamental question of whether the axioms are successful in formalizing the concepts they are intended to capture." (Binmore 1994, p. 150). To be useful in empirical analysis, the standard interpretation of the TRC needs to support itself on two assumptions: (1) the postulate of rationality and (2) the allocation of a selfish utility function. The first assumption takes for granted that agents are rational in that a coherence of beliefs, desires and intentions of the agent is supposed. For example, the intention to do A must be able at least in theory to be deduced from the beliefs and desires of the economic agent as regards A. For the second assumption, individuals possess a utility function that leads them to prefer their individual interests, usually defined in monetary terms. The addition of this auxiliary hypothesis allows the TRC auxiliary to leave simple mathematical formalism and to perform empirical predictions: therefore, an agent's preferences can be inferred from their utility function as can the action that would likely be selected in a certain context.

1.2 Game Theory

Game theory is part of decision-making theory, and its underlying assumption is that in making decisions, an individual attempts to maximize his own benefits while minimizing the costs (cost-benefits analysis). According to this approach, individuals attempt to achieve the maximum benefit, taking into account the given constraints and other participants' behavior. The first significant studies in the field were developed by the mathematician John von Neumann, who developed in 1928 the first theorem of game theory, considering the so-called zero-sum games. These games consisted of subtracting the sum of the participants' costs from the sum of the same participants' benefits, leading to the result of zero. Every participant in the game must know the game rules and must be aware of the other players' moves and of their possible consequences; he must substantially have perfect information. The case of imperfect information is provided when a player is not thoroughly aware of the moves that have already occurred in the game.

Hence, game theory is intended to be a mathematical science that analyzes and describes particular types of conflict situations and seeks competitive and cooperative solutions by means of mathematical models. Game theory describes the study of individuals' decisions in situations in which interactions among different subjects exist so that a subject's decisions can influence those of others (partners and rivals) according to a feedback mechanism. Game theory has this specific feature: the theory accounts for strategic interactions in which an agent is not confronted with a passive environment but one that is composed at least partially of other agents. Specifically, in situations of strategic interaction, an agent is conditioned not only by his own actions but by those of other agents as well. The rationality of a decision-maker in a passive environment is sometimes distinguished from that of a decision-maker interacting with other decision-makers. In the first type of situation, the term "parametric rationality" is used, and "strategic rationality" is used in the second type.

In a typical game theory model, all participants know the game rules and are aware of the consequences of every single move. In fact, in the simplest formulation of the games (normal or strategic games) and with complete information, the temporal structure of the game is not explicitly described (as opposed to extensive games), and the uncertainty of the agents is restricted to the prediction of others' actions. A fundamental notion of game theory is the Nash equilibrium (Nash 1949), conceived by John Forbes Nash while still a student at Princeton and first reported by him in a 1949 article. A profile of actions reaches a Nash equilibrium if none of the players is (strictly) interested in a unilateral result. To better explain this notion, we refer to the classical example of the "dispute between the sexes" game. Luisa and Mario wish to spend an evening together, and they can choose between attending the ballet or a boxing match. Luisa prefers the ballet, whereas Mario prefers the boxing match; however, both prefer to spend the evening together rather than separately. This game situation can be represented by the following matrix:

		Luisa	
		Boxing	Ballet
Mario	Boxing	(2,1)	(0,0)
	Ballet	(0,0)	(1,2)

In this game, two Nash equilibria can be achieved, that is, the two profiles of action thanks to which Mario and Luisa choose to spend the evening together both in case 1 (boxing–boxing) and in case 2 (ballet–ballet). Therefore, in the case of a Nash equilibrium, each agent responds his best with regard to the anticipation of the other players' actions, an anticipation that is assumed to be correct. Briefly, in the Nash equilibrium, each agent 1) correctly anticipates the choices of other players (correct anticipation) and 2) positively responds to this anticipation (best response).

In contrast, the prediction that a certain profile of action does not reach a Nash equilibrium implies that at least one player makes a "mistake" in predicting his adversary's moves. The dispute between the sexes explains this point very well: if Luisa anticipates that Mario will choose the boxing match, her best response is to choose the boxing match as well. If she chooses the boxing match, and Mario chooses the ballet, then Luisa mis anticipated the move Mario "played," or she correctly anticipated but did not "play" her best move or response (the ballet).

Therefore, the pair "correct anticipation/best response" is the foundation of the Nash equilibrium. We can extend game theory to an arbitrary number of players, showing that under certain conditions, a situation of equilibrium always exists and is achieved when every player chooses a strategic move to maximize his benefit. This intuition undermines the dominant logic in Adam Smith's classical economics theory according to which a group achieves the maximum result when every component achieves what is best for himself.

To better illustrate this point, let us refer to one of the most well-known games, the "Prisoner's Dilemma" game. This game is the best example of the infringement of the so-called "Pareto optimum", which is a strategy adopted jointly by all players so that every player can achieve a positive pay-off (final outcome) without hampering his adversaries' pay-off. In other words, adopting a strategy that infringes the Pareto optimum simply means building a game in which the benefit of one player is reduced without increasing any of the other participants' benefits. The Prisoner's Dilemma is actually a classical example of a non-cooperative game: a type of game in which players cannot agree in advance to adopt the most favorable strategy for all of them.

The most notorious version of the Prisoner's Dilemma is the following: two suspects A and B are arrested by the police. The police do not have sufficient evidence to discover who is guilty, and after jailing the two prisoners in two different cells, they interrogate them by offering the following possibilities: if one confesses (C) and the other does not (NC), the one who does not confess will serve 10 years in prison, whereas the other one will go free; if neither confesses, the police will sentence them to only 1 year of prison; if both confess, the sentence will be 5 years in prison. Every prisoner can reflect on the strategy to choose—to confess or not to confess. In any case, none of the prisoners can know the other prisoner's choice. The pay-off matrix of the Prisoner's Dilemma is as follows:

Prisoner A/Prisoner B	Confesses	Does not
Confesses	−5,−5	0,−10
Does not	−10,0	−1,−1

Therefore, in this game, all the combinations of strategies represent "Pareto optimum" except for the strategy of mutual accusation. If the prisoners were able to communicate with each other, the best strategy would clearly be to not confess because both would be sentenced to only 1 year of prison. However, given that communication is impossible, for the second player (who is questioned second), it will always be most favorable to confess regardless of the first player's choice.

In fact, if A confesses, it is convenient for B to confess because in this way, both players would serve 5 years in prison. However, if A does not confess, it is more convenient for B to confess in any case because he would be set free, whereas the first player would serve 10 years. For this reason, it is always better for both players to confess because independently of the other player's choice, the pay-off is always higher in this condition.

Such a situation is called "Pareto inefficient" because although it is the most rational scenario, it does not represent the best possible situation. In fact, although it is much more convenient not to confess (both would serve only 1 year), this strategy is the least played because it is very risky (if the adversary confessed as is rational for him to do, one would risk serving 10 years; moreover, the adversary would be set free). Thus, the Prisoner's Dilemma indicates a fundamental aim: even if the Pareto optimum is rational from a collective perspective, it is not necessarily rational from an individual perspective. In essence, the participants of a game compete with each other and act according to their individual rationality whose goal is to maximize their personal profit, but doing so by no means guarantees that they obtain Pareto optimality, with the immediate consequence that their actions might lead to a dispersion of resources. We can illustrate this last point through a simple example. Consider the question of payment for a means of public transportation such as a train. Public transportation can work and provide service as long as a sufficiently large number of travelers are willing to pay for tickets. However, it is not necessary that every single traveler buy a ticket, only that a sufficient number of travelers buy tickets to cover expenses. For each traveler, it is more advantageous is not pay. However, if this strategy were to allow all passengers to use the public service, the service would fail, and you should rely on more expensive private service that is certainly an option without a doubt the worst cost (usually a private service costs more than a public service) than that would if all travelers would pay the ticket (a case study of Nash equilibrium).

One possible solution might offer an escape from the Prisoner's Dilemma by avoiding the deceptions of a perverse Nash equilibrium and repeating the game several times. In this case, it might be cheaper for each player to create a positive reputation for future meetings of the game: the abandonment of selfish self-interest aimed at seizing an immediate advantage could pave the way for cooperation with the enemy for an increasingly great future advantage (Akerlof 1970). Another possible solution might be the definition of a set of rules that impose some form of cooperation between players. Compliance with these regulations may be entrusted to any authority having the power to impose them (for example, the Legal State) or to leverage informal rules, such as internalized rules, i.e., traditions, habits, a sense of duty, a sense of ethics, which assumes the agents are naturally led to observe whether they have authority.

However, according to John Harsanyi, one of the most important and interesting theoretical rationality, relying on standards, institutional constraints, social values, is completely erroneous because the explanations that use these approaches, which are based among other things on unclear and poorly defined concepts (i.e., class consciousness), overestimate the degree of consensus and integration that truly exists in society and thus ignore conflicts of interest and individuals' disagreement over non-economic values. Furthermore, in the same way, this type of explanation appears to be static and conservative and therefore not suitable to explain changes. According to Harsanyi, social interaction must be explained in terms of the objectives and interests of people who are united and divided by certain things, who cooperate in one field and contrast in another. Ultimately, as game theory states, individuals seem to have "mixed interests." In Harsanyi's model, the optimal explanation is therefore reversed "are individuals who choose and change the rules or institutions based on their motivations and their goals in their own interest should do this in a rational way, where the term is understood in rational utilitarian sense: they are decisions made to maximize the welfare of individuals and of society as a whole" (Giorello and Morini 2008, p. 34). In Harsanyi's view, the problem becomes how to assign the proper weight that individuals attach to their motivations to ensure that one can predict their decisions. According to the author, because there is a theory capable of specifying the utility functions that individuals assign to economic assets and non-economic values of various kinds, the only way that remains is to infer from observations of individuals' behavior. By examining the behavior of individuals engaged in various social situations, Harsanyi derives his four postulates of rationality. The first postulate states that individuals can be impartial when they or their reference group are not directly involved (impartiality low cost). The second postulate states that a third person not involved in the dispute between two others directly concerned can assess the situation more fairly, prompting a behavior that is equivalent to maximizing a social welfare function (the principle of a sympathetic and impartial third person). The third postulate specifies that people also decide based on previous commitments related to family, friends, and social groups (constraint of previous commitments). Finally, the fourth postulate specifies that personal goals can be traced to two main reasons, namely, economic gain and recognition of social status (economic and social reasons). These postulates definitively identify the main variables that guide individuals' decisions, forming the basis of Harsanyi's "cognitive-utilitarian" empirical model: "Cognitive, as it tells the acceptance of certain values by the people with the uniformity of their beliefs about the consequences of the existing social system compared with those of alternative systems. Utilitarian, as it assumes that individuals choose between alternative social institutions and values on the basis of how these values and institutions have proved unable to satisfy their selfish personal interests or altruistic (Giorello and Morini 2008, p. 36).

Ultimately, the rational choice model by Harsanyi has a normative abstract because it does not necessarily reflect the actual behavior of choice or how individuals think under certain conditions (certainty, risk or uncertainty). The individuals mentioned by Harsanyi are abstract entities, fictitious, can recognize and follow the four postulates of rationality is reinterpreted in the right way.

1.3 Teleology, Instrumentalism and Interpretivism

The rational choice theory and the utility function of agents have been widely criticized, especially regarding their descriptive capacity because in the real world, many individual decisions depart from the perfect logic the theories invoke. The first to highlight the shortcomings of neoclassical economic models was Simon (1972), who emphasized that no limits were placed on the rationality of the subjects in these models given that the only restrictions theorized were structural or environmental and therefore did not influence decision making because they were not subject to the decision-maker's will or intervention. As a result, the decision-maker was required to have perfect knowledge of environmental constraints and extremely high-level computing capabilities. To formulate a theory of rationality that was more consistent with reality, Simon's first step was to criticize the assumptions of the rational choice theory, which had been judged incorrectly because it was based on an unrealistic model of rationality (defined as "Olympic"), and then to show the limits requiring emphasis in a theory of rationality that instead corresponded more to real-life facts. According to Simon, because of decision-making agents' cognitive limitations, which distanced them greatly from the abstract idea of the omniscient and rational *Homo Economicus*, such agents subjectively elaborate their perceptions of reality, namely, partial information at their disposal. Therefore, the choices that follow reflect different ways of achieving individual objectives. Simon believes that there is thus a considerable gap between rational and normative theories postulated by bounded rationality that he believes characterize real human behavior. One must consider human cognitive limitations during decision-making in terms of the selection, attention, acquisition, processing and storing of information. Through examples drawn from the game of chess (a closed domain game), Simon shows how a rational agent with Olympic rationality should be able to compute all the moves (10^{120}). However, in reality, the task proves to be difficult even for an extremely powerful computer. In fact, chess champions proceed in quite a different manner based on the considerations of no more than one hundred alternatives in the choice of a move or strategy. Simon calls the results "procedural rationality", i.e., a rationality that is defined based on adopted resolution procedures rather than on the final solutions obtained. The cognitive limitations of real decision-makers actually cause maximization to be impossible; thus, according to the author, maximization must be replaced by a "satisfactory" solution so that the decision-maker can consider options individually until he finds and chooses a satisfactory or sufficiently good option according to his individual minimum level of acceptability. A decision is a resolution of problems in a context in which the complexity of the information exceeds the computational capabilities of the agents, so their solutions must be simple and parsimonious. Therefore, it is necessary to replace the ideal agent model with another that is compatible with the information access capabilities and calculations that organisms possess in their given environments.

However, rather than considering the TRC false, economists and to a lesser extent epistemologists in the social sciences have preferred to adopt one of three attitudes: (1) teleology, (2) instrumentalism, or (3) interpretivism. According to the first (teleology), agents situated for a long period in a dynamic market will eventually adopt the prescriptions of the TRC. If the agents do not adopt these prescriptions, the agents will be eliminated. In other words, in this case, the market will eventually favor rational individuals and companies. Therefore, the TRC prescribes a set of rules that agents tend to adopt. According to the second interpretation (instrumentalism), the TRC does not describe the behavior of real agents but of ideal ones. Specifically, the TRC constructs mathematical models of decision-making in which the errors (computational, time, information) and the cognitive limitations of agents are not considered. From an epistemological point of view, the ideal agent plays the same role as the ideal gas or the perfect lens. As noted by Milton Friedman and Leonard Savage (Friedman and Savage 1948), the behavior of a professional snooker player could be predicted adequately if we could use physics formulas to calculate different trajectories to evaluate the most advantageous one. In other words, we assume that players will behave "as if" they know the mathematical formulas designed to calculate the angles and the best trajectories.

According to the authors, economics does not need to adopt or to comment on the level of an ontological commitment so that entities can only be considered parameters of an economic calculation. Additionally, reality certainly shows that real agents are not *Homo Economicus*, but that is no more surprising than discovering that a real gas does not behave as an ideal gas. Ultimately, the important element is the character of the predictive model. According to the third approach (Interpretivism), rationality is a rule of interpretation that allows the understanding of others' behavior, and the TRC in this case plays a marginal role (Davidson 2001). Even if agents do not conform to the TRC, such agents can always be interpreted by assuming that the agents' mental states are always motivated by reason. The rationality of the agents is therefore fundamental to a set of interpretive principles rather than to explicit rules. Although one can encode the rules of inference and decision-making (logic, Bayesian theory of rational choice, game theory, etc.), there are no universal principles that allow us to uniquely respond to questions such as "What should we believe?", "Which belief should we reconsider?", "What should we do?", and no principles that allow us to uniquely interpret a person's behavior. An agent always has the option to choose and then to reconsider a local belief. Nothing but the general and fundamental principles remain, that is, attributing a mental activity, rationality, beliefs, desires, knowledge, etc. In this way, the theory of rationality is nothing more than basic knowledge that we require to interpret behavior rather than to produce it. The theory suggests the epistemic norms and practical usage of the concepts of beliefs, desires, actions and rationality. The reality shows not the irrationality of agents but rather that the reasons that motivate agents do not meet the code: therefore, although technically irrational, the subjects are always interpreted as rational agents. Underlying these three interpretations is the same epistemological basis, namely, normativism, the idea that the theory of rationality is primarily a normative theory. This fundamental idea has largely modulated the relationship between formal theories of decision and the experimental results. This relationship will be illustrated with an example: the Ultimatum Game (Sanfey et al. 2003). In the Ultimatum Game, an agent A must propose a fraction $f > 0$ of the amount of money m to an agent B. If B accepts, he receives f, and A pockets m-f. If B refuses, both remain without anything, and the game ends there. The standard interpretation of game theory predicts that rational agents will behave as follows: A will propose the smallest fraction possible, and B will accept any amount rather than remain without anything (on the rationale that the monetary amount offered to him is preferable to zero). However, the experimental data show that agents A offer average amounts ranging from 20 to 50 % of the amount of money and that agents B generally refuse values of less than 20 or 30 %. The result is the same even when certain parameters are changed: the amount of money, the culture, the degree of anonymity, experience, etc., have little importance. The only cases in which agents behave in a manner approximately consistent with the standard interpretation is when B (but not A) is a computer, and A and B are groups that make collective decisions. Therefore, we can apply the three interpretations of the TRC to these results:

- *Teleology*. If agents play Ultimatum repeatedly, they will ultimately adopt the optimal strategy; all offers, even minimal ones, will be considered as possible, and any offer will be accepted.

- *Instrumentalism*. If agents do not follow an optimal strategy, this is due to their imperfection; however, the theory will generally provide useful forecasts.
- *Interpretivism*. Subjects can be interpreted as having reasons for choosing non-optimal strategies, for example, to enhance the fairness or even the absence of risk.

Implicitly, each of these interpretations tends toward the same idea: a theory of rationality is primarily a theory of the norms of rationality, which are reduced to rules that a rational agent will eventually follow (teleology), to those agents who follow an ideal agent (instrumentalism), and to those rules by which we interpret a rational agent (interpretivism). The idea that a theory of rationality is a normative theory is still so rooted in economic and philosophical practice that when experiments showed that agents did not behave as the theory predicted, scholars preferred to discuss "paradoxes" (Allais, Ellsberg, etc.) rather than empirical refutations or significant counter-examples. However, the normative outlook placed the TCR in an epistemologically problematic position. In fact, teleology and instrumentalism ultimately take for granted the usefulness of the TRC: both as a practical purpose and as an abstraction, the theory always describes an ideal state of rational agents, and this ideal is not subject to refutation. Theorems and axioms formally prove the character of rational decisions: the validity of the TRC is not evaluated by the measure of its predictions but by its formal virtues (completeness, consistency, etc.). In the Ultimatum Game, if subjects submitted irrational bids (always proposing a substantial portion of money at their disposal) and refuse insignificant offers, this occurs not because the theory predicts incorrectly but because agents demonstrate irrational tendencies. Economists skeptical of the results of the Ultimatum Game experiment state, "increase the amounts and you will see that they will adopt optimal strategies." However, in a version in which players had to divide a sum of $100.00, even $30.00 offers were refused (Camerer and Thaler 1995).

An argument often advanced to justify the standards of rationality is that just as grammar consists of rules of language, theories of rationality consist of decision rules that recommend courses of action. Ultimately, the TRC is a coherent set of rules and not an empirical study (Marschak 1951). If a person does not choose what the TRC recommended, he makes a mistake in the same way that a person who provides an incorrect result when multiplying 234 by 92 would blame the arithmetic and not the individual. However, if this is true, then economic theory should not be considered a scientific theory open to falsifiable proposals but only a science that can describe how traders would behave if they were fully rational individuals. In this way, a business practice is considered rational only if it can be rationalized within a theoretical model proposed. However, in this sense, the theory excludes itself from science because as we know, science involves a systematic attempt to verify, falsify and compare theories. Otherwise, we face a complex tautology justified by its formal rigor. Ultimately, interpretivism only clarifies this perspective: rationality cannot be analyzed based on rules and codes but through the principles of interpretation.

1.4 Experimental Economics

However, since its debut in the 1950s, experimental economics has clarified that the TRC, or at least its auxiliary hypotheses, are not verifiable, thus designating a long list of possible dissonance, for example, the aversion to risk and uncertainty, the violation of the independence axiom, and the preference for morally preferable but not economically optimal solutions. In particular, it dramatically scales back the role that is assigned to rationality in the choice process, a role that we have found in previous theories to assume the compelling possibility of calculating the objective of several utilities achievable with different alternative choices. The birth of experimental economics is traced to the work of Herbert Simon and his theory that economics should not address the study of rational behavior in an abstract way but, on the contrary, should address the empirical study of the limits of individuals' ability to calculate when faced with a choice (and therefore not with presumed objective rationality) and how these limits subsequently affect real economic behavior. Daniel Kahneman and Amos Tversky follow Simon's reasoning, although with different assumptions. The authors integrate economics and cognitive science, creating the famous 1979 work the Prospect Theory (Kahneman and Tversky 1979). The starting point of the Prospect Theory is the observation that the concept of expected utility, the basic normative notion introduced by von Neumann and Morgersten, expressed the assessment of the consequences of a choice, including a definition of the probability that these consequences actually occur, and the theory is ultimately inadequate to predict the actual behavior of real decision-makers.

Essentially, the Prospect Theory is based on three fundamental assumptions. The first assumption postulates the existence of an asymmetry between the impact of earnings and the impact of loss (loss aversion). In practice, the authors believe that a subjective evaluation that people conduct of the value function is concave for gains and convex for losses:

$$Uselessness\ of\ loss\ x \ > \ utility\ of\ gain$$

This aversion to loss might involve choices in contradiction with expected utility. In a systematic way, people overestimate the probabilities of catastrophic events (the strong emotional dimension related to these events is no stranger to this phenomenon). Moreover, it is interesting that Adam Smith described this emotional dimension in 1759: "We Suffer more…when we fall from a better to a worse situation, than we ever enjoy when we rise from worse to a better…" (Camerer and Loewenstein 2002, p. 4). This asymmetry can sometimes lead to sub-optimal decisions that do not conform to the expected utility theory. For example, Shefrin and Statman (1985) indicate the type of error generated by this curvature of the value function: As investors are much more sensitive to losses than to gains, they hold on for too long to securities that fall below their purchase prices, and to avoid considering the effect of a loss, they refuse to sell (this is explained by the convexity of the value function for losses). In contrast, when the price of a security situation remains at its purchase value, investors have a tendency to sell too quickly (this is explained by the concavity of the value function for gains). This phenomenon is known as the “disposition effect”. The same result was shown for consumers, who are more sensitive to price increases than to price reductions (Chen et al. 2006). Kahneman and Tversky suggest that the value function is applied to gains and losses (especially monetary) and that they are evaluated according to a neutral point; the loss is more painful than the happiness generated by a similar gain. The figure below illustrates the asymmetry between gains and losses (Fig. 1.1).

To better understand decisions in a situation of uncertainty, Kahneman and Tversky propose replacing the probabilities associated with a choice that aims at satisfaction with a decisional pondering that would take greater account of what has been observed empirically. In fact, agents do not assess well the probabilities associated with an event and have difficulty applying the means of statistical analysis beyond purely mathematical applications. In particular, agents have a tendency to overestimate the probability of rare events and to underestimate the probability of frequent events. This poor estimate of the likelihood of events derives from what the authors call heuristics. In their 1974 article, Tversky and Kahneman do not define a heuristic opinion but describe it as a set in a process that reduces the degree of difficulty associated with assessing the evaluation of value and probability. In their article, three types of heuristic judgments, anchoring, availability, and representativeness and a dozen biases associated with them, are outlined.

For Kahneman and Tversky (1974), the anchoring heuristic means and works as an “anchor” of judgment: subjects based their judgments on a fact first information and do nothing, then you adjust it to produce a final answer. In their experiment, the authors asked subjects to estimate the number of African countries that are part of the United Nations. Before the participants answered, the researchers asked them to turn a wheel of fortune that contains the numbers 1–100, then to judge whether the number of countries is greater or less than that provided randomly by the wheel, and then to provide their estimate. The authors showed a correlation between the random number data from the wheel and the number proposed by the participants. For example, the number 10 was presented to one group of participants, whereas the number 65 was presented to another group. The median estimate of the percentage of African countries resulting from the two groups was found to be 25 and 45. The authors interpreted this discrepancy of judgment with the explanation that the first group with the number 10 performed an adjustment upwards, whereas the second group with the number 65 executed a subsequent downward adjustment.

However, the availability heuristic requires the subject to estimate the probability of an event based on the ease of recalling past cases in which this event occurred. Therefore, this heuristic is related to the ease with which information of such events is retrieved from the memory. In fact, when individuals attempt to assess the probability of the occurrence of a future event, look in the memory of similar events have occurred, according to the ease with which they can recall of similar events, or more, consider it less likely to occur again for that event in

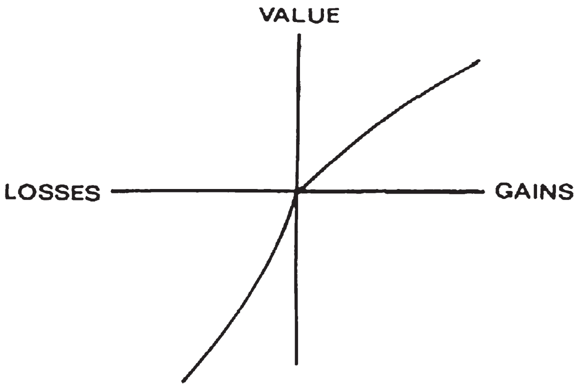

Fig. 1.1 *Source* Kahneman and Tversky (1979)

question. A typical example of heuristic availability is that most people consider it more dangerous to travel by plane than by car despite statistics showing the exact opposite. This belief can be explained simply by the fact that the media grant significant attention to a plane crash so that people vividly remember that event rather than the memory of a car crash, which normally is not subject to as much publicity.

Finally, the representativeness heuristic is a search for an image, an internal representation that individuals are the objects of a sample to be evaluated. Therefore, the probability that they attribute to these objects depends on the similarity between the sample, the event and the source from which it originates or that produces it. In a famous experiment, Kahneman and Tversky (1973) indicated that when the parties involved in an experiment are brief descriptions of personality (which they stated were drawn at random from descriptions of approximately 100 professionals of whom 30 were engineers and 70 lawyers) such as the following:

- Jack is 45 years old. He is married and has four children. He is rather conservative, careful and ambitious. He has no interest in politics and social issues. He spends most of his time cultivating his many hobbies, such as bricolage, sailing and mathematical puzzles.

The probability that Jack is one of 30 engineers in the sample….%.

The subjects evaluated the probability that the description corresponded to an engineer rather than a lawyer according to the degree of representativeness of their stereotypes of the description without regard to the likelihood provided by initial primary data (i.e., 30 and 70 %). In contrast, when subjects were not presented with a description of personality, they considered the primary data.

According to the authors, this inability to assess the probability of an event's occurrence based on primary probability in the presence of a description is a flagrant violation of the principles of normative theories and evidence that the subjects used heuristics when they were issued a judgment of belonging or inclusion.

For example, a type of judgment that seems to be based on the heuristics of representativeness is the "categorical prediction." This type of judgment is required to estimate the probability that a person with certain characteristics belongs to a category or class target. The most famous example is the problem of Linda. The participants in this experiment (Tversky and Kahneman 1983) read the following character description:

Linda, 31 years old, is very intelligent and has a degree in philosophy. Furthermore, she is socially committed to end discrimination and has participated in anti-nuclear demonstrations. The two researchers then asked the participants to rank the following response options in order of probability:

(1) Linda is active in the feminist movement (A).
(2) Linda is a bank teller (B).
(3) Linda is a bank teller and is active in the feminist movement (A and B).

The problem of Linda became famous because the majority of the participants (89 %) chose outcome 3 as more probable than outcomes 1 and 2. This evaluation constitutes a violation of one of the fundamental principles of probability theory known as the "rule of the conjunction" and is called the "conjunction fallacy". In this fallacy, the occurrence of the conjunction of two events (h1 & h2) is considered more likely with respect to the probability of the presentation of one of the two constituents (h2), an opinion that is contrary to the regulation provisions of the rule of conjunction, which states precisely that the probability of the conjunction of two events, p (h1 & h2), cannot exceed the probability of its constituents considered individually, p (h1) and p (h2), because the extension of the conjunction is included in the extension of its constituents. Specifically, the violation of the conjunction is a violation of the following Pascalian principles, the "multiplication rule for the conjunction" that $P(A \& B) = P(A) \times P(B \mid A)$, from which it follows that $p(A) \geq p(A \& B)$. In particular, given that $p(A \& B) = p(A) \times p(B \mid A)$ and $p(B \mid A) \times p(A) \leq 1$, then $p(A \& B) \leq p(A)$, and $p(A \& B) \leq p(B)$. Moreover, in addition to violating the probabilistic rule of conjunction, the participants' solution in the experiment on the Problem of Linda reveals a strictly logical error as the case of Linda also involves a problem of relationship inclusion between classes. The whole unit "bank teller and feminist" is a subset of the broader class "bank teller". The inclusion relation determines in this case a report of quantity: if a class A includes class B, then it is safe to conclude based on the cardinality principle (Gallistel and Gelman 1978) that A is larger than B; therefore, to be an element of class A, something is more likely to be an element of class A and class B simultaneously.

Gerd Gigerenzer has continued the work of Kahneman and Tversky by introducing a new element into the debate, namely, the role of the environment and of natural selection; thus, Gigerenzer has found a new type of real rationality, which is ecological rationality, a type of adaptive rationality viewed as a set of "fast" heuristics appropriate for the environment. For example, Gigerenzer specifies that when people with little knowledge of U.S. geography were asked whether Detroit or Milwaukee has more inhabitants, the majority of the participants (German students) responded correctly by choosing Detroit. According to the author, the subjects answer by following a very simple heuristic (known as the recognition heuristic). When someone is required to assess a certain thing, if one object is recognized and the other is not, then the detected object is inferred to have a larger value (Gigerenzer 2007). This heuristic works in certain contexts because we recognize or are aware of certain information concerning large cities as opposed to small ones. In this regard, a heuristic is far from the classical canons of rationality precisely because it relies on ecological regularities (in the case of Detroit and Milwaukee—on the correlation between the size of large cities and the number of inhabitants).

According to Gigerenzer, we can find a form of structural isomorphism between the environment and whether the rationality paradigm is presented in the form of natural frequencies rather than conditional probabilities (Gigerenzer and Murray 1987). We illustrate this aspect through the well-known paradigm of mammography (Eddy 1982) presented initially as conditional probabilities (with percentages) and then through natural frequencies:

> One percent of women who are 40 years of age participating in a finding they have a routine breast cancer. Eight out of ten women with breast cancer will have a positive mammography, and 9.6 % of women without breast cancer will also have positive mammograms. A woman who belongs to this age group has a positive result of a routine mammogram. What is the probability that she actually has breast cancer?…%
>
> Ten of one thousand women who are 40 years of age participating in a finding they have a routine breast cancer. Eight out of ten women with breast cancer will have a positive mammography, and 95 of 990 women without breast cancer will also have positive mammograms. A woman who belongs to this age group has a positive result during a routine mammogram. What is the probability that she actually has breast cancer?…%

Cosmides and Tooby (1996) have shown that the neglect of the base rate disappears in a frequent format, with 76 % of participants successfully solving the problem in frequency compared with only 12 % of the cases resolved through the odds. Therefore, the formats can be equivalent from a mathematical point of view but not from a psychological point of view (Feynman 1967). Laura Martignon and colleagues (Martignon et al. 2003) have shown that although equivalent, Roman numerals and Arabic numerals are not treated equally. For Richard Feynman (1967), various representations of the same mathematical formula can evoke different mental images and bring new mental representations that lead to different solutions. These considerations are coupled with new discoveries by neuroscientist Stanislas Dehaene and many other cognitive scientists, according to which human beings enter the world equipped with two systems of numerical representation: the first is "natural-rough" and no influenced by culture, and the second is "language-dependent," which is the basis of our accurate knowledge (Dehaene and Brannon 2011). This hypothesis is supported by numerous experimental data revealing that an important numerical knowledge exists even before the linguistic phase and is regulated by two "psychological laws": "the effect of distance and the effect of size," which establish the principle that the more minimal the difference between the two groups to compare, the greater is the level of difficulty in distinguishing (for example, 6 and 4 or 12 and 8 are distinguished better when the quotient is 1.5 rather than 5 and 4 or 10 and 8 when the quotient is equal to 1.25). In the same way, distinguishing groups/numerosities will be more difficult when the size is larger (therefore, 2 and 3 are easier than 24 and 25).

In any case, beyond the recognition of numbers, for our present purpose, the important issue in the distinction between natural frequencies and conditional probabilities is the answer, the theoretical and methodological debate between rationality/human irrationality. In fact, if one of the conditions enabling people to successfully solve problems posed, then it is not legitimate to speak of mistakes, or worse, human irrationality. Therefore, the alternative that remains is that limited human rationality is submitted to an effect of context and presentation. It seems that the latter condition applies both to individuals naive man-of-way as to subject experts, such as doctors who are likely to have scientific knowledge of probability.

The result of the pioneering work of Kahneman and Tversky, the numerous attempts to develop and refine the theoretical concepts of heuristics and biases (Griffin and Kahneman 2002), and research attempting to extend the theories to other areas of psychology, decisions in the legal field (Saks and Kidd 1980) or to the sphere of public policy (Thaler 1983) is that such research finds opposite results from the findings of economic studies: the issue is no longer a question of predicting the behavior of agents but of building on these behaviors to discover or explore the cognitive biases of utility functions. For example, the research program of Kahneman and Tversky was not concerned with the prediction of actions but the ways in which actions revealed cognitive biases (Kahneman et al. 1982). Research conducted by Werner Güth, Ernst Fehr, Frans van Winden, Georgre Ainslie and other experimental economists aim to use the real decisions of agents to describe their utility functions more precisely (for example, taking into account dynamic preferences or the existence of social preferences). This inversion is similar to that implemented by the analysis of revealed preferences used to interpret the behavior of consumers (Samuelson 1938). That analysis did not attempt to predict the behavior of agents from a hypothetical utility function but sought to explore the agents' utility function by observing their real choices.

1.5 Criticism of Experimental Economics

The application of theories to the economy (such as the applications of Kahneman and Tversky or Gigerenzer) and methods derived from cognitive psychology have provided a new and radically different perspective compared with standard interpretations of decisions. Moreover, for approximately the last thirty years, resources have been invested in the development of "experimental economics", i.e., the application of the "controlled experiment" method to economic decision-making, a method that has greatly contributed to the progress and development of the natural sciences. These experimental practices have a radically different approach compared with those that use logical-deductive methods of traditional investigation. The aim of these experimental practices is to understand how decision-makers actually make real decisions and to "describe" the manner and psychological mechanisms that lead them to make choices. Therefore, whereas philosophy and rational choice theory focus primarily on regulatory issues, prescriptive, experimental economics is concerned with the descriptive aspects of decision-making. Only in this way can the contradictions in the assumptions adopted by traditional economic models be revealed, the first being the groundlessness of the notion of "utility function" according to which it is possible to assign each action a value that specifies its desirability.

Experimental economics has had the merit of introducing psychological concepts to the debate on decision-making processes, operationalized to produce behavior that is predictable and testable experimentally, thus allowing emphasis on so-called non-rational behavior (previously considered "anomalies") and the important role of emotions and unconscious and implicit cognitive processes. However, this progress does not mean that experimental economics is beyond criticism or that it can single-handedly solve all epistemological problems related to decisions. Contrarily, this field has obvious limits of applicability, and certain of its methodological procedures have created new problems, which still await solutions. Indeed, there is an enormous gap between the economic phenomena that occur in real life, phenomena that are multi-faceted and complex, and phenomena that one can replicate, manipulate and control in experimental contexts. For example, economic phenomena such as the relationship between technological progress and growth or the relationship between money supply and inflation cannot be replicated in a laboratory. However, although we postulate that with creativity, we might one day reproduce on a small scale phenomena that initially seem unsuitable to study through an experimental approach, enormous simplifications would be necessary to control the variables of the game and to ensure the conditions for the feasibility of the research (for example, we must minimize the time, knowledge, background and commitment required of subjects involved in an experiment to ensure understanding of the proposed tests). Moreover, in a laboratory experiment, a situation must always be sufficiently familiar to the participants so that they can recognize and empathize, and simultaneously, the same situation must be sufficiently new so that it does not trigger an automatic response. In addition, the experimental situation must not recall experiences that the subject may have had previously and/or idiosyncratic reactions that produce distortions in his answers. To avoid these types of problems, the experimental situation must necessarily be clear and structured with all the involved elements made explicit, i.e., the elements must be declared by the investigator. However, in that process, it follows that the experimental situation tends to lose the background of tacit assumptions that are a part of decisions in real situations.

The concern behind this type of argument is primarily that when judging the performance of subjects engaged in reasoning tasks in a laboratory setting, you end up neglecting completely the practical goals instead of their reasoning in a naive natural context. This criticism of experimental economics implies that outside the laboratory, human inferences are always directed to the finalization and implementation of a specific objective. Moreover, as numerous studies (Sperber et al. 1995; Bagassi and Macchi 2006; Sher and McKenzie 2006) have emphasized, certain major phenomena traditionally considered as exemplary cases of the limits of human thought actually result from sophisticated inferential processes largely influenced by semantic considerations and attributions of meaning, which depend critically on the situation in which the arguments unfold. Thus, the roles of the physical context (the environment in which the subject exists), the rules relating to certain social situations, the linguistic context (the use of special codes or stylistic registers) and the co-text (the elements that accompany the text of the problem itself influencing the interpretation) become crucial. In particular, from a pragmatic perspective, the representation of the problem of the subject results from the text of the communication problem, not only what is literally stated (sentence) but also what is involved and then effectively communicated in an utterance. The distinction between sentence and utterance is crucial in the communication theory of Grice (1975), who has identified the "conversational rules" to which communicative acts naturally comply in accordance with a general principle of cooperation. Stated differently, according to Grice, communication is a basic orientation towards active cooperation in the exchange of information, which gives rise to certain implicit rules that guide speakers in choosing the most effective ways to transmit information. The observance or violation of these rules produces the "conversational implies true", i.e., the implicit information that is obtained from the relationship between linguistic expression and the context in which it is used. Therefore, the implicit knowledge of rules and frameworks of reference provide "additional information" to complete the information transmitted by an explicit expression. In contrast, in the formal language of logic, the only source of information is produced by "conventional implications", which depends exclusively on the conventional meaning of the words used in speech. Thus, according to Grice's approach and the two main authors who have proposed a revision, namely, Levinson (2000) and Sperber and Wilson (1986), communication is achieved when a receiver recognizes the particular intention with which a communicative act is produced, but in formulating a sentence correctly, it is necessary that the speaker assume the listener's ability to infer the correct "meaning". Thus, in the laboratory studies of experimental economics, the comparison between the experimenter and the subject possibly involves very different assumptions. This situation occurs particularly when an experimenter poses a problem as a purely logical exercise to be solved using only formal rules, whereas the subject interprets it in terms of "everyday language", then using conversational rules and other assumptions are not strictly logical governing disclosure daily.

The considerations that influence the rules of discourse have on processes of thought does not seem surprising. However, this immediately obvious intuition did not influence the psychology of thought or experimental economics. In fact, with the exception of the pioneering research of Mosconi (1990), the difficulties encountered by subjects in performing the tasks that were proposed were almost always interpreted as experimental demonstrations of human error when they would be explained more simply as the effect of using two different codes, the "natural" code adopted by the subject and the "formal" code endorsed by the investigator. We illustrate all these concepts with a concrete example by returning briefly to the problem of Linda and the conjunction fallacy. As mentioned, the conjunction fallacy violates the regulations' provisions because the subject is assessed by the occurrence of the conjunction of two events (Linda is a bank teller and is active in the feminist movement) as more likely than the probability of the occurrence of any one of its constituents (Linda is a cashier in a bank). Several authors (Fiedler 1988; Macdonald and Gilhooly 1990; Mosconi and Macchi 2001), who are proponents of the pragmatic school, have speculated that the conjunction fallacy could be due to pragmatic factors that lead individuals to interpret the experimental problems differently from the investigator's intentions in creating the experimental material. Specifically, as evidenced by Moro (2009), one of the authors of the "misunderstanding Hypotheses", one factor that influences the subject differently from the investigator's intentions concerns the interpretation of the connective "and". According to Moro, participants, correctly following certain pragmatic rules of communication and grammar, might have interpreted the conjunction 'and' as a disjunction that unlike the logical meaning of the connective element, suggests a union instead of an intersection. In this view, the evaluation of the subject would not be misleading because the probability of the disjunction of two events is always greater than or equal to the probability of the events considered individually.

Therefore, the use of the experimental method in the study of economic decisions is an enterprise not without difficulties and limitations in data interpretation. From a methodological point of view, the experimental evidence must indicate unambiguously that if a hypothesis is false, the probability of observing this type of evidence must be very low. In the literature, this criterion is explained through the concept of "strict control". According to the theory of severity control (Mayo 1996) (D) h and indicates if and only if the conditions are such that control T, T, P (e/h) > P (e/−h), i.e., one must generate a type of evidence (s) when the hypothesis is true and another type of evidence (−e) when the hypothesis is false. When a procedure meets the requirements of (S), we are faced with a "strict control". The first consequence of this control is that normally, a certain set of data can strictly control a limited or "local" hypothesis, and this occurs for one simple reason, namely, data collected in an experiment are usually compatible with a wide range of theories. This implication of the strict control of the method is in stark contrast with what is sustained by certain philosophers of science, such as Lakatos' argument that on the contrary, success in predicting new facts can support an entire research program. However, for our purposes, the difference is minimal. In fact, for Lakatos, any research program is upheld by a series of "metaphysical" considerations, which means that research programs in their entirety need not be supported by evidence; on the contrary, significant movements of interest from one research program to another are also influenced by non-empirical considerations (Guala 2005). Therefore, it is plausible that the experimental method is most effective when testing hypotheses "locally" or at a "low level". The construction of more general models is only possible in the course of a long research program and after several different experiments.

The consequence of this minimum is that despite all the limitations and doubts, we continue to require normative theories because they continue to be a system of reference for experimental-descriptive theories: for example, only through the concept of rational behavior can we detect violations of this principle (Piattelli Palmarini 2005). Kahneman and Tversky themselves emphasized the indispensability of both theories even if their task was to perform two different functions: normative theories helped to shape the rules of rational behavior, and descriptive theories clarified the real processes at the basis of decisions. Ultimately, for a normativist, a theory of rationality is only a theory of the norms of rationality, and a theory that is merely descriptive is simply an exercise in psychology. In any case, normative theory's focus on specifying how things should be as opposed to how they really are seems to be incompatible with naturalism, which instead proposes that rational agents are part of nature; therefore, appeals to entities other than those of the natural sciences are unnecessary. The corollary of this principle is that the project of the naturalization of rationality is controlled by the following dilemma: (1) either rationality contains rules, but these cannot be naturalized (not being the rules of natural facts) (2) or rationality does not involve rules, and all that can be naturalized must be of a descriptive nature. In fact, when we have a mathematical formalization of the alleged behavior of a rational economic agent, these laws are not laws of nature or empirical regularities because only the behaviors are specified or the reasoning, which must be according to what the norms require. As well as what must be the case is not considered to be natural facts but according to the values of what the norm prescribes. In this sense, the tenets of the theory of rational choice (or game theory) are simply "obligations".

Thus, the real question appears to be as follows: are the standards of rationality enunciations under a law (or obligations), or are they universal regularities in the scientific sense? The theories, models and arguments explained thus far by the normativists and the descriptivists create a situation in which the theories of rationality vacillate between the status of rules and descriptions of norms on the one hand, and on the other, between the difficulty in reconciling them with normative and descriptive projects that are credible and applied to idealized systems. Therefore, an integrative approach that seeks to clarify the normative/cognitive-descriptive relationship is proposed to define a new line of analysis for decision-making contexts. According to the terms outlined in this book, the enunciations of decision theory and game theory are not limited to the two possibilities listed above, that is, that they are either only general descriptions of natural facts or obligations. A third possibility can be proposed that is compatible with the significant uses demonstrated for theories of rationality. The statement, "rationality is a normative concept" can mean two things depending on what "standards of rationality" signify. The standards can be understood in a non-instrumental sense, a "must" absolute, but can also be interpreted in an instrumental sense, such as requirements for a goal-oriented action. In this sense, the modal vocabulary used can be considered normative but in a conditional and not absolute way: the modal force in this case only results from the conditions of satisfaction of the action. Thus, a distinction must be made between "ought" and "ought in order to". In this distinction, we follow what Duncan Luce and Howard Raiffa wrote of game theory: "We belabor this point because we feel that it is crucial that the social scientist recognize that game theory is not descriptive, but rather (conditionally) normative. It states neither how people do behave nor how they should behave in an absolute sense, but how they should behave if they wish to achieve certain ends. It prescribes for given assumptions courses of action for the attainment of outcomes having certain formal *optimum* properties. These properties may or may not be deemed pertinent in any given real world conflict of interest. If they are, the theory prescribes the choices which must be made to get that optimum." (Luce and Raiffa 1957, p. 63). However, normative economics itself is not a branch of economics that promotes standards of right action but rather an activity that attempts to determine the best policy measures to maximize or minimize some economic variables: GDP, inflation, unemployment, etc.

Of course, these topics are not aimed at reducing all forms of normativity to instrumental normativity: rationality is present and desirable in many aspects of social life. However, what I wish to propose in this discussion is that for an economist, a rational agent is not someone who thinks rationally but someone who appears to have rational "preferences" such that if he prefers A to B, he simply chooses A and not B. The entity being evaluated by the criteria of rationality is not a thought but a behavior. For example, a consumer prefers to buy item A as opposed to item B because the former is less expensive, which shows a form of rationality that is different from the ability to articulate a rational thought. In this case, the consumer is not an agent, an individual who limits himself to only existing in the world; he exerts an influence on the world through an action that itself expresses something that is inherent to the agent (nothing more than that which is prescribed by our ordinary psychology and biology) and moves it to such action. All these processes constitute rational action, i.e., confer on such action the qualities of consistency, validity, and above all pragmatism.

References

Akerlof, G. (1970). The market for "Lemons": Quality uncertainty and the market mechanism. *Quarter Journal of Economics, 84*(3), 488–500.

Bagassi, M., & Macchi, L. (2006). Pragmatic approach to decision making under uncertainty: The case of the disjunction effect. *Thinking and Reasoning, 12*(3), 329–350.

Bentham, J. (1789). *An Introduction to the Principle of Morals and Legislations*. Oxford: Blackwell.

Bertuglia, C. S., & Vaio, F. (2011). *Complessità e modelli*. Torino: Bollati Boringhieri.

Binmore, K. (1994). Rationality in the centipede. In M. Kaufmann (Ed.), *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning about Knowledge* (pp. 150–159). San Francisco: Morgan Kaufmann Publishers.

Camerer, C. F., & Loewenstein, G. (2002). *Behavioral economics: Past, present, future*. Pasadena: Mimeo, Division of Humanities and Social Science.

Camerer, C. F., & Thaler, R. (1995). Anomalies: Dictators, ultimatums, and manners. *Journal of Economic Perspectives, 9*(2), 209–219.

Chen, M. K., Lakshminaryanan, V., & Santos, L. R. (2006). The evolution of our preferences: Evidence from capuchin monkey trading behavior. *Journal of Political Economy, 114*(3), 517–537.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1–73.

Davidson, D. (2001). *Subjective, intersubjective, objective*. Oxford: Oxford University Press.

Dehaene, S., & Brannon, E. (2011). *Space, time and number in the brain searching for the foundations of mathematical thought*. San Diego: Elsevier.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge: Cambridge University Press.

Feynman, R. P. (1967). *The character of physical law*. Cambridge: MIT Press.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*(2), 123–129.

Friedman, M., & Savage, L. (1948). The utility analysis of choices involving risk. *The Journal of Political Economy, 56*(4), 279–304.

Gallistel, C. R., & Gelman, R. (1978). *The child's understanding of number*. Cambridge: Harvard university Press.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.

Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. New York: Erlbaum.

Giorello, G., & Morini, S. (2008). *Harsanyi visto da Giulio Giorello e Simona Morini*. Roma: Luiss University Press.

Grice, P. (1975). Logic and conversation. In P. Cole, J. L. Morgan (Eds), *Syntax and semantics. Speech acts* (pp. 41–58). New York: Academic Press.

Griffin, D. W., & Kahneman, D. (2002). Judgment heuristics: Human strengths or human weaknesses? In L. Aspinwall & U. Staudinger (Eds.), *A psychology of human strengths: Perspectives on an emerging field* (pp. 165–178). Washington: APA Books.

Guala, F. (2005). *The methodology of experimental economics*. Cambridge: Cambridge University Press.

Ingrao, B., & Istrael, G. (1987). *La mano invisibile. L'equilibrio economico nella storia della scienza*. Roma-Bari: Laterza.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237–251.

Kahneman, D., & Tversky, A. (1974). Judgement under uncertainty. Heuristics and biases. *Science, 185*, 1124–1131.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

Levinson, S. C. (2000). *Presumptive meanings*. Cambridge: MIT Press.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.

Macdonald, R. R., & Gilhooly, K. J. (1990). More about Linda or conjunctions in context. *European Journal of Cognitive Psychology, 2*(1), 57–70.

Marschak, J. (1951). Why « should » statisticians and businessmen maximize « moral expectation » ? In J. Neyman (Ed.), *Paper presented at the Second Berkeley Symposium on Mathematical Statistics and Probability (Cowles Foundation Paper 53)*. Berkeley: University of California Press.

Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 189–211). Chichester: Wiley.

Mayo, D. G. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.

Moro, R. (2009). On the nature of the conjunction fallacy. *Synthese, 171*(1), 1–24.

Mosconi, G. (1990*). Discorso e Pensiero*. Bologna: Il Mulino.

Mosconi, G., & Macchi, L. (2001). Pragmatic rules and the conjunction fallacy. *Mind & Society, 2*(1), 31–57.

Nash, J. F. (1949). Two person cooperative games. *Econometrica, 21*, 128–140.

Piattelli Palmarini, M. (2005). *Psicologia ed economia della scelta*. Torino: Codice Edizioni.

Saks, M. J., & Kidd, R. F. (1980). Human information processing and adjudication: Trial by heuristics. *Law and Society Review, 15*, 123–160.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica, 5*(17), 61–71.
Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science, 300*, 1755–1758.
Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long. *Journal of Finance, 40*(3), 777–790.
Sher, S., & McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition, 101*, 467–494.
Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & C. Radner (Eds.), *Decision and organization* (pp. 161–176). Amsterdam: North-Holland.
Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31–95.
Sperber, D., & Wilson, D. (1986). *Relevance communication and cognition*. Oxford: Basil Blackwell.
Thaler, R. (1983). Illusion and mirages in public policy. *The Public Interest, 73*, 60–74.
Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–315.

Chapter 2
Neuroeconomics

2.1 Neuroeconomics and Causality

The rational choice theory (TRC), in its standard interpretation, is not presented as a causal theory, but as a formal-normative theory. Experimental economics, replacing the standard assumptions regarding the interpretation of the TRC's more realistic assumptions—for example, a utility function that includes social preferences—has allowed us to improve the prediction. However, experimental economics has failed to overcome another theoretical difficulty that affects the TRC, the absence of a formal causal link to the entities hypothesized. In the end we are before the same problem Chomskyan linguistics faced. Chomsky, in fact, proposed a formal model of explanation of linguistic competency assuming a division into elements (e.g., verbal or nominal groups) and operations (the rules of formation and transformation of sentences). However, in this theoretical model neither the elements nor operations corresponded to neural, anatomical or molecular structures, but only to a set of computational processes. Therefore, the Chomsky model explains human linguistic competence, but it cannot explain (and moreover Chomsky did not intend for it to do so) how, for example, auditory and motor areas interact to produce or understand sentences. Conversely, there are phenomena that can be explained by leveraging counterparts' materials, namely a model in which each element of the explanation can be identified with a real material structure, capable of causal interaction. The diagram of a combustion engine, for example, represents parts (piston, valve, etc.) which can be identified in a real engine. In Chomsky's model of explanation, by contrast, the elements and operations are described without being coupled to material structures; in practice we consider the adequacy of the functional relationships (input/output) of the model and those of the system under consideration.

Neuroeconomics, the science through which the study of brain processes, enables us to find new foundations for economic theories, and in recent years has received particular attention from researchers and scholars. This "new" science, placing itself at the intersection of neuroscience and economics, makes it possible

M. Graziano, *Epistemology of Decision*, SpringerBriefs in Philosophy,
DOI: 10.1007/978-94-007-5428-7_2, © The Author(s) 2013

to understand the neural basis of decision-making. It would also be able to offer privileged access to what, going beyond the formal models of economists, actually produces human behavior.

The Neuroeconomy, therefore, combines the knowledge of different approaches such as cognitive science, neuroscience and economics with the aim of studying the neural correlates of tasks in economic decision using brain imaging techniques, in particular the Resonance Functional Magnetic Imaging (fMRI). The latter uses the nuclear properties of certain atoms in the presence of magnetic fields. The technique came into use in the seventies in order to obtain detailed images of the anatomy brain. Through ultra-fast data detection techniques, it became possible to take pictures of very small increments of time (one hundredth of a second), which allowed them to follow in carrying out certain aspects of metabolism. Applied physiology of the brain, fMRI has allowed to display on a time scale very fine changes in oxygenation of cortical regions, changes that are considered to be closely related to the degree of activity in those regions. The magnetic properties of the hemoglobin molecules, which differ slightly depending on whether or not this is related to oxygen, are exploited for this purpose. Therefore, it is assumed that the fMRI images faithfully represent regional changes in neural activity which are apparent when a contrast is highlighted between regions that are rich in oxyhemoglobin, that is when blood flow is increased, and regions that exhibit normal blood flow.

There are strong reasons to be attracted to neuroscience. We owe our lives and all that we are to our brain and it is therefore legitimate to try to see how it works. But at the same time, we should wonder about the multiplication of the branches of scientific knowledge that use the prefix "neuro": Neuropolitics, Neuromarketing, Neuroethics, Neuroaesthetics, to name a few. To what extent are we dealing with really new areas of knowledge? To what extent are we instead faced with attempts by some researchers who work outside neuroscience to take advantage of the prestige and interest that neuroimaging techniques arouse? As specified by Paolo Legrenzi and Carlo Umiltà: "As all fields of human knowledge depend on the functioning of the brain, there is nothing to prevent the application of neuropsychology to disciplines such as economics, aesthetics, pedagogy, theology, etc. In fact, neuropsychology could have been (and was, to a certain degree) extended to these disciplines without the need to invent new terms by the pleonastic use of the 'neuro' prefix." (Legrenzi and Umiltà 2011, p. 9). And in particular for what concerns neuroeconomics: "It must be said, out of intellectual honesty, that neuroeconomy is very fashionable. Frequently the enthusiasm of the supporters of this new field of research leads them to reformulate what is already acquired knowledge thanks to the experiments conducted by psychologists, simply by substituting 'mind' by 'brain', and believing that in doing so they have enriched the reputation of this field of study." (Legrenzi and Umiltà 2011, p. 77).

Moreover, as with other methods that aim to explain the complex relationship between the brain and behavior, it must be said that brain imaging includes, within this theoretical framework, a number of ways to make observational data as transparent as possible and to limit the proliferation of possible explanations of the phenomena examined. Actually, in the case of business decisions, these are crucial features for the simple fact that models are applied behavioral interpretations of neurophysiologic data in nature. The methodology that is well accepted in neuroscience is known as "cognitive subtraction". To understand the role of subtractive methods one must take a step backward in the history of cognitive science up to Frans Cornelis Donders' studies (1869) in experimental psychology and reaction times of mental processes. In these studies Donders developed a rudimentary subtractive technique that would allow him to isolate the different operations performed by a subject during a particular cognitive activity. The basic logic of Donders' method consisted of the idea that the duration of a processing step can be measured by comparing the time required to resolve a version of a particular task (for example, press a button after the recognition of a particular visual stimulus), with a second version of the task that differed from the first only by omitting the step of processing (the pure reaction to the visual stimulus). The difference in time required to resolve the two versions represented the time that was spent in the stage of development taken into account. Donders' subtractive method was resumed and completed a century later by Saul Sternberg by the method of "additive factors". Stenberg (1969), in fact, demonstrated that the reaction times of mental processes were subject to variations as a result of the manipulation of certain variables: for example, in the case of the time necessary for a subject to determine whether a number belongs to a list first choice, influential variables may be the clarity with which the number is visually presented (therefore the clarity of the signal), the search in the active memory of the length of the stored list, the activation of the response of its compatibility, and so on. So assuming that a task that requires a sequence of operations, said Stenberg, it can be assessed to what extent manipulating variables are affecting the duration of individual operations. The importance of Donders' subtractive method and Sternberg's additive factors in brain imaging studies is due to the collaboration between the neurologist Marcus E. Raichle and psychologist Michael Posner (Posner and Raichle 1994). In the two authors' studies, the subtraction is used in the reconstruction of the factors that generate the neural activity detected. In fact, the principle adopted by Posner follows Donders and Sternberg and is subtracted, the activation maps, the values for the state control than those relating to the activated state. For example, according to the two authors, it can be assumed that you can isolate the brain activity related to a subject's passive fixation on a single set of visual stimulus by subtracting the control values recorded from the same subject while keeping his eyes closed. From the above, therefore, it must be inferred that the study of physiology and functional anatomy provides information independent of the psychological models, which departs from the architecture of cognition. Thus, when we use cognitive subtraction, if we make mistakes in the choice of the control task, we run the risk that research results will be worthless if not misleading (Legrenzi and Umiltà 2011).

These objections are legitimate. The researchers of the standard disciplines (economics, marketing, aesthetics, etc.) must not surrender to the charm that the brain images obtained through the use of functional magnetic resonance imaging can arouse. However, there is a simple reason for thinking that the neurosciences are necessary for the understanding of the processes of decision-making. The observation of human behavior “to the naked eye” is likely to reveal regularity, but does not tell us how these are derived from unobservable processes and structures and, in particular, from how they derive from the nervous system. The claim of neuroeconomics is thus first to open the “black box” showing how decision-making is carried out by the brain. To use a more properly philosophical jargon, neuroeconomics investigates decision-making in a “causal” outlook. It tries to show that as a high-level phenomenon, decision-making can be produced by underlying components and assets. Explaining decision-making, thus becomes first of all describing the mechanism that produces it. Neuroeconomics therefore wants to reintroduce to the study of decision-making, causal considerations which have been neglected or simply set aside, from standard economics. Unlike the neoclassical paradigm, which only provided the tools which allowed to predict the behavior, the claim of neuroeconomics is to explain it.

From the neoclassical point of view, the utility function is a formal term used to designate a preferable relationship between baskets of goods. To say that basket A is more useful than basket B, according to the neoclassical theorists, is simply affirming that there was a relationship between the two, that a rational agent—i.e. whose utility function that meets certain formal constraints—will select basket A rather than the basket B. On the contrary, neuroeconomics aims to understand the choice of basket A as a result of a neural mechanism that consists of components, activities or structures of the brain. The utility of a choice is not determined by formal preference relationships, but rather it is the result of a complex mechanism, which implicates, for example, specialized components in the production of pleasure, motivation, learning, attention or of cognitive control. Recent discoveries on the brain’s function allow us to reformulate, in a more realistic and accurate way, the fundamental insights surrounding the concept of utility function according to the standard assumptions which rest on the idea that the marginal utility is positive and decreasing, i.e. that $U' > 0$ and $U'' < 0$.

For example, an experiment Platt and Glimcher (1999) conducted on rhesus monkeys illustrates that the brain can actually encode this kind of function. The experiment was to teach the monkeys to choose between two bright spots (which appeared on screens placed to the right or left of the observation point the monkeys were). When, with a head movement, the monkeys made the right choice they got a reward (food). To maximize its usefulness, a monkey had to remember the probabilities associated with his earlier choice and of course the value of the reward. The experimental task was performed on blocks of 100 tasks. In some of these blocks, the probability of movement towards the right side, was steady at 80 % and at 20 % towards the left. In other blocks of the experiment, these probabilities were reversed. In this way, the anterior and posterior probabilities in each of these blocks had to be continually assessed by the monkey brain.

The goal was to change the probability of reward by maintaining constant visual stimuli and motor movements in such a way as to verify whether the activation of neurons in the lateral intraparietal region was related in one way or another with

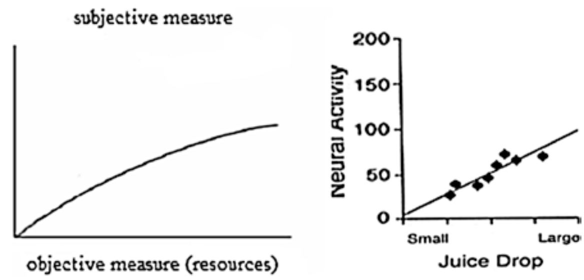

Fig. 2.1 *Source* Glimcher (2003)

the front or rear probabilities. Maintaining constant stimuli and the movements of the monkeys, the authors concluded that the probability of reward and their changes was related to neuronal activation. Platt and Glimcher interpreted this signal as evidence that first of all the apes were seeking to maximize their chances of gain.

After showing that the monkey brain can encode the probabilities associated with a reward, a second part of the experiment was devoted to ascertaining whether the monkey is able to do the same with the value of the reward. The authors, therefore, kept movements, the stimuli and the probability of reward constant (set at 50 %) and varied, from one block of tasks to another, only the amount of rewards (in some blocks, the amount of food obtained was 0.2 ml by looking left and 0.1 ml by looking right, in other blocks, the quantities were reversed). The results underlined how the value of earnings, when the elements were kept constant, were encoded by neurons practically speaking: the neurons were more active when the hope of gain was high. The fact that the brains of monkeys were able to simulataneously encode both the magnitude of the rewards and their probability, shows two things. Firstly, the concept of hope of utility seems to have found a neurological corollary. Secondly, that some concepts whose basis can be found in economic theory can be accommodated within the framework of neuroscientific analysis. Economics and neuroscience can therefore benefit from each other. Figure 2.1 shows a certain similarity between how economic science understands this basic concept and the data which were recorded from the brains of the monkeys.

What is surprising compared to the research of Platt and Glimcher is that subsequent studies using fMRI in humans come to quite similar conclusions. For example, activation of the posterior parietal cortex is correlated to the magnitude of monetary gain (Paulus et al. 2001) so that its anticipation would be positively correlated with the activation of the ventral striatum region (Knutson and Peterson 2005). Two important observations are indicated by Brian Knutson and Richard Peterson. Firstly, the ventral striatum is active only in anticipation of monetary gain; gain, loss, or anticipation of monetary loss apparently has no effect on this region. The second observation concerns the medial prefrontal cortex region. According to the two authors, this region keeps track of monetary gains and is deactivated when the monetary gain is zero. It should be noted, that it is neither

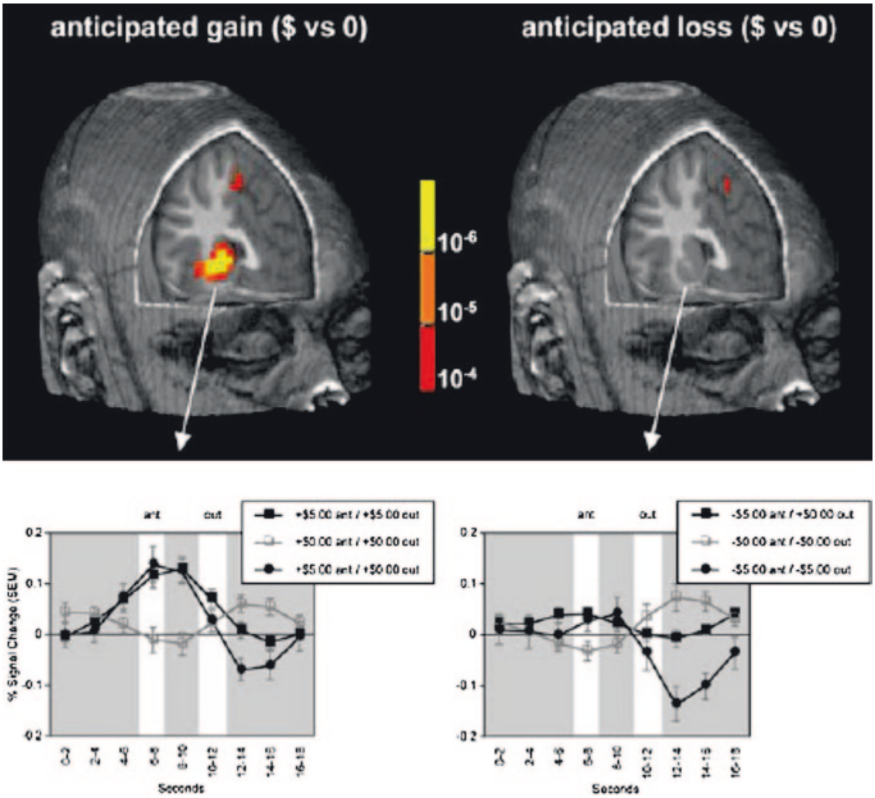

Fig. 2.2 *Source* Knutson and Peterson (2005)

active for anticipations (gains or losses), nor for losses: the prefrontal cortex is therefore not implicated in the anticipation of a reward previously known of.

Figure 2.2 shows the region of the striatum when activated by the anticipation of a monetary gain (left side of the figure) but not activated by the anticipation of a loss (right side of the figure). Therefore, the Knutson and Peterson framework for interpretation partially diverges from that of Platt and Glimcher. In fact, Knutson and Peterson interpret their results in light of Kahneman Tversky's prospective theory, namely that the prospects of gain and loss are not supposed be treated by the same neural mechanisms. In short, some interpretations and research allow insight to confirm some of the intuitions or precepts of the neoclassical approach (utility function, hope of utility as in Platt and Glimcher) but also some competing or complementary approaches underlined by Knutson and Peterson (prospective theory).

However, from both searches, what is revealed is that the concept of reward is a key construct in understanding human behaviour. Indeed, the ability to search and get rewards for their actions as a goal is, from a developmental point of view, essential to the successful breeding of complex organisms. As specified by Wolfram Schultz (2004), in an evolutionary sense, we can describe at least three functions of reward:

(a) Produce learning as it promotes the recurrence of the same conduct;
(b) Production and consumption behaviour approach;
(c) Generate positive experiences.

One of the most important discoveries made by Neuroeconomy relating to the important role of the concept of reward, is the one that has highlighted how this important process takes place in the neural structures situated in the most ancient part of the brain, namely "dopaminergic systems", which are involved in motivation and evaluation (Montague 2006). The first evidence that systems related to dopamine receptors are of primary importance were made by Olds and Milner (1954). Their research showed an increase of dopamine in certain brain regions in mice when they were involved in rewarding activities. This allowed a glimpse of a link between dopamine and hedonistic pleasure. In fact, these initial assumptions established a direct causal link between the feeling of pleasure and dopamine. This interpretation has now been called into question. Recent neuroscientific discoveries, rather, have demonstrated the link between dopamine and learning through experience, in which the dopaminergic response is transferred from an unconditioned stimulus (the reward itself) to a conditioned stimulus (the reward announcer). The dopaminergic neurons, which initially are triggered by the arrival of the reward, therefore, are activated later particularly in light of the conditioned stimuli (Schultz et al. 1997; Schultz 1998). The decision and the choices would then be oriented towards satisfying the dopaminergic neurons. As Roy A. Wise explains: "It is the return to a reward previously experienced that is the essence of habit and addiction. […] the return to a previously experienced reward involves the return to reward-associated landmarks as much as it involves return to the reward itself. […] The sounds, sights, and smells associated with the food are clearly predictors of reward, and the efficiency of the animal increases with the identification of more and more distal predictors of reward, predictors that guide the foraging and that are important for the "error signals" that guide corrections to the foraging path." (Wise 2002, p. 233). To ensure its survival, all species must be able to satisfy the vital functions, which vary from simply feeding to responding to aggression and reproducing. The brain's reward circuit allows the attainment of these objectives. The ventral segmental area, a group of neurons located in the centre of the brain, is particularly important in the operation of this circuit. It receives input from many other regions that inform the level of satisfaction of basic needs (or more specifically human needs). At the arrival of a signal announcing a reward, then after treatment of the sensory cortex, the activities of ATV are increased. This region then transmits this information by means of a chemical messenger, dopamine, that is "released" from the "accumbens nucleus" but also by the amygdale and the prefrontal cortex. The circuit formed by these structures selectively responds to primary rewards (food, herbs, sexual stimuli) and secondary rewards (money, music, cars, or social stimuli like faces attractive, pleasant tactile stimuli or emotionally connoted words). In particular, based on the experimental task, the accumbens nucleus has been demonstrated as having specific activation in relation to pleasant stimuli whose occurrence is predictable (on the contrary, it is not activated by stimuli that are always pleasant in nature but are not predictable). This feature seems to demonstrate the presence of a neuro functional structure capable of mediating the cognitive processing of stimuli in relation to their predictability. In fact, as it has been shown in research by Samuel McClure and colleagues (McClure et al. 2004), the complex-striatum accumbens seems to reflect the so-called prediction error (the difference between the probability value expected with respect to a certain reward and the actual value found) at the Neuro—functional level.

The daily experience of the subjects with expectations and their subsequent evaluation of errors with respect to certain contingencies that are repeated over time, allows individuals to learn more and become more refined, making them able to train the body to conduct a more appropriate and optimal response to its environment. This learning mechanism, functional optimization (maximization) of their behaviour in situations of uncertainty, seems to be mediated by the accumbens nucleus that plays a key role in determining the subjective behaviour in situations where the individual needs to evaluate alternatives with probability values and uncertain utility.

The amygdale, however, is the sub cortical region of the brain best studied for its role in emotional processing, with particular relevance to the formation of conditioned responses to dangerous stimuli. This thesis has been tested, among others, by Michael A. Paradise and colleagues (Paradiso et al. 1999), through an experiment that would establish the role of the amygdale in the evaluation of negative stimuli (presented in visual form through photographs). The author, noting a greater activation of the amygdale in negative evaluation of pictures (not experiencing the same activation in the evaluation of positive and neutral pictures), concluded with the assertion that the amygdale was involved only in the evaluation of a wide range of negative stimuli, but not in the assessment of positive stimuli. However, this interpretation is inconsistent with a growing number of recent results from research conducted both in animals (Davis and Whalen 2001; Everitt et al. 2000) and human beings, suggesting the involvement of the amygdale in the treatment of positive stimuli. For example, faces expressing joy (Gorno-Tempini et al. 2001), positive words (Hamann and Mao 2002) positive pictures (Hamann et al. 2002), excerpts of erotic videos (Beauregard et al. 2001; Karama et al. 2002). The role of the amygdale in emotional processing is therefore very wide, being involved in the development of both positive and negative emotions. The amygdale also seems related to the formation of estimates with respect to financial rewards, showing how this anatomical structure appears to correspond to the intensity of the stimulus, showing that peaks of activation are implicated as positive reinforcement (Hommer et al. 2003).

However, the brain area that undoubtedly represents an area of particular interest for the study of decisions in general and for studying higher cognitive processes is the prefrontal cortex, which is located in the frontal lobe and in front of the drive and premotor region. It is characterized by a late development from both phylogenetic trees; in primates, much of the ontogenetic myelination continues, in fact, even during the early years of childhood, and the dendritic and synaptic maturation reaches a stable level only in adolescence. In particular, this region provides important reciprocal connections with the sub-cortical regions (diencephalon, midbrain and limbic system) as well as with several cortical areas, mainly somatic, auditory and visual (Fuster 1989). Several routes also link the prefrontal cortex to the basal ganglia and the thalamus that are, themselves, the main regions involved in the motor activity of man. However, several studies have highlighted the different tasks performed by different areas that make up the prefrontal cortex, classically distinguishing three major prefrontal regions, which are linked to specific behavioural and cognitive functions. The first of these areas, the anterior cingulated cortex (Brodman area of 24, 25 and 32) seems to be implicated in the control of autonomic functions, the initiation of the response, intention, the treatment of the conflict or error, and the allocation of cognitive resources (Bush et al. 2000; Holroyd and Coles 2002; Botvinick et al. 2004). The orbit frontal cortex (Brodman area 12 and 13) seems to play an important role in the functions instead of needing a front control of the limbic system, such as inhibition, the encoding of the motivational value of an object or of a stimulus, decision-making and control action based on reward, impulse control and interference, mood and social behaviour (Bechara et al. 2000; Rolls 2000). Finally, the lateral prefrontal cortex, particularly the dorsum—lateral (Brodman areas of nine forty-six) is usually associated with functions involved in executive control, such as changes or representations of all the current rules (set-shifting), the resolution of complex problems, the recovery of memories in long-term memory, organization and strategies of working memory (Goldman-Rakic 1987; Fuster 2001; Watanabe et al. 2005). The prefrontal cortex, along with other structures such as the anterior cingulated cortex and the insula, seems to be involved in the regulation of social interactions and behavioural conduct so that they are handled in a timely manner.

In conclusion, the approach aims to analyze neuroeconomy in the more substantial economic structures of the mind through the study of the brain during its operation. Many neuroeconomists are convinced that a better understanding of the mechanisms by which the brain assesses and compares alternatives and various forms of rewards can assist us in learning what determines our choices and our behaviours. As we shall see later, neuroeconomics does not attempt to analyze the brains of individual decision makers, but it tries to extend its domain to the interactions between different agents using game theory.

2.2 Game Theory and Neuroscience

In recent years, game theory models have been used by neuroscientists convinced that by so doing new findings could emerge that might prove useful both to economists and scientists that study the brain. From this perspective it is interesting to consider the works of James K. Rilling, who along with his colleagues used a version of the Prisoner's Dilemma in order to observe the cerebral activity of 19 participants (11 girls and 8 boys) when they played both against human counterparts and against a computer. Rilling found that participants cooperate more (in the 81 % of cases) when their adversaries are human beings. The researchers subsequently identified the cerebral regions that were activated in the cooperative acts and in the defection acts in order to analyze their dopaminergic impact, recording a BOLD activation (Blood Oxygen Level Dependent Signal) of part of the striatum and of the prefrontal ventromedial cortex. This activation, it was also noted, was not meaningful when player A had to deal with a computer: the nature of the interaction therefore plays a relevant role in the BOLD activation (Rilling et al. 2002). Other research has confirmed the assumption that a positive social interaction with others is particularly gratifying. For example, Tania Singer and her colleagues (Singer et al. 2004) use the Prisoner's Dilemma themselves in order to demonstrate that simply seeing the face of a person who had previously cooperated activates reward circuit areas of the brain.

Hence, these neurological observations complete empirical observations previously reported: adopting cooperative behavior and obtaining mutually beneficial behavior will produce an activation of the reward circuit implying a sort of gratification. Moving in the same direction, but also useful for testing the role of emotions in economical decisions, is the research carried out by Alan Sanfey and his colleagues (Sanfey et al. 2003) who used the "Ultimatum Game—UG" to evaluate the neurological foundations of economic decisions. The protagonists are two players who are given the chance to share a certain amount of money. One of them, player A, makes the offer, the other, B, can accept it or not. If B accepts the offer, the money will be shared as proposed, but, if he refuses it, both remain without anything and the game ends there. It would seem natural that player A would make the most favorable proposal for himself, and that B would accept it, rather than remain with nothing (on the rational basis that the money he was offered is in any case preferable to zero). This, however, is not the case. Regardless of the amount at stake, the proposed sharing is most of the times fair, and not only. Low offers have approximately 50 % chance of being rejected, a detail which demonstrates how under certain circumstances people are motivated to refuse an economic benefit.

But, why does this happen? Excluding the possibility that that the players did not understand the rules of the game or that they have difficulties in conceptualizing the match played in just one move, it is meaningful that when confronted with an unfair offer, the rejection is often associated to an angry reaction to an offer perceived as unfair. In fact, studies carried out in several countries, have demonstrated that the vast majority of people offer, within the third day, half the amount (Camerer and Loewenstein 2002). Moreover, it has been remarked that the rejection of an offer, by the second player, is usually combined with a feeling of anger (Pillutla and Murnighan 1996). The expression of feelings, during the game phases, has its relevance indeed: an offer can be rejected, for example, in order to keep a good reputation in the game (the acceptance of low offers damages the

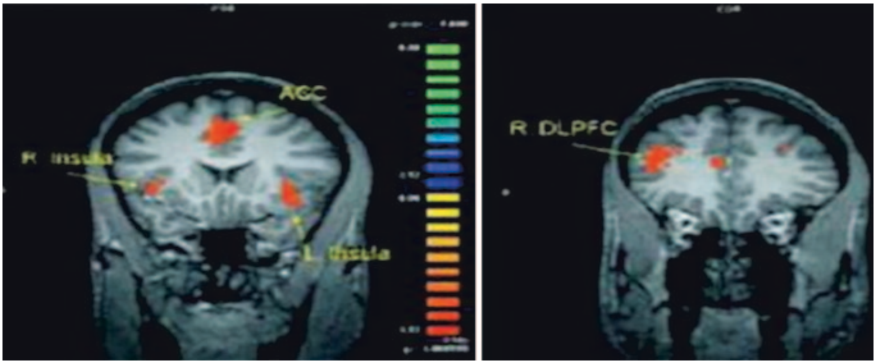

Fig. 2.3 Activated cerebral regions during the Ultimatum Game (Sanfey et al. 2003)

player's reputation, increasing as a consequence the chance of being proposed further low offers (Nowak et al. 2000). The only exceptions, where people made and accepted low offers, were found in some autistic adults: hence, paradoxically, a rational behavior is observed, mostly in subjects having cerebral deficits (Camerer 2003). It is Sanfey and his colleagues who are credited with using the "Ultimatum Game" in economic studies (Sanfey et al. 2003). In fact, thanks to the use of the fMRI technique (functional Magnetic Resonance Imaging), Sanfey studied the cerebral activation of 19 subjects taking part in the "Ultimatum Game", paying more attention to the cerebral activations of the participants who were proposed low offers (20 % of the sum at stake).

Assuming that such an offer would cause a conflict between the emotional desire of not accepting on the one hand, and the will to accumulate as much money as possible on the other, Sanfey and his colleagues identified the areas of the brain potentially involved in these mental processes. The authors noticed, in particular, a higher activation of three cerebral areas: the anterior cingulate cortex (ACC), the anterior bilateral region (right insula and left insula) and the dorsolateral prefrontal cortex (DLPFC) (Fig. 2.3).

During their experiment, Sanfey and his colleagues, asked their participants to complete 30 rounds, each of 36 s, of the Ultimatum Game (10 with people, 10 with a computer and 10 as a control). Before the partner's offer was disclosed, a picture of the sparring partner (whether a person or a computer) was shown on a display. Researchers found very similar results to those mentioned above: the acceptance rate was negatively correlated with the unfairness degree of the offer. Interesting to remark was also that the players' behavior was different depending on whether the offer was made by another player or by a computer. The rejection in fact was meaningfully higher when the offer was made by a human being. This demonstrated how the players were sensitive, not only to the amount offered, but also to the context (i.e. whether the low offer was made by a human adversary or by a computer). In particular, the regions of the insula were those showing the greatest activation. From this, Sanfey deduced that, being the region of the insula associated to negative emotional states, its activation, following an unfair offer, had to be interpreted as an effect of a negative emotion.

Hence, according to the authors, the activation (or not) of this region, allows the prediction of the rejection of an economic offer; on the contrary, the other cerebral region of interest, the DLPFC, which is supposed to be linked to the will of maximizing the monetary benefit (hence the rational part), on its own (not being correlated with the acceptance rate of low offers), does not allow the prediction of the behavior. Nonetheless, the authors suggest that it can enter into competition with other cerebral regions (the insula for example) for control over decision-making. So, to sum up, the findings show that if the activity of the insula is higher than the activity of the DLPFC, the offer is rejected; on the contrary, if the activity of the DLPFC is relatively higher, the offers are accepted. For the purpose of our study, we are simply interested in highlighting the role played by emotions in decision-making in an economic situation, and emphasizing how these are now (rightly so) being taken into consideration in the studies of this field.

Moreover, it is worth mentioning how corroborations to Sanfey's study have come from the evolutionist perspective as well. We refer to an experiment carried out by Sarah F. Brosman and Frans De Waal (Brosman and De Waal 2003), where the authors demonstrated that when a monkey, after working hard in order to be rewarded, gets less than another monkey (or when the other monkey did nothing to deserve a better reward), has an emotional reaction inducing it to reject a deal that, he would have almost surely accepted if he had been alone. In other words, the monkeys seem to measure, as men do, their rewards in relative terms: comparing their level of offer and of benefit with that of others. The studies we have just described have had the merit of including emotions as a factor in the study of economic decision-making behavior, and have highlighted how man cannot be represented as a calculator with no limits, recalling the inevitable complexity requested by the interaction among different levels of study.

A clear understanding and investigation cannot operate without research competences and methodologies that concern different disciplinary sectors, in which neuroscience clearly plays a major role. The integration of these disciplines does not just generate an exchange of knowledge and expertise, but entails the change and specification of the object of study as well. It is in this sense that the *Homo Economicus* is replaced by the *Homo neurobiologicus* (the neurobiological man), whose behavior derives from a neurobiological development able to generate sentiments, belief, actions and the capacity to make decision. Moving beyond Colin Camerer and other strong ideas put forth by other neuroeconomists that claim that just the measurements of cerebral activity during decision-making allows us to verify economic notions (primarily utility), there can be other fundamental reasons according to which the proposal of crossing economic models with psychological and neurobiological data might prove to be beneficial.

An example comes from Ernst Fehr, and it consists in exploring systematically the neuronal basis of altruistic behaviors in order to discover whether they are conditioned by strategic attitudes and considerations rather than by purely pro-social inclinations and emotions (Singer and Fehr 2005). Another interesting question could be that of understanding whether the perception of others' intentions is really an indispensable resource for the execution of a rational strategy in a game, or if on the contrary, the development in infancy of this cognitional capacity would instead have the tendency of making us deviate from the Nash equilibria. Thanks to the emergence of such research programs, in fact, it is being demonstrated more and more, in a convincing and rigorous manner, how the classic universal material self-interest model presents rather frequent violations when put under the test of experimental analysis.

In particular, with specific regard to game theory, the idea that individual preferences could be, besides being driven by the pursuit of a personal interest, also not directly self-interested, is being consolidated. There are, in fact, some social preferences (preferences of a social type) that are positively or negatively influenced by the behaviors, preferences, or intentions of other subjects. For example, thanks to the Ultimatum Game it has convincingly been demonstrated that decision-makers are in general not self-interested, being disposed to punish adversaries who make offers perceived as being unfair, despite this being costly for them. In fact, the profile of the subjects that take part in the UG is, as emphasized by Sacco and Zarri (2003), self-interested from a motivational point of view, but in fearing that a low offer might be rejected by the adversary, the player takes precautions and makes offers which could be perceived as fair and that could reasonably be accepted by the other player.

It appears therefore, worthy of note, how the tendency to reason strategically at the level of the motivational states of adversaries, emerges from the experimental evidence. In the UG, the proponent, in fact, does not offer large amounts to the decision-maker because of an innate sense of equity, but rather on the basis of a belief leading him to predict that the other player, from a motivational point of view, is not acting out of self-interest either and that the offer made is satisfactory. Hence, the beliefs of subject A, with regard to the motivations of player B, influence the choices of A himself, through a judgment that A formulates with regard to B's intentions. If, in fact, A is an individual who decides to cooperate in a conditioned way, or in other words in virtue of an expectation that the other player will cooperate (positive reciprocity), but that will instead refuse to cooperate in the case the other does not (negative reciprocity), in the case in which his beliefs suggest to him that B intends to defect, A will trigger his choice to defect rather than cooperate.

Therefore, if in a certain interaction, both A and B are driven by reciprocity based on their intentions, then their preferences will be directly interdependent. From what has been said an interesting parallel can be established between the psychological mechanisms involved in decision-making within game theory and the psychological mechanisms involved in the resolution of the ToM (Theory of Mind) tasks. To this purpose let us consider respectively, the classic version of the unexpected transfer of the false belief task (Wimmer and Perner 1983) along with the two games we have discussed so far: the Ultimatum Game and the Prisoner's Dilemma.

The concept of theory of mind (ToM) was introduced in 1978 by David Premack and Guy Woodruff, in an article provocatively entitled "Does the chimpanzee have a theory of mind?" (Premack and Woodruff 1978). Their study intended to demonstrate the capability of primates to understand mental states and hence to predict human behavior in situations finalized to a purpose. The two primatologists concluded in their essay that monkeys are endowed with "Machiavellian intelligence" allowing them to plan behaviors of alliance or of deception in order to achieve their goals. The theory of mind was hence defined as that capability of understanding, of inferring and attributing mental states (desires, beliefs) to oneself and to others, in order to understand and predict their behaviors.

The issue concerning the presence or not of a mentalizing capacity in species other than human, has aroused and continues to arouse debate. It seems that some species of great apes attain a certain level of comprehension of the thoughts of others without having a theoretical capacity of the mind as complex and developed as that of the human mind. But beyond this very specific problem, researchers in the psychology of development rapidly became interested in the mechanisms and the development of this cognitive function in human beings.

The very first studies in this direction were those conducted by Heinz Wimmer and Josef Perner. Their purpose was to test the ability of children to attribute a false belief (Wimmer and Perner 1983). The classic version of their experiment is as follows. A first child, named Max, puts some chocolate in a green box and goes out of the room. His mother appears on the scene and moves the chocolate to a blue box. At this point Max re-enters the room and looks for the chocolate. Now imagine asking a second child who witnessed the whole scene where in his opinion Max will look for the chocolate or where he'll think it is. What will the second child answer? Any normal adult understands that Max's knowledge will not allow him to look for the chocolate in the right place and, thus, in order to respond correctly must adopt Max's point of view of the false belief and not the one of reality.

The two researchers hypothesized that until the age of 4, children tend to answer adopting the point of view of reality, hence stating that Max should look for the chocolate in the blue box. In the several replications of the test that have followed the original, the notion of false belief has become a central criterion to establish when children completely develop a theory of mind structurally similar to that of adults. Nevertheless, success in a task of false belief can depend upon several factors rather than just theory of mind (for example, the comprehension of counterfactuals and the ability to carefully follow complex situations); moreover, false belief is not the only existing mental state, and studies of the mental states of others suggest a gradual development of this ability, characterized differently at different ages.

Mentalizing is therefore a complex ability which is made up of several components such as shared attention (which develops in the first year of life) and other more complex abilities (such as counterfactual reasoning for example). Nevertheless, despite all these limitations the scientific knowledge obtained through the test of false belief still remains strong (Leslie 2005); although, what the mechanisms are that drive children to identify themselves with others and to reflect on their mental states continues to be a matter of dispute among researchers of the ontogenetic development of theory of mind.

The ToM requires a structured set of innate abilities that gradually allow humans to construct representations of internal states of other individuals of their species (but also others), in order to be able to predict the behaviour of others. But regarding the specific issue of how information processing occurs which then leads to the understanding and prediction of behaviour, two different theories have tried to give an answer: the theory of theory and the theory of simulation. According to the theory of theory, every human being is built through the experience of a folk theory (folk psychology) about mental representations, motivations, goals and emotions that lead to different behaviours. The theory of theory presupposes that every individual has available a rather sophisticated inferential system, under which one foresees what will happen starting from one's data of experience, which are the axioms of the folk theory. The peculiarity of the theory, then, is precisely to consider the capabilities expressed by the psychology of common sense as part of a theory. Obviously, this is a theory that for most individuals would remain in an implicit state. The theory of fact does not suggest that everyone is able to reflect on individual aspects of the theory or check back to make the right inferences. But the systematic aspect of the predictor would remain, even if latent, in each of us. The limits of the theory of theory seem very obvious. It presupposes the existence of a folk theory of behaviour which, in turn, seems to presuppose the accessibility of complex inferential processes by one's consciousness and one's ability to abstract representation of rules. The theory of simulation, which opened in the mid-eighties by Robert Gordon, however, rejects the use of such a chain of logical inferences to explain inter subjectivity, which it considers too rich, wasteful, uneconomic, preferring a more modest model based on imitation and imagination, and on the belief that beyond individual differences all humans are endowed with a mind that works similarly in similar circumstances.

Therefore, in order to pass a task of false belief, one has to temporarily suspend the knowledge of factual reality (known to the participants of the test and to the experimenter) in order to represent the mental state of the protagonist or, in the terms of the simulationist approach, one has to mentally "put oneself in the other's shoes". In other words, the task of false belief requires making a prediction with regard to a behavioral outcome, in response to the question: "where will the child look for the chocolate?" In order to correctly answer this question and predict the child's decision, one has to set aside one's knowledge for a moment, that is knowing perfectly well where the chocolate was moved to, in order to represent the state of knowledge of the other, who instead does not know that the chocolate was moved and who thus has a false belief that will drive his behavior.

If the theory of theory shows its major limitation as being entirely directed towards the other person, whose mind is trying to understand the operational concept in terms of chains of arguments and logical inference, the simulation shows the theory to be, on the contrary, too ego centred, since it plays all its cards on the ability of the ego to imagine itself in place of the other. On closer inspection, however, that selfishness is at the same time a de-centring of the ego, since one has put aside one's beliefs in merely assuming the beliefs of others. In doing so, the theory of simulation, which is able to understand and predict behaviour based on the closeness of the affinity between the observer and the observed agent, fails to explain how it is possible to predict the behaviour of agents outside of oneself. Moreover, it seems quite plausible that the obvious purpose of a simulation is not understanding how you would act in the place of someone else, but just imagining how one would behave just the other would. In other words, when one puts himself in the shoes of another he does not bring along his own mind, but he tries to see the world through the eyes of the one whom he is imagining (Perconti 2003).

The current state of the art in the debate of theory of mind is most certainly polarized by the positions of some of the supporters of the theory of pure theory and other extreme proponents of the theory of simulation. However, there have been different attempts at hybridization between the two poles and reciprocal concessions on certain specific points in the theoretical modelling. For example, the simulationist is willing to admit that, in one's understanding of others, one ends up with recursive logical-inferential arguments that refer to a system of conceptual knowledge about the way the other's mind works (which even does not depart from the fundamental point that this happens only after one takes on the other's first-person perspective and one just puts himself in the other's shoes with an imaginative non-inferential logic-simulation).

At this point we want to hypothesize that, in a subject who has to predict the decision in a game, a mechanism triggers a meta-representation of the mental state of the other, similar to the one we have just discussed for the tasks of false belief. Let us use the example of the Ultimatum Game: the player has to think of the sum he should offer to the other player, that is whether to propose a high, low or fair offer, because if the other player should reject it, he will lose everything himself and will not gain the remainder. So, in thinking of how much he should offer, the player has to evaluate the other person's attributes. If he knows him, we can imagine he will evaluate some character traits, such as, whether he is a "take it all" subject, or in other words, a type of person that no matter what the offer will be most probably will accept; or if he is a subject that carefully evaluates received offers and therefore will not likely accept an offer under a given threshold, in such a circumstance the proposal of a low offer would mean to risk leaving with empty pockets.

We can similarly imagine what is in going on in the mind of the player receiving the offer. If he receives a low offer he could for example think: "he made me a low offer because he thinks I'll accept it anyway?" In the most extreme case he could think: "rather than letting him gain so much, I will reject the offer. I'll gain nothing, but neither will he". The same mechanism seems to take place in the Prisoner's Dilemma, the other game we have discussed. When we are questioned, we have the chance to decide whether to cooperate or not; nonetheless, this choice cannot be made without taking into consideration the move made by the partner: in fact, if you choose to cooperate you risk a lot because you appear to trust the other from which you expect a reciprocity of the behavior. If the partner decides not to cooperate (and to act like a spy) everything will be at the expense of the one who cooperated (Marchetti and Castelli 2006).

Nevertheless, from what has been discussed, it is clear that it is not enough to put yourself in someone else's shoes and to wonder what you would do under the same circumstances in order to adopt a "winning" strategy. To this first simulation level, we have to add a second higher level, made up of the set of mental assumptions that enable us to imagine the "internal life" of the other, for example, the counterfactual imagination (i.e. the ability to assume that things can take a course that is different from the actual one). In other words, we need an "integrated" approach of simulation and common sense psychology. Therefore, it is necessary to integrate the theory of mind and simulationism in a further theory where simulation processes are the basis for behavioral prediction. At the same time, a major role could also be played by the processes regarding our psychological intuitive knowledge and those regarding the particular mind of the person we are trying to simulate (Perconti 2003).

The theory we are alluding to is the one known by the name of "social cognition", whose main components are: mind reading, imitation, shared attention, empathy, language, self-awareness, one's own ability to lie and to detect the lies of the others, and imagining the point of view of others. The set of these components make the typical forms of human society possible, rendering possible the complex cultural practices which earn man a central place in the animal kingdom (Ferretti 2006).

2.3 The Role of Social Cognition

In general, social cognition is the process that allows people to think about and give meaning to themselves, others and social situations (Fiske and Taylor 1991). In particular, it concerns the ways in which we form an impression (positive or negative) of the personality, role and identity of others. The notion of social cognition must therefore take into account a number of features of human cognition: (1) the recognition by the subject of an actor who works in an environment intentionally; (2) the consideration that the beliefs and representations of others are related to their actions; and (3) changes in the constitution of beliefs and representations about the goals of a subject. Therefore, the way we define other people affects our social interaction, but at the same time, the other is also influenced by social interaction, that is, we form the view that is both an effect and a cause of social interaction.

Usually, people think that social cognition primarily serves a practical purpose (Fiske 1992). According to studies of social cognition, people must balance their impressions of others with the requirements of appropriate social interaction as a result of the limitations of the cognitive system. As a result, people engaged in social interaction are usually "motivated tacticians", who mostly use "quick and dirty" judgments to conserve cognitive capacity but can be trained to use strategies yielding thoughtful and detailed impressions (Fiske and Taylor 1991). An example of this comes from the issue of trust (distrust) versus social reputation. In social interactions, trust (distrust) of others plays a crucial role. People expect others to be competent, friendly, honest and trustworthy. When we need to make quick decisions, we use heuristics, stereotypes, habit patterns, and other gimmicks to evaluate others. Many experimental data show that one of the key assumptions in our social interactions is the so-called "positivity bias" (positive bias); that is, most people expect from others, ceteris paribus, a kind of benevolence. People emphasize the pleasant and avoid the unpleasant; they communicate good news more often than bad and are more likely to judge unpleasant events as pleasant (Rothbarth and Park 1986). Similarly, in a phenomenon known as "positive bias of the person" (person positivity bias), people are evaluated more favorably than corresponding abstract entities; that is, students evaluate individual teachers more favorably than the courses they taught, or individual politicians more favorably than their political party in general (Sears 1983). All of this, of course, is reflected in language; in fact, in most languages, positive terms outnumber negative terms (Zajonc 1998).

The assumption of positivity is clearly present in our expectations about other individuals and events. In social situations, positivity encourages interaction with our fellow humans and the environment.

Compared to trust, reputation seems to need more time to build. In fact, a person's reputation (positive or negative) must have time to stabilize before it can become a form of capital for related actors. This stabilization reduces uncertainty in the expectations of those who attribute to the other qualities of a certain type (for example, absence of opportunistic behavior or skills and ability to carry out commitments made previously). Social reputation can be represented as the stable expectations of a plurality of agents (another fundamental difference compared to simple trust, which is generally based on dyadic relationships) relating to certain qualities. Moreover, unlike trust, due to risk aversion (Savadori and Rumiati 2005), a bad reputation spreads faster than a good reputation, and we tend to accept information about a person's bad reputation without independently verifying it; the limiting case is represented by prejudice. On the contrary, a good reputation tends to be accepted only after the positive qualities of the subject have been carefully checked.

From the foregoing it becomes clear that indefinite cooperation in a repeated game (such as a Prisoner's Dilemma) is not in any way "irrational". On the contrary, over time, a reliably cooperative player builds a good reputation that may prove useful in later negotiations.

An alternative explanation has been proposed by neoclassical economic theory that is based on the interpretation that agents do not want to behave in a selfish way because they are driven by altruistic behavior to maximize the profits of others. This interpretation, however, conflicts with a large body of experimental evidence showing that the percentage of cooperation in a finite Prisoner's Dilemma game decreases progressively (Guala 2006). The apparent paradox of the experiments conducted using models of game theory in which there is a high level of cooperation in the early rounds that decreases gradually with the progress of the game can be easily explained by the fact that people tend to initially build a good reputation, but as they begin to play the game, they come to understand what is the best move, altering their choices by the end of play. This explanation has led to the creation of games built on models with “error and learning” in which subjects are initially involved in the experiment above the Nash equilibrium, but gradually align themselves to it (Kreps et al. 1982). Rules of conduct are similar in the case of the Ultimatum Game. The answers provided by participants in the game seem to contradict the predictions of neoclassical theory, with bids higher than the percentage of waste and significantly high offers. However, in this case, it was noted that if we introduce to the game an element of competition among the players, the kind of behavior predicted by neoclassical models is restored (Guala 2006). Therefore, although the hypothesis that individuals are always rational maximizers of their utility is discredited by some experimental data, it is true that the neoclassical predictions are supported when you are able to interpret the true reasons that motivate individuals to behave in certain ways. These behaviors may seem irrational at first glance, but they possess some degree of meaningfulness when the goal is to achieve certain economic or social objectives, without aspiring to optimality. Thomas Ulen writes in Rational Choice Theory in Law and Economics that the agents are probably influenced by social context such that, even if you expect them to behave in a selfish way to maximize their own interests, they instead cooperate unexpectedly: “These experimental results present a puzzle for rational choice theory: why do people cooperate when there appears to be a rational basis for not cooperating? One possibility is that people start any given interaction from the presumption that it is better to cooperate than not; they continue to cooperate until the evidence shows this to be ill-advised; and then they quit cooperating.” (Ulen 1999, p. 803). In the literature, imitation re-establishes a sort of balance between individual rationality and social rationality. Imitation, which has been a topic of interest to great researchers across the social sciences, from Piaget in psychology to Keynes and Hayek in economics, is still central to the interests of many social psychologists, ethologists and philosophers of the mind. Despite the many differences between the various approaches, we can break down imitation into two main categories: automatic or unconscious processes present in human beings from birth and imitative and reflective processes that qualify as rational imitation. Contrary to Piaget’s theories, recent studies have shown that humans have a rudimentary representational capacity at birth, measured in terms of deferred imitation, which develops later during interactions with their environment.

Imitation is primarily concerned with a few superficial behaviors and is then extended to phenomena that are not directly visible (for example, the intention behind an action), even before the appearance of language. This has been highlighted through a series of experiments that show this ability in children at 14 months of age. In one such study by Meltzoff (1988), an adult sat at a table, on which a lamp was placed. As the child watched, the adult bent over the light and turned it on by pressing it with his forehead. Two-thirds of the children imitated this behavior a week later. In fact, they did not use their hands to light the lamp, although that would have been much easier. This experiment can be interpreted in two ways: the first interpretation is that children have not considered the adult as an intentional agent and have simply imitated his behavior without really understanding his goals, and the second interpretation is that children have understood the intentions of the adult and therefore used the same means to fulfill the same purpose (in this case, the goal may have been to turn the light on with one's forehead). This second conclusion was subsequently confirmed by other experiments (Zelazo and Lourenco 2003), suggesting that children's imitation is more sophisticated than mere reproduction of an observed behavior; on the contrary, imitation in prelinguistic children is already a selective and interpretive process (in this sense, we speak of "rational imitation"). These studies in developmental psychology have been corroborated by some neuroscientific studies of strategies and intentions and their products in the form of actions in the environment (Chaminade et al. 2002). Imitation is therefore a feature that enables human beings to interact with the world beginning at birth. Over the years, to the extent that the child learns to distinguish the content of knowledge, imitation becomes less and less automatic and more the result of deliberation. Therefore, imitation pervasively characterizes the social dimension of human existence by intervening at multiple levels. In fact, mutual mimicking increases in more intimate relationships and occurs when a dyad must interact successfully for the sake of the group; if the dyad fails to meet this goal, imitation increases in future attempts (Ferguson and Bargh 2004). Mimicry appears to be a key tool by which human beings attain social satisfaction. Far from being a simple, passive registration process (as evidenced by developmental psychology research), imitation plays an important role in understanding other individuals.

We have seen so far that imitation plays an important role in the characterization of the processes of social transmission in human beings. We just have to show that imitation may prove to be an interesting approach of some emerging phenomena within the decision-making. We begin with a very intuitive definition of imitation that will serve us later for a more precise definition. Consider two agents, O (observing subject) and M (model), involved in a single activity, such as playing football. Suppose that O recurrently observed that M is better at drawing penalties. Then O will seek to know the reason: If he comes to think that M's success depends on a trait T O can imitate, such as following a specific diet or wearing a certain type of shoes, chances are that O mimics M when trying to acquire this property. In this case, the interest of O in T is not direct, but is aroused by the fact that M has this feature, after the success that O binds to M in his personal reading of world events. Thus, we find the typical triangular structure of imitation, Model-Based Object, most of which is located in the literature of modelling and imitation with a slightly different approach in René Girard (in the terminology of the latter the triangle is composed of Subject-Intermediary—Object; Girard 1961).

Imitation thus proceeds in three stages:

(1) choice of a model, according to a certain criterion that belongs to the subject, in the course of an act of observation,
(2) selection of a feature that belongs to the model which the subject contributes to the fulfilment of this criterion,
(3) an attempt by the party to copy the feature.

These three stages suggest several characteristics of imitation. The first is that the choice of model depends on the particular feature that will be imitated, since the latter is not necessarily known at this stage, but is a general criterion that allows the assessment of individuals, which reflects the purpose or intent of the subject. Moreover, the subject engages in an act of imitation only if it believes that the model is better than him on one of the dimensions covered by this policy. This therefore requires a comparison between the model reflective of the subject and himself, on this criterion. The second characteristic is that the subject has to identify some characteristics of its potential model and infer the one or ones that are involved in its positive assessment of the model. This step will therefore be favoured by the subject's ability to categorize and reflect on these various categories. This requires in particular meta-cognitive and reflective skills. Finally, the third stage brings into play the capabilities of the individual learning of the subject from a model.

From the foregoing, imitation (like reading the mind of others outlined in the previous paragraph), denotes different levels and modes of interaction by which the individuals establish sensory ties with the other. Also speaking to the mimetic phenomena, another capacity, especially in recent years that has been the subject of much research (including neuroscience) is undoubtedly Empathy, namely that special ability to understand how others feel, their feelings, their emotions, and in a sense to share them. At a superficial level it seems normal to think that mind-reading, imitation and empathy differ significantly. In each of these three types of interpersonal relationships, it seems we are confronted with apparently different objects: in the case of imitation of the actions of someone else, translating the observed into executed movements; in the case of mind reading, recognizing the reasons which have produced a behaviour; while in the case of empathy, experiencing the emotions and feelings of others. Based on these superficial differences, it may appear legitimate to assume that imitation, mind reading and empathy depend on different mechanisms. In what follows we will support the contrary view, namely that imitation, mind reading and empathy share many more of those things that separate them, primarily the same specific functional mechanism: the embodied simulation (Rizzolatti et al. 1996).

2.4 Empathy Basic and Empathy Re-Enactive

The word "Empathy" appeared in Anglo–Saxon languages in the early twentieth century and translates the German word "Einfühlung" (which appeared in proto-romantic Germany from the turn of the late eighteenth century until the 1860s), which was originally used to characterize a form of aesthetic experience in which the subject is projected in the act of imagining a work of art. This aesthetic theory was developed by Lipps (1903), who later extended the use of the term "Einfühlung" to the domain of interpersonal relationships, designating Empathy as the ability one has to put oneself in the other's shoes. If one succeeds, then he/she has a empathetic personality. However, when we move away from the use of the term in ordinary language, we can easily see that the term Empathy is used as an "umbrella name" for a whole class of terms only seemingly overlapped such as, for example, identification, imitation, emotional contagion, sympathizing. It should, therefore, be distinguished, even if at times it becomes somewhat arbitrary insofar as this term appears to be floating, but so that it still designates real differences between the actual phenomena observed. For example, it is undoubtedly an excellent thing to distinguish between Empathy and emotional contagion. The latter term, for the most part, is the phenomenon of propagation of an emotion from one individual to another. This phenomenon is well known from the psychology of crowds and is also found in children who hear to the cries of another baby and respond by starting to cry themselves. It is generally agreed, that emotional contagion is characterized by a form of non-differentiation between oneself and others, both in the case of children, where the basis for this differentiation is not yet sufficiently well placed, and in the case of phenomena of crowds, where we witness a form of temporary abolition of the distinction of the individual self that merges into a collective I. Empathy differs from sympathy, however, in another dimension. In both cases, the distinction of self/others is preserved. The essential difference between the two phenomena, according Wispé (1986), lay in the purposes intended in each. With sympathy, as indicated by its etymology, we assume that the emotions are felt by another, sharing one's pain or, more generally, his emotional experience. Sympathy brings into play altruistic purposes and presupposes the establishment of an emotional connection within what it covers. Empathy on the other hand is a process of imagination that seeks to understand the other and not to the establishment of emotional bonds. Empathy can certainly entertain sympathy, but this is not a necessary consequence of the first. Empathy can also help for altruistic reasons. Understanding the regret another feels does not mean that one agrees with it or is trying to reduce it. As Wispé said: "The object of Empathy is understanding. The object of sympathy is the welfare of others". (Wispé 1986, p. 318). Therefore, Empathy is used for the most part today when we have the ability to know the other perspective, and assume the other's role: the ability to put oneself in another's place, to see the world as seen by another while retaining, however, always a clear separation between who has experienced and who empathizes.

Having outlined these differences, the crucial point becomes how to characterize this particular form of understanding between individuals. The fact that we are all more or less able to interact with others in itself does not excuse us from trying to understand how we do it. The answer to this question came from two Italian neuroscientists at the University of Parma, Giacomo Rizzolatti and Vittorio Gallese, who in the early nineties found a particular population of deputy vasomotor neurons in the cerebral cortex that process information concerning the behaviour of others, the so-called “mirror neurons”, forcing everyone involved in inter subjectivity to confront a new paradigm centred on the hypothesis of a neuronal correlate of empathy and the idea of a biological basis of sociality. In particular, through a series of recordings of individual neurons, the two researchers discovered that in a sector of the premotor cortex of the monkey (area F5) there are neurons which are activated during both the active execution, by the monkey’s actions such as grasping or manipulating objects, and during the observation of similar actions performed by another agent (another monkey or a human). These neurons were then called mirror neurons to emphasize their dual nature of action and action observation. Therefore, from a neuronal point of view there is some difference between when a subject completes an act and when he/she simply observes another individual accomplishing the act. Over the past 10 years, numerous neuroimaging research studies have allowed us to locate the operation of the human mirror system in different situations and to verify the existence of an anatomical correspondence between the cortical circuit involved in visual-motor transformation as studied in macaques and that also exists in humans (Rizzolatti and Sinigaglia 2006). When a human being observes an action, there is an activation of complex brain areas that include the occipital, temporal, parietal lobe of the visual areas, and two areas with a mainly motor function: the rostral portion of the parietal lobe and the lower part of the pre-central gyrus together with the rear part of the inferior frontal gyrus (IFG). Comparative studies have suggested that at the end of the human homologue of area F5 of the monkey can be identified in the pars opercularis (Brodmann area 44), an area that has always been assumed to be related to speech production. However, more recent studies have actually shown that it is also implicated in the processes of organizational actions, foreshadowing of motor acts and understanding of actions, demonstrating how the area 44 is not exclusively used for the language (Iacoboni et al. 1999; Buccino et al. 2001).

The discovery of mirror neurons has corroborated the theory of social psychology and the automaticity and pervasiveness of imitation and empathy. It is not only the vision of the motor patterns of others that trigger the activity of mirror neurons, the observation of emotional response also generates a mirror response: when seeing other people’s emotions, the observer can determine the activation of the cortical region that is normally active when the observer tests that emotion (Rizzolatti and Vozza 2008). In people, the two emotional experiences, direct and observed, cause an activation of the same areas of the cortex. The properties of these neurons reflect not only movement but also dissolve the problem of the emotions of other minds: humans are social animals because our neural activity is coordinated with, and depends on the neural activity of people who are round. This feature leads to intentional consonance; by virtue of the mechanisms of mirroring, the other is experienced as another self (Gallese 2007). To understand the importance of reflection at the cognitive level of brain activity it is interesting to report an experiment that explored the role of mimicry in the expression of the cognitive skills of general knowledge (Ferguson and

Bargh 2004). It was asked of healthy adult human volunteers to answer general knowledge questions from the board game "Trivial Pursuit". Those who before being subjected to questions had been engaged for thirty minutes reading articles on Hooligans showed significantly lower performance than subjects who had read for thirty minutes narratives of scientists or writers. Specifically, according to Vittorio Gallese: "If—through simulation—we go even for half an hour in the cognitive framing of a Hooligan, our pre-existing cultural knowledge produces a poorer performance than when we enter a period equivalent to the cognitive framing of an intellectual" (Gallese 2007, p. 203). Mirror neurons allow an implicit form of understanding the actions of others, establishing a sort of bridge between the observer and an actor of an action and enabling very important imitative behaviour. This neuronal identity allows us to explain the particular affinity that was observed experimentally between the behaviour of a player and behaviours that he observed in another player, but there was a limit because of the strictly motor function of these mirror neurons. What interests us, beyond the actions performed by players, are the intentions that guide them and make them intelligible to others. If the character of such knowledge among the players can be explained by the mirror system, its intentional content, however, is elusive. Some experiments that link the intentionality of an action to its contextualization suggest that this system could also be extended to the recognition of intentions (Iacoboni et al. 2005).

However, the answer to our question is not simply limited to players and to identifying the intentions of others, but that their intentions are also identified by other players. The mirror system allows us to recognize, for example, that another person is angry or wants to grasp an object. However, we must also be able to explain the subsequent conduct of an individual in a complex social situation (such as those alleged by game theory). It is not enough to recognize that the other is angry but one must try to realize why he's angry, tracing and updating it in oneself the reasons for his state of mind, reconstructing the context that caused the error. Along these lines, a very interesting position is advanced by the philosopher Stueber (2010) who proposes to distinguish between empathy (basic) and re-enactive empathy. The first defines the unmediated mechanisms that underlie our theoretical capabilities to perceive and recognize other creatures directly as essentially thinking beings like us. Re-enactive empathy, in contrast, contains our cognitive capacities and resolutions that allow us to actualize and imitate the thought processes of others in our mind and then to design the complex social behaviour of other people's behaviour as rational agents that act on the basis of reason. While basic empathy can be explained by considerations related to neurological mechanisms (mirror neurons) and the results derived from experimental psychology, empirical, re-enactive empathy requires a philosophical reflection from the vantage point of the investigations on the nature of rational action (folk psychology). Specifies the author: "…. use the term "empathy re-enactive" to referrer to the required Simulative Capacities Because, in contrast to Goldman, for example, I see the Notion of rational agency to be at the centre of our folk psychological practices. Rational agents are not creatures who act merely, because something is happening inside them. Rather, they are able to take a reflective stance towards their own agency and to take ownership of their action in terms of their reasons for acting. Being able to do this, I maintain requires that One's agency is potentially intelligible to oneself in terms of one's understanding of the world, one's long-term plans, and the standards and rules of conduct that one is committed to…" (Stueber 2012, p. 59).

While it is true that we cannot explain all the complex inter-subjective experiences only with the help of mirror neurons, actually something has been accepted by the discoverer himself when he says: "Mirror neurons do not empathize anything, simply for the fact that they are not people. They know nothing about actions, intentions, beliefs, emotions, know only the exchange of sodium and potassium, and electrical impulses. Give the subject a direct mechanism, automatic, non-predicative and non-inferential simulation". It seems to Stueber, that we are asking a bit too much of re-enactive Empathy, so that, with good reason, Shaun Gallagher writes about Stueber, "Third, a problem that is both terminological and conceptual. Stueber (2012) uses 'empathy' as just another term for social cognition, even the most basic, default, and automatic resonance-based mode of social cognition. To associate it with mirror neuron activation, as Stueber and many others do, is to make empathy the automatic default mode of social cognition. On this view, if it is true that my mirror neurons activate whenever I see you engage in intentional action, and this generates some basic form of empathy, then, in effect, I cannot help but empathize with you—even in the case of seeing you engage in what I take to be an obnoxious action. Stueber (2012) also considers high-level re-enactive processes as a further variety of empathy. In either respect, however, there is no difference between empathy and ordinary, everyday understandings of others, or what some theorists call 'mindreading.' Other theorists, including myself, however, and in some agreement with ordinary language, suggest that empathic behavior involves an other-directed feeling of concern or interest, distinct from both sympathy and mindreading. Is this not another variety of empathy to be considered?" (Gallagher 2012, p. 65).

Gallagher rightly makes mention of the term social cognition because this expression of empathy is more comprehensive since it refers, as we have seen, to all the cognitive processes that mediate interpersonal relationships (not only empathy, but also mind-reading and imitation, along with language and consciousness). Trying to capture the essence of social behaviour and avoid the temptations reductionism implies necessarily taking into account the different levels of description of the matter: from neural dynamics to social dynamics, the adaptive logic of its path of development in the phylogenetic and its appearance during the ontogenetic development of different species. With regard to our specific field of investigation, therefore, it seems once again that when we want to understand and anticipate the behaviour of others, calibrating their actions accordingly, we need all the components of social cognition with the result that we need to integrate different investigation techniques in a multidisciplinary perspective.

2.5 Doubts, Feasibility and Future of Neuroeconomics

Neuroeconomics seems to have secured an important place in research, and numerous laboratories and centers have been created to develop research and studies in this new field. Now that its concepts and methods have been clarified, it has become possible to examine and question the relationship that it might establish with traditional economic research. Many economists, in fact, have expressed doubts about the relevance of neuroeconomics and in general of neuroscience in traditional economic studies. The more radical criticism is to deny the relevance of the neurological and psychological causes of behavior to economics. This criticism has been made forcefully by Gul and Pesendorfer (2008). In fact, they argue that economics is not normally interested in such matters as neuroscience and psychology: economics pursues different objectives and must use abstractions suitable to its own purposes. Gul and Pesendorfer are particularly important in this debate because they put forward a distinction between “true utility” and “choice utility”. In the opinion of the two authors, economics should be interested only in the second (as is the standard economic theory), by defining the utility function of an agent only in relation to the notion of choice. To say one option is more useful than another, in this sense, simply means to say that a rational agent would choose it. Gul and Pesendorfer oppose, by contrast, the tendency of many neuroeconomists wanting to return to the idea that there is “true utility” linked to pleasure or to some other motivating factor, which is similar in many respects to the cardinal concept present in the early utilitarian philosophers. The neurological or psychological variables that cause choices are not simply relevant from an economic point of view. The economist does not need to know, for example, that to prefer x to y is motivated by the pursuit of happiness, a sense of duty, a religious obligation or an impulse (Gul and Pesendorfer 2008). The only thing that should matter to economists is the relationship of preference revealed by the choices.

This defense of the standard approach does not imply that the economy is entirely isolated from psychological research. However, it must incorporate only relevant data, i.e. data on choices and behaviors. The data produced by psychology or by experimental economics can thus be used as a model to measure and predict future choices or balances. In contrast, data or variables unrelated to choice, such as neuronal activity data, must not find space in the theoretical model. To paraphrase Hilary Putnam, it is irrelevant to the theory proposed by Gul and Pesendorfer whether the brain is made of gray matter or Swiss cheese.

There are many ways to respond to criticisms made by Gul and Pesendorfer. The first of these, advanced by Ross (2005, 2008), is to argue that neuroeconomics is not limited to the approach that they criticize. According to Don Ross, in fact there are at least two ways to study neuroeconomics. The first one which he called “behavioral economics through the scanner”, which is also the best known, seeks to identify the neural mechanisms which produce decision-making. But there is also a second way, less known but no less important, namely "neurocellular economics". This approach aims to apply methods and economic techniques to unconventional systems: networks of neurons. The originality of this position consists in stating that the neoclassical approach to economics is better suited to the construction of behavioral models of neural networks than to those of traditional economic agents (individuals). Neural networks, in fact, are less likely to violate the axioms of the theory of revealed preference, which makes neoclassical theory particularly useful to building models of their behavior. This paradigm assumes, for example, that a given cortical area or a dopamine circuit acts as a market whose behavior can be modeled thanks to the theory of general equilibrium. Neurons, instead, are regarded as agents whose activation reveals preferences. The application of neoclassical theory to neuronal circuits is promising: it allows us to question an idea according to different abstractions used by neuroscience and economics, related to the various objectives they pursue (Bourgeois-Gironde and Schoonover 2008).

However, this approach shows a greater enrichment of neuroscience through economic theory and, therefore, the criticism raised by Gul and Pesendorfer remains intact with regard to the relevance of neuroscience for economics itself. A decisive answer from this point of view comes from Colin Camerer, according to whom economic theory is not obliged to disavow its neoclassical matrix, it has only to gain and draw from the conceptual and empirical wealth of psychology and neuroscience to build models of adequate explanation. In his reply to the text of Gul and Pesendorfer, Colin Camerer, recognizes that "since behavioral economics is meant to be a generalization of rational choice theory which incorporates limits on rationality, willpower and self-interest in a formal way. These generalizations allow the possibility that conventional rationality is an adequate approximation, and often permit a parametric way to measure the "degree" of limitedly rational behavior and its economic impact." (Camerer 2008a, p. 44). In particular, according to Camerer, neuroeconomics adds a neuro-causal component to the theory of rational choice, but his approach is not fundamentally different from that of neoclassical theory. Let's give an example. The theory of revealed preference claims that it is necessary to draw a utility function from the observation of the behavior of the agents, which allows one to predict their future behavior. If you choose, for example, to eat, in different contexts, an orange rather than an apple and if your choice resists various forms of interference and is firm, we can conclude that you prefer oranges to apples. Neuroeconomics only expands the data and the variables considered in the previous example to include the behavior of neuronal components. Consequently, the central question, one to which Gul and Pesendorfer do not respond, is whether the data and neuronal activations allow as is always stated by Camerer to improve the ability to understand and predict the choices, while maintaining the discipline of mathematics and the use of behavioral data (Camerer 2008a).

Certainly, a good part of economics remains heavily influenced by the instrumentalist interpretation which sees in the forecast (rather than in the explanation) the highest goal of the models it develops. But even if one accepts this interpretation, one must bear in mind that the safest way to predict the behavior of any system is to have a good understanding of its functioning. Gul and Pesendorfer emphasize strongly that the economist should rely on "the data on the choices" to measure the reliability of its models. The problem is obviously that it is impossible to interpret the behavior of an individual such as, for example, his choices without a variety of underlying assumptions, which are based in particular on how this individual will represent the options and their probability (the beliefs) and on the way in which he attributes them a value (his wishes). Without these options on the way in which states of mind change, a set of behaviors cannot simply reveal the relationship between preferences: we would be faced with the practical problem of indeterminacy of Quine-Duhem (Guala 2005). This affirms that the economist does not need to know if the agent "is motivated by the pursuit of happiness, a sense of duty, an obligation or a religious impulse" (Gul and Pesendorfer 2008, p. 24), which introduces considerable confusion. In fact, even at the level of ordinary language, the concepts "the pursuit of happiness" "sense of duty", "religious obligation" or "on impulse" define determining characteristics of the utility functions of agents. To say that someone acts on "impulse", for example, means attributing to the agent a utility function which is sensitive to the passage of time. So there is behavioral data relevant to determining whether or not someone acts on impulse: just measure the impact of the passage of time on his behavior. At the same time, it is interesting for an economist to know if a person acts out of a "sense of duty" or "interest", because "the sense of duty" refers to the idea that his preference would remain the same even if his interest changed. One can reply by saying that the economist does not just want to know whether or not a given behavior is or is not sensitive to the passage of time, but he wants to be able to quantify the impact of the time variable in a formal model. Or again, he wants to know exactly to what extent the sense of duty makes the behavior resistant to the interference of personal interests. Once the question is presented in this way, however, it becomes much less interesting to try to determine if the economist has the right to use a concept such as "impulse" to describe a utility function or whether he should limit himself to formal representations of preference relationships. You could support the second option, as do Gul and Pesendorfer, but this implies a narrow vision of the conceptual means available to the economist. Moreover, this also leads the economic theorist to move away from the concrete practice. Contrary to what happens in theory, the naive and scientific psychology terminology is ubiquitous for describing utility functions.

However, if we admit that neuroeconomics may be relevant, we must understand if it is already or if it can potentially become so. On this subject, opinions differ. Some, like Harrison (2008), appear largely pessimistic due to the considerable gap between the claims and achievements of neuroeconomists. Others, like the aforementioned Colin Camerer, are quite optimistic and believe that the contribution of neuroeconomics is already real. Harrison's skepticism derives from a characteristic of neuroeconomic research referred to above. We saw, in fact, that neuroeconomics is often limited to a search for neural correlates of known and studied phenomena in psychology and behavioral economics: for example, the aversion that comes from receiving unfair offers, the pleasure of punishing incorrect behavior, the suffering related to social exclusion. But, if a phenomenon has already been studied on a psychological and behavioral level, what need is there to know the neural correlates? Harrison raises precisely this point. Let us consider variables such as risk aversion or aversion to ambiguity. Not only are the phenomena well-known on a psychological level, but there are numerous experimental methods which allow determining the value of the agents. What then can neuroeconomics add? Ideally, it can help to identify the components of the mechanisms involved in the production of these two phenomena. For example, Brian Knutson and colleagues (Knutson et al. 2005) showed that the medial prefrontal cortex (CPFm) is activated in proportion to the objective probability of gain. For their part, Ming Hsu and colleagues (Hsu et al. 2005) have established that there is a positive correlation between the ambiguity of a choice and activation of a circuit connecting the amygdale and the orbital frontal cortex (COF). These studies show, therefore, that two phenomena are identified as distinct from behavioral economists—risk aversion and aversion to ambiguity—are produced by distinct mechanisms. What can the economist derive from this discovery? Does it allow him to better assess the utility functions of the methods traditionally used by behavioral economics? Once again Camerer (2008b) seeks to give an answer to this question by trying to clarify what the economist can derive from the analysis of neurons. Camerer's idea is to use data on the brain to decide empirically between theories that are difficult to distinguish by means of tests of market forecasting (using the usual data). Camerer's claims therefore remain relatively modest.

Neuroeconomics allows us to see if the models are on the right track, by verifying for example, if separate variables correspond to components or separate processes. Neuroeconomics, according to Camerer, however, may be destined to play a greater role, leading economists to consider the importance of variables hitherto neglected. In this case, it would cease to trail psychology and behavioral economics, and would drag the economy directly to a beneficial progress.

If the idea of producing decision-making neuroscience is seductive, it is however, necessary to recognize the limitations of the research conducted so far. Neuroeconomics meets the same methodological difficulties as cognitive neuroscience. As shown in detail by Craver (2007), all neuroscience methodology has limitations. The visualization of brain imaging and recording unitary neuronal activity, for example, are used to locate areas in which brain activation coincides with decision-making, but are not enough to prove the causal implications of these surfaces. Further arguments must be added to the techniques of localization. Some authors have criticized the tendency of neuroeconomics to draw hasty conclusions based on very limited experimental data (Rubinstein 2006; Harrison 2008). If the criticism is justified, it is more about the rhetoric that often accompanies these studies, and the neuroeconomists are certainly not to blame for wanting to promote their research programs. The limits of the various approaches do not justify the abandonment of the search for causal mechanisms, but encourage us to be prudent and to remember the importance of the integration of the different methods.

References

Beauregard, M., Lévesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *The Journal of Neuroscience, 21*, 1–6.
Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex, 10*(3), 295–307.
Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539–546.
Bourgeois-Gironde, S., & Schoonover, C. (2008). Une hybridation de l'économie et des neurosciences a-t-elle un sens. *Revue d'économie politique, 118*, 35–50.
Brosman, S., & De Waal, F. (2003). Monkeys reject unequal pay. *Nature, 425*, 297–299.
Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400–404.
Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215–222.
Camerer, C. F. (2003). *Behavioral game theory: experiments in strategic interaction*. Princeton: Princeton University Press.
Camerer, C. F. (2008a). The case for mindful economics. In A. Caplan & A. Schotter (Eds.), *The foundations of positive and normative economics* (pp. 43–69). New York: Oxford University Press.
Camerer, C. F. (2008b). The potential of neuroeconomics. *Economics and Philosophy, 24*, 369–379.
Camerer, C. F., & Loewenstein, G. (2002). *Behavioral economics: Past, present, future*. Pasadena: Mimeo, Division of Humanities and Social Science.
Chaminade, T., Meltzoff, A. N., & Decety, J. (2002). Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *Neuroimage, 15*, 318–328.
Craver, C. F. (2007). *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press.
Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry, 6*, 13–34.
Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431, (originally published in 1868).
Everitt, B. J., Cardinal, R. N., Hall, J., Parkinson, J. A., & Robbins, T. W. (2000). Differential involvement of amygdala subsystems in appetitive conditioning and drug addiction. In J. P. Aggleton (Ed.), *The amygdala: Neurobiological aspects of emotion, memory and mental dysfunction* (pp. 354–390). New York: Wiley-Liss.
Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trend in Cognitive Sciences, 8*(1), 33–39.
Ferretti, F. (2006). The social mind. In M. Marraffa, M. De Caro, & F. Ferretti (Eds.), *Cartographies of the Mind* (pp. 295–308). Dordrecht: Springer.
Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laser photo. *Journal of Personality and Social Psychology, 63*(6), 877–889.
Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill International Editions.
Fuster, J. M. (1989). *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe*. New York: Raven.
Fuster, J. M. (2001). The prefrontal cortex-an update: time is of the essence. *Neuron, 30*(2), 319–333.
Gallagher, S. (2012). Three questions for Stueber. *Emotion Review, 4*(1), 64–65.
Gallese, V. (2007). Dai neuroni specchio alla consonanza intenzionale. Meccanismi neurofisiologici dell'intersoggettività. *Rivista di Psicoanalisi, 53*(1), 197–208.
Girard, R. (1961). *Mensonge romantique et vérité romanesque*. Paris: Grasset.
Glimcher, P. (2003). *Decisions, uncertainty, and the brain: the science of neuroeconomics*. Cambridge: MIT Press.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology* (pp. 373–417). Bethesda: American Physiological Society.

Gorno-Tempini, M. L., Pradelli, S., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., et al. (2001). Explicit and incidental facial expression processing: An fMRI study. *Neuroimage, 14*(2), 465–473.

Guala, F. (2005). *The methodology of experimental economics*. Cambridge: Cambridge University Press.

Guala, F. (2006). *Filosofia dell'economia. Modelli, Causalità, previsione*. Bologna: Il Mulino.

Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin & A. Schotter (Eds.), *Foundations of positive and normative economics* (pp. 3–40). New York: Oxford University Press.

Hamann, S. B., Ely, T. D., Hoffman, J. M., & Kilts, C. D. (2002). Ecstasy and agony: Activation of the human amygdala in positive and negative emotion. *Psychological Science, 13*(2), 135–141.

Hamann, S., & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *NeuroReport, 13*(1), 15–19.

Harrison, G. W. (2008). Neuroeconomics: A critical reconsideration. *Economics and Philosophy, 24*, 303–344.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679–709.

Hommer, D. W., Knutson, B., Fong, G. W., Bennett, S., Adams, C. M., & Varnera, J. L. (2003). Amygdalar recruitment during anticipation of monetary rewards: An event-related fMRI study. *Annals of the New York Academy of Sciences, 985*, 476–478.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science, 310*, 1680–1683.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J., & Rizzolatti, G. (2005). Grasping the intentions of others with one's owns mirror neuron system. *PLoS Biology, 3*, 529–535.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526–2528.

Karama, S., Lecours, A. R., Leroux, J. M., Bourgouin, P., Beaudoin, G., Joubert, S., et al. (2002). Areas of brain activation in males and females during viewing of erotic film excerpts. *Human Brain Mapping, 16*, 1–13.

Knutson, B., & Peterson, R. (2005). Neurally reconstructing expected utility. *Games and Economic Behavior, 52*, 305–315.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *Journal of Neuroscience, 25*(19), 4806–4812.

Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated Prisoner's Dilemma game. *Journal of Economic Theory, 27*, 245–252.

Legrenzi, P., & Umiltà, C. (2011). *Neuromania: On the limits of brain science*. Oxford: Oxford University Press.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends Cognitive Science, 9*(10), 459–463.

Lipps, T. (1903). *Ästhetik. Psychologie des Schönen und der Kunst*. Hamburg-Leipzig: Voss.

Marchetti, A., & Castelli, I. (2006). *Teoria della mente e Neuroeconomia: quali affinità?*. Milano: Università del Sacro Cuore.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 304*, 503–507.

Meltzoff, A. (1988). Infant imitation after a 1-week delay-long-term-memory for novel acts and multiple stimuli. *Developmental Psychology, 24*(4), 470–476.

Montague, R. P. (2006). *Why choose this book? How we make decision*. New York: Penguin.

Nowak, M. A., Page, K., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science, 289*, 1773–1775.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology, 47*(6), 419–427.

Paradiso, S., Johnson, D. L., Andreasen, N. C., O'Leary, D. S., Watkins, G. L., Boles Ponto, L. L., Hichwa, R. D. (1999). Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *American Journal of Psychiatry, 156* (10), 1618–1629.

Paulus, M. P., Hozack, N., Zauscher, B., McDowell, J. E., Frank, L., Brown, G. G. (2001). Prefrontal, parietal and temporal cortex networks underlie decision-making in the presence of uncertainty. *Neuroimage, 13*(1), 91–100.

Perconti, P. (2003). *Leggere le menti*. Milano: Bruno Mondadori.

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of Ultimatum offers. *Organizational Behavior and Human Decision Processes, 68*(3), 208–224.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*, 233–238.

Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York: Scientific American Library.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515–526.

Rilling, J. K., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron, 35*, 395–405.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*(2), 131–141.

Rizzolatti, G., & Sinigaglia, C. (2006). *So quello che fai. Il cervello che agisce e i neuroni specchio*. Milano: Raffaello Cortina.

Rizzolatti, G., & Vozza, L. (2008). *Nella mente degli altri. Neuroni specchio e comportamento sociale*. Bologna: Zanichelli.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*(3), 284–294.

Ross, D. (2005). *Economic theory and cognitive science: Microexplanation*. Cambridge: MIT Press.

Ross, D. (2008). Two styles of neuroeconomics. *Economics and Philosophy, 24*(3), 473–483.

Rothbarth, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*(1), 131–142.

Rubinstein, A. (2006). Comments on behavioral economics. In R. Blundell, W. K. Newey, & T. Persson (Eds.), *Advances in economic theory* (pp. 246–254). Cambridge: Cambridge Univeristy Press.

Sacco, P. L., & Zarri, L. (2003). Complessità motivazionale, interazione strategica e gestione del conflitto. *Etica ed Economia, 1–2*, 73–89.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*, 1755–1758.

Savadori, L., Rumiati, R. (2005). *Nuovi rischi, vecchie paure. La percezione del pericolo nella società contemporanea*. Bologna: Il Mulino.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*(1), 1–27.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioral ecology. *Current Opinion in Neurobiology, 14*(2), 139–147.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593–1599.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology, 44*(2), 233–250.

Singer, T., & Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *American Economic Review, 95*(2), 340–345.

Singer, T., Seymur, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not the sensory components of pain. *Science, 303*, 1157–1162.

Stenberg, S. (1969). The discovery of processing stages: Extensions of Donders methods. *Acta Psychologica, 230*, 276–315.
Stueber, K. R. (2010). *L'empatia*. Bologna: Il Mulino.
Stueber, K. R. (2012). Varieties of empathy, neuroscience and the narrativist. Challenge to the contemporary theory of mind debate. *Emotion Review, 4*(1), 55–63.
Ulen, T. S. (1999). Rational choice theory in law and economics. *Encyclopedia of Law and Economics*, 790–818. http://encyclo.findlaw.com/0710book.pdf.
Watanabe, M., Hikosaka, K., Sakagami, M., & Shirakawa, S. (2005). Functional significance of delay-period activity of primate prefrontal neurons in relation to spatial working memory and reward/omission-of-reward expectancy. *Experimental Brain Research, 166*(2), 263–276.
Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128.
Wise, R. A. (2002). Brain reward circuitry: review insights from unsensed incentives. *Neuron, 36*, 229–240.
Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology, 50*(2), 314–321.
Zajonc, R. B. (1998). Emotion. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 591–634). Boston: McGraw-Hill.
Zelazo, P. D., & Lourenco, S. F. (2003). Imitation and the dialectic of representation. *Developmental Review, 23*(1), 55–78.

Part II
The Biological Approaches

Chapter 3
Evolutionary Economics and Biological Complexity

3.1 Biology and the Economy

According to some theorists, economic phenomena that determine wealth are prioritized in research. This "substantial" definition is often contrasted with the "formal" definition by Lionel Robbins (1935), stating that economic science owes its unity and specificity to the fact that it studies contradictory choices. The agent has limited resources to distribute between different objectives, and he must choose to sacrifice some objectives for the benefit of others. This definition, by intrinsically linking economics to the theory of choice, has led economics, as a science, to focus on human behavior as a relationship between ends and means. In other words, economics is the science of choosing the most advantageous option among several alternatives, depending on one's context and needs. The aim is therefore to make the most favorable long-term choice. However, in everyday life, many possibilities arise when we make an important decision; thus, the consequences of our choices are not clearly predictable or known a priori. In this complex and uncertain environment, our choices have consequences that become more or less attractive over time. Experience allows an agent to accumulate knowledge about the consequences of different choices and to develop preferences for some. If an agent understands the consequences associated with each choice, uncertainty decreases, and decisions are driven by the agent's preference or by risk aversion.

Consequently, we can speak about the decision-making process as an adaptive process that manifests itself in fundamental steps as an individual negotiates a complex and dynamic environment. Decision-making allows us to control events and make choices that are profitable. When making a good choice, an agent learns about a new situation, considers prior errors while progressing forward, and modifies the actions that have previously proven inappropriate. In contrast, when agents make choices that have systematically unfavorable consequences for themselves and/or for their group, their decisions are said to be inappropriate (or pathological). The most complicated aspect of decision-making is that the values that we assign to each choice are influential. The preference for one action over another

M. Graziano, *Epistemology of Decision*, SpringerBriefs in Philosophy,
DOI: 10.1007/978-94-007-5428-7_3, © The Author(s) 2013

depends on the way the brain interacts with external and internal stimuli. These same mechanisms are necessary to constrain our choices by anchoring them to our biological needs. The ability to assign values to different choices becomes more sophisticated over time, and assigning values to choices represents a valid strategy for solving problems that affect our survival and need for flexibility in dealing with the uncertainties of the surrounding environment (Montague 2006). Living things also tend to develop increasingly complex structures by virtue of accumulating changes or adjustments that alter their previous configurations.

This approach constitutes a critique of "orthodox" economics, which considers data at the micro level, including all possible production plans. In the history of economic thought, since the studies of Cournot (1838), who was the first economist to apply the physical model and its powerful means of investigation—differential calculus—in economics, the act of economic exchange has been viewed as analogous to the transmission of movement through a machine. This physical–mathematical explanation of economic phenomena was developed largely by Marginalists Leon Walras and Stanley Jevons. The economic orthodoxy, trapped in an idealized vision of the world, is committed to revealing an increasing number of truths about how the world works, rather than developing possibilities about how the world might work. Not only is there no such thing as theorems in orthodox economics, but the field cannot address questions about, for example, the process that allows us to structure the catalog production process, the factors that are likely to lead to change, or the reasons for differential access to economic realities. More generally, the field of economics has been criticized of building models that are insufficient to explain reality. Models that drew inspiration from physical facts had the major limitation of not taking into account that phenomena and processes of change happen in a historical context and thus cannot be contained by rigid and outdated mathematical models.

It was therefore apparent that contemporary economic theory would have to stop relying on the theories of classical physics and turn to biology, the only discipline capable of accurately describing economic development (Hodgson 1993). However, as Dobzhansky argued (1973), we cannot understand the important concepts in biology without the idea of Darwinian evolution by natural selection. This idea has demonstrated great explanatory, and sometimes predictive, power, bringing together a wealth of experimental data within a single theoretical framework through which we can interpret the diversity of life and its many transformations. However, before evolutionary biology and economics could establish a fruitful interdisciplinary dialogue, both fields had to reach a certain level of maturity. This dialogue became possible through a few basic assumptions that the two sciences share:

1. Rarity. Individual economic/biological research resources are limited.
2. Competition. Individuals are competing to acquire limited resources.
3. Maximizing. The individual maximizes its value (utility/attitude) in the attainment of resources.
4. Emergency collective. The processes (market/evolution) are not direct agents but still compete against one another.

This approach considers not the superficial similarity between two constructs, but the use of mathematical models in the description of behavior, in particular the "optimization theory" and "game theory". The latter, for example, was first used by economists, and was later used to add a time dimension to models of the evolution of animal behavior. These models, in turn, allowed economists to model the dynamic interactions between economic agents. For models to be translated from one discipline to another, it was necessary for both disciplines to have the same predictive or explanatory goals. This assumption holds in the present study.

In economics, however, the term "evolution" is used in different ways, and sometimes has the opposite meaning as it does in biology. There is indeed a sense in which the term is not used in the same way in the two fields. In such cases, the biological terms are used simply to criticize or justify standard economic theories or notions. Another promising approach is the evolutionary paradigm, which allows researchers to specify a system's characteristics and then determine its development in terms of breeding and selection. Although several authors have referred to the economy as a biological system, Alfred Marshall puts forth an enhanced evolutionary paradigm. In fact, he was the first to argue in favor of an "organic economy" based on the assumption that economic and biological phenomena share a large number of similarities, from their complex and organic nature to their involvement in an evolving world. Both qualitative and quantitative influences on both kinds of systems imply that future events never reproduce prior conditions. His original idea of economic dynamics (or rather of economic development) derived from a biological model was long ignored until other areas of research had been permeated by evolutionary theories. Its main merit was that it used the theory of evolution not to justify a posteriori the basic postulates of neoclassical economics (as Alchian and Friedman did later), but on the contrary, to create new directions for the discipline. As Schumpeter pointed out, the greatest merit of the work of Alfred Marshall is that it provided a foundation for subsequent important ideas. More than any other economist, Marshall was credited with having paved the way for later economic research (Schumpeter 1954). In particular, Marshall integrated the spatial dimension, by translating it in terms of adaptation to an environment, and the temporal dimension based on the common descent of organisms of the same species.

3.2 Economic Progress and Evolutionism

In mainstream economic theory, change is interpreted as a fundamentally exogenous phenomenon that is caused by external shocks to the economy, is largely unpredictable, and generates changes in the data on which decisions are based. The theory must merely note changes in tastes, technology, capital resources and/or the expectations of market agents without altering its analytical approach.

The research program initially formulated by Marshall (1890) in Principles of Economics after he decided to devote himself to economic studies showed his intention to rethink this classic theoretical framework from the inside out. He aimed to update, extend and formalize the Millian-Ricardian traditional theoretical paradigm. Eager to remedy what he considered two specific limitations of classical economics—the use of a conceptual language away from the market and the lack of quantitative analytical instruments—but reluctant to introduce radical changes in method or theoretical perspective, Marshall could only approach the problem through incremental innovations and the progressive loss of meaning. He then decided to combine the classical economic categories, appropriately revised in content, with some new and empirically relevant concepts. The exploitation of biological evolutionary theory is the foundation on which this new research program was centered. In addition to the Principles, Marshall's later work, Industry and Trade (Marshall 1919), is marked by the chronic dissatisfaction that is revealed every time Marshall found analytical economic instruments not to be immediately applicable to practical situations. As we have said, however, building a satisfactory economic theory of evolution does not only involve searching for similarities with the biological notions of natural selection, of change, and of the unit of selection, but should include considerations beyond the simple concepts. In fact, applying the principle of natural selection to the evolution of patterns of social organization and industry, both of which are theoretically relevant in economics, requires upstream integration and an understanding of Darwinian evolutionism as part of a global issue.

Alfred Marshall worked from this perspective, asking how Darwinian evolution could be used in the study of economic phenomena. In fact, Marshall is unique in being one of the first economists to take explicit recourse to a dual approach, including both static and dynamic components, to the study of economic phenomena. Both approaches are structured around different reference systems, namely the physical model and the biological model, respectively. Marshall writes: "There is an analogy between the relatively strong first stages of an economic reasoning and the static step in the physical sense. But there is a step so profitable in the last stages of an economic reasoning and dynamic methods? I think not. I think that in advanced stages of reasoning, biological analogies are more appropriate than physical" (Marshall 1991, p. 106). In this step, the two reference systems (i.e., the physical and biological) do not involve the same kinds of underlying theoretical reasoning. According to Marshall, in fact, set formal similarities in physics or mathematics have the advantage of providing static solutions, with an emphasis on some economic aspects. In this case, researchers can establish a formal correspondence between two fields of knowledge (physics and economics), but the contribution of physics to economic analysis is exhausted in providing a series of arguments that do not lead to useful conclusions. In contrast, Marshallian biological analogies establish a substantial correspondence between two fields of study through a coherent network involving similar relationships between objects and properties in both domains. Although the analogy of form is an integral part of the construction of a logical theory, it is essential that the pre-theoretical assumptions also fit the analogy. Frequently, an emerging field adopts the scientific paradigms of another field that is more consolidated or considered more classical. In the case of Marshall, this may seem paradoxical given that biology as a discipline was established in the nineteenth century (by Lamarck), whereas economics is an older discipline. In fact, Marshall uses biological analogies to address questions about economic dynamics that predate the discipline of biology. However, the real question is whether the use of biological analogies surpasses simple imitation to become a useful tool for understanding new knowledge. Our analysis aims to demonstrate how Marshall approaches the study of economics from the Darwinian evolutionary perspective, identifying the economic forces as driving forces and progress in a changing world. In Marshall's era, most of the theoretical constructs developed by the fathers of the equilibrium theory were based on an abstract notion of time borrowed from rational mechanics. For example, the mathematical conditions on which Walras' Law based its concept of balance canceled out at each time horizon.

As highlighted by Claude Ménard, "the action and the consequences are mixed, the dimensions are perfectly continuous; […] we are in a world less time and at no cost". From a technical point of view, "the time the image is copied on the kinematics of the machine without friction […] and the overall time is only moments (of arrests over time) juxtaposed" (Ménard 1979, p. 3). Thus, as time is materialized continuously and uniformly, an economist can afford to abstract this variable in the determination of economic laws.

Incorporating the notion of time is one of the most demanding aspects of constructing any economic model, and traditional economic analysis ignores this factor. In Darwinian evolution, time is inseparable from the origin of the living world and its evolution and is associated with ideas of continuity, instability and contingency. According to Bergson, time evolution is a real-time envisioned as a stream, or in other words, as the mobility of being itself, and is opposed to abstract time, time that intervenes in our speculations on systems artificial, that die and are reborn forever (Bergson 1996). Similarly, Alfred Marshall defined economic progress as organic growth, limited, restricted and sometimes opposed by a host of factors that affect each other and the effect of which varies depending on the state of growth already achieved by each of them (Marshall 1991). His analysis of the balance has different bases and differs from the analyses of other equilibrium theorists (like Walras or Jevons) mainly because he considered the time period. The choice of the time period, in fact, determines the point of view of the observer and therefore determines the theoretical explanation. The two crucial Marshallian periods, the short term and the long term, define profoundly different modes of regulation. In the long term, for example, the determination of specific regulatory mechanisms of the market becomes more difficult because the market's temporal behavior depends on the duration considered. So, whereas according to Walras, the "real" market plays the role of a logical construction, the markets are in equilibrium in time in the Marshallian tradition. Ménard noted that the very foundations of equilibrium in Marshallian marginalist economic thought involve biological analogies: "the problems raised by the integration of decentralized markets and drive led him to seek the active side of the living expression of the most suitable models of economic processes" (Ménard 1979, p. 51).

Marshall shows us how the whole organization is characterized by its transitional form, which compromises any claim that economists could identify universal laws similar to those of physics. As an example, Marshall presents the principle of division of labor. This principle, at the time of Adam Smith, was identified as a “routine” that favored high-quality products through standardization. This principle gradually took the form of mechanization, a process by which man is gradually replaced by machines. Marshall has shown that, regardless of its form, the principle of division of labor must remain suitable for the purpose for which it is applied. Adam Smith had already explained the advantages of this method, but, as Marshall shows, Smith always hoped to develop a universal law that would guarantee the prosperity and welfare of the people who followed it. This form of industrial organization, which met the needs of Smith’s time, owes its success to temporary benefits that outweighed its drawbacks (especially on social matters). By contrast, Marshall’s economic model is inseparable from the social reality; therefore, more must be known about the social reality in which the model operates and its likely impact on daily life.

The problem for Marshall is therefore to find a classification system that allows the economist, using a small number of terms in common use, to express a large number of fine distinctions. The main difficulty is expressing all of this information in a language that is intelligible to the general public while structuring a system of universally valid definitions. The solution proposed by Marshall is of utmost importance to the goal we set for ourselves because it makes direct reference to Charles Darwin and his classification system. We quote a significant step in which Marshall, after having adopted the Millian idea to develop a scientific classification for economic objects, described the precise nature of this enterprise: “But we meet at starting with the difficulty that those propositions which are the most important in one stage of economic development, are not unlikely to be among the least important in another, if indeed they apply at all. In this matter economists have much to learn from the recent experiences of biology: and Darwin’s profound discussion of the question throws a strong light on the difficulties before us. He points out that those parts of the structure which determine the habits of life and the general place of each being in the economy of nature, are as a rule not those which throw most light on its origin, but those which throw least. […]. And in like manner those properties of an economic institution which play the most important part in fitting it for the work which it has to do now, are for that very reason likely to be in a great measure of recent growth” (Marshall 1890; II, I, 2). In the context of economics, this means that the real “affinity” (Darwin’s term) or a particular notion of the fundamental properties determine its adaptation to the medium, although this pattern is limited to those properties that are “the result of hereditary community of descendents” (Darwin 1967). Marshall wanted to prove that the current social organization is the product of slow development over many generations, during which basic properties of the organization were conveyed as a sort of code. Such is the influence of heredity for Marshall, and this influence works both for living things and for business organizations. He attempts to show how most of the distinctions that are expressed in economic terms are based on differences of degree, not of nature.

However, since the publication of Darwin's theory, there has been a misconception that originated in the identification that was made between trends and progress. This is partly explained by the influence of the writings of Herbert Spencer, a philosopher and contemporary of Darwin, who tried to unify under one principle the law of evolution and the phenomena described by the natural and human sciences, thereby distorting Darwin's original ideas. The law of evolution, according to Spencer, expresses a tendency inherent to the increasing complexity of the organization of living things. Darwin never argued this, but he contributed, in a way, to ensuring that this misunderstanding endured by replacing, albeit with some reluctance, the notion of natural selection in the sixth edition of "The Origin of Species" with that of "survival of the fittest", borrowed from Spencer. Some have interpreted this new formula as the quintessential example of the victory of the strong over the weak. However, Darwin's key idea of natural selection has, as its ultimate goal, continual improvement, which inevitably leads to gradual progress in organizations and in most living beings. As Pievani stated: "It is important to remember however that Darwin never tired of his theory resolutely away from any social and political implications: the struggle for survival, for him, was a complex scenario of interrelations between organisms in an ecosystem and had nothing to do with the metaphor of the survival of the fittest that will suffer dire applications in the social and racial field" (Pievani 2005, p. 8). Consequently, what we call progress or adaptation is only the necessary result of inevitable interactions between the system and its surroundings. The idea of evolutionary progress, therefore, does not imply that of an internal principle of improvement. Many have misunderstood the mechanism of natural selection, partly because they have integrated the pattern that underlies it. Natural selection is not deterministic; instead, it must be interpreted as a statistical concept: to have more than a genotype does not guarantee survival and abundant reproduction; this only gives a higher probability (Mayr 1982).

In a chapter devoted to the analysis of social and industrial organizations, Marshall seamlessly integrated this dissociation between progress and development that exists in the Darwinian theory of evolution. He writes: "in the same way, that greater factor of economic prosperity, the organization of a well-ordered state, is the product of an infinite variety of motives; many of which have no direct connection with the pursuit of national wealth" (Marshall 1890; IV, VIII, 5). This definition includes any economic models that include two characteristics from the main philosophical concept of Darwinian evolution: first, the integration of a historic time for the study of economic forces such that scholars must consider irreversible phenomena, and second, economic evolution does not presuppose the idea of a linear change oriented toward perfectibility. Evolution and progress are two separable notions. This dissociation explains how the different forms of social organizations and industries undergo a process of worldwide economic natural selection that gradually eliminates those agents that do not properly adjust to changes while simultaneously promoting the development of new organizational forms. In addition, the economic mechanism of natural selection, similar to the analogous mechanism in biology, is understood probabilistically.

Outside influences may prevent the emergence of an organization that appeared perfectly adequate a priori because "Conversely, the struggle for survival may fail to bring into existence organisms that would be highly beneficial: and in the economic world the demand for any industrial arrangement is not certain to call forth a supply, unless it is something more than a mere desire for the arrangement, or a need for it" (Marshall 1890; IV, VIII, 1). Finally, Marshall retains the uncertainty surrounding the mechanism of natural selection. Progress in the division of labor, as a special form of organization, is mainly due to external factors, such as "It is the largeness of markets, the increased demand for great numbers of things of the same kind" (Marshall 1890; IV, IX, 3). He then proposes an analogue to the Darwinian principle of divergence, derived from Milne-Edwards' physiological division of labor, which provided a preferred direction change. According to this principle, the conquest of new ecological niches created new opportunities for innovative organizations to thrive. When an organism creates the ecological conditions in which it can thrive, it not only diverges from the group from which it was derived, but also opens up new possibilities for its existence and reproduction. This principle, when applied to economics, explains the gradual disappearance of similarly structured firms in the same market.

Marshall, in considering the various forms of organizations, is placed in a context in which real variability is the norm. The environment does not begin with an arbitrary definition of industrial enterprises and reject all deviant forms. In contrast, it defines a company through its life cycle, "But here we may read a lesson from the young trees of the forest as they struggle upwards through the benumbing shade of their older rivals. Many succumb on the way, and a few only survive; those few become stronger with every year, they get a larger share of light and air with every increase of their height, and at last in their turn they tower above their neighbours, and seem as though they would grow on forever, and forever become stronger as they grow. But they do not. One tree will last longer in full vigour and attain a greater size than another; but sooner or later age tells on them all. Though the taller ones have a better access to light and air than their rivals, they gradually lose vitality; and one after another they give place to others, which, though of less material strength, have on their side the vigour of youth. And as with the growth of trees, so was it with the growth of businesses as a general rule before the great recent development of vast joint-stock companies, which often stagnate, but do not readily die. Now that rule is far from universal, but it still holds in many industries and trades. Nature still presses on the private business by limiting the length of the life of its original founders, and by limiting even more narrowly that part of their lives in which their faculties retain full vigour" (Marshall 1890; IV, XIII, 1). These laws of nature act on a company by limiting how long it will operate before it loses part of its strength, flexibility and innovativeness in competition against its leading competitor.

The idea of a struggle for survival is an old idea, the use of which dates from the seventeenth and eighteenth centuries. At that time, it was considered a benevolent formula, which allowed the necessary corrections to the balance of nature (e.g., Line, Cuvier). The Darwinian theory of the struggle for existence calls into question the idea of a harmonious consistency of the world. Just as Darwin views the adaptation of organisms to be a dynamic process and a static status over time, organizations are doomed to extinction unless they continually change and adapt. Returning to the economic context, the Darwinian principle of the struggle for existence expresses a similar idea about the competition for resources between several agents. Note that this implementation is only possible because Marshall views the company in terms of its organizational structure. Marshall applied this idea to the competition between organizations within a specific industry. In particular, as Keynes, who was a student of Marshall, once said: “The volume as a whole also serves to illustrate what Marshall was always concerned to emphasise, namely, the transitory and changing character of the forms of business organisation and of the shapes in which economic activities embody themselves. He calls particular attention to the precarious and impermanent nature of the foundations on which England’s industrial leadership had been built up” (Keynes 1924, p. 213). Specifically, Marshall considers large companies that have evolved methods of mechanization to be most advantaged because they can take advantage of both external economies related to the overall development of the industry and domestic economies linked to their resources.

Some authors, while recognizing the innovativeness of the biological analogies mobilized by Marshall, have criticized the fact that they have a consistent explanatory form and have not been systematically integrated into the body of the analysis. For example, Geoffrey Hodgson (1993) argues that the expressions are nothing more than biological metaphors of style and that, therefore, Marshall’s biological model of the economy is more a promise than a fact. Similarly, Clark and Juma (1988) show that despite the use of biological metaphors, Marshall’s work is primarily non-evolutionist: “[A]lthough he advocates the use of biological concepts, his own work paid only token allegiance to the approach. Much of The Principles of Economics is non-evolutionary except for the sections which deal with industrial organization and the division of labour where he draws on the concepts of survival of the fittest and psychological view of human behaviour. He sees large-scale industries as trees of the forest which grow, compete for light and water, lose vitality, grow old and die” (Clark and Juma 1988, pp. 203–204). Nelson (1995), on the other hand, while crediting Marshall with developing a new kind of economic analysis, notes more mechanical than biological analogies. Although Marshall was attracted to “biological conceptions” of economic ideas, it is apparent that he found himself forced to fall back on “mechanical analogies”. Therefore, he must have found it very difficult to develop a formal theory based on “biological conceptions” that he thought adequate for economic analysis (Nelson 1995, p. 49).

However, we cannot dismiss the ideas of irreversible change and organic growth present in the work of Marshall. Through the analysis of long-run supply curves with increasing returns, Marshall offers a clear example of irreversible phenomena. This “Marshallian vision” marks a clear departure from tradition, in which such abstract economic terms as law, normal trend, and average power do not express what really happens, but what might happen based on certain assumptions that are never exactly realized. The real novelty of Marshall was that he passed this limit by drawing parallels between biological and economic laws. He gives us a monistic view of the world in which economic and biological phenomena are governed by comparable laws.

3.3 The Computational Methods and the Engineering Approach

According to Foucault (1998), each "age" possesses special conditions that allow us to reflect on the kind of knowledge of a particular historical moment. Over time, epistemology has identified a common basis for various forms of scientific knowledge across disciplines. As economics and biology are traditionally placed on opposite sides of the frontier between the natural and social sciences, it is possible to think of their relationship as simply an exercise in reductionism.

Reductionism implies a hierarchy between the disciplines. Thus, the reduced discipline sees its traditional methods scientifically discredited, and the substance of his research "best explained" by the scientific models and methods of the reducing discipline. What framework would then be reduced further?

It is known that the logical empiricists built their entire approach on the idea that scientific theories are the very heart of science. A theory must be understood as a formalized system based on axioms from which one can deduce the consequences of an action, which in turn, can be verified experimentally. These philosophers have almost always chosen examples from physics, in which most of them were educated. Therefore, the problem of reductionism is viewed in this framework in terms of relations between theories. The classic work of Nagel (1961), Structure of Science, offers perhaps the best illustration of this concept and also largely dictated the content of subsequent discussions. According to Nagel, the reduction of one theory into another requires that two criteria are met. First, the laws of the reduced theory must be able to be derived from the reducing theories (the condition of differentiability). Second, the vocabulary of the two theories must be able to be related by translation (the condition of connectability). These criteria are known as the principles of the bridge, or bridge laws. This second condition would appear to be easily met because these philosophers defined every theoretical term at the level of observation. This consistency guarantees that two theories can be translated and compared. However, it became clear quite quickly that these laws raised a great number of difficulties. Philosophers such as Thomas Kuhn and Paul Feyerabend have shown in their works the different theories that could not be compared to each other with all the beauty and simplicity of logic that characterized the logical empiricists. Countless debates gave rise to the idea of incommensurability between the theories proposed by Kuhn, but it became clear that development in relation to theories and their vocabulary was far from being a matter of course. Because these issues were initially discussed using examples in physics (as was long the case in philosophy of science), the application of this reductionist biological framework created additional difficulties.

The case attracted the attention of philosophers of biology interested in the possible reduction of classical genetics to molecular biology. When philosophers of science began to debate this topic, toward the end of the 1960s, molecular biology had grown into a thriving, successful field that seemed capable of explaining biological processes at the molecular level. The field had discovered that genes, which had been theorized by Mendel and geneticists in the first half of the twentieth century who had studied in detail the modes of genetic transmission, were in fact made up of a sequence of nucleotides in a double helix of DNA. It therefore seemed reasonable to expect that molecular biology could reduce Mendelian theory, including its laws on the segregation of characters and the concepts of the gene, the allele, recombination or even dominance. The most successful attempts to apply the reductionist framework in this example can be found in the work of Schaffner (1967, 1969) and Ruse (1971), but here I summarize the questions that have arisen since the 1970s. The fundamental problem that theoretical reductionism encounters when applied to biology is that each entity of classical genetics can be achieved in various ways at the molecular level. Expressing a functional entity in molecular terms would require an infinite separation of structures and molecular processes, which would certainly be very heterogeneous. This is a one-too-many relationship. This problem is also called multiple reliability. To better understand this idea, consider the case of the gene (Hull 1974). The more knowledge in molecular biology progressed, the more it became obvious that it was not possible to determine, in molecular terms, the definition of a gene in general. In fact, this problem exists at two levels. First, if one admits that genes correspond to parts of DNA, one must recognize that the category of a functional gene corresponds to an infinite number of DNA sequences. But the problem is much deeper if you consider that the molecular definition of a gene should include not only the sequence itself, but also the molecular mechanisms that enable its effects. The inclusion of context, then, presents two difficulties. On the one hand, we risk worsening the problem by integrating additional mechanisms into our multiple reliability analysis. On the other hand, we refer to the molecular aspects of the context, undermining the foundations of reductionism. The case of dominance is equally illuminating because it seems to exist in a number of ways at the molecular level. We can therefore say that the central problem of this first form of reductionism is related to the complexity of the relations linking the two fields of biology of which we speak.

The feasibility of the idea of multi-level generalizations that have a certain independence from the lower-level mechanisms that implement them is quite important. Thus, the principles outlined by Mendel possess an explanatory power that does not depend on the details that molecular biologists can describe. Waters (1990) speaks in terms of details that are extremely complex, but would not be relevant to high-level phenomena. According to Kitcher (1984), the law of segregation is completely explained by reference to the fact that there are two copies of each chromosome and that a copy of each is found in each gamete. It is in this sense that the functional explanations in biology are irreducible. The positivist framework in which this debate arose, with its insistence on laws, has now been largely abandoned. As many authors have shown, there are no laws in biology (with the possible exception of the principle of natural selection), and biologists seem more interested in studying particular phenomena and mechanisms that describe universal principles. After all, it is reasonable to conclude that if there are no laws or general theories in biology, the issue of reductionism must be addressed in a different way than in terms of relations between theories. Because we are concerned with individual cases and universal generalizations, the problem of multi-feasibility ceases to be a concern. In fact, it is no longer a reduction of type to type, which corresponds to the reduction of one theoretical term to another, but a reduction of occurrences (token–token), that is, of one particular case to another.

For several years, the problem is more room in terms of theoretical reductionism but in that of explanatory reductionism. Explanatory reductionism can be qualified in so far as it can be reduced, but the theories cease to provide explanations. William Wimsatt suggested a very clear definition of explanatory reductionism: "A reductionist explanation of a behavior or property of a system is an explanation that shows that they are explicable in terms of mechanical properties of parts of the system and their interactions" (Wimsatt 2000, p. 293). According to this definition, in fact, trying to reduce a complex phenomenon to its simpler constituents, or to explain reality through the macroscopic relationships among smaller elements, is a good heuristic.

This definition clearly shows that explanatory reductionism is fundamentally linked to the question of mechanism. Philosophers have begun to pay attention to particular explanatory models, as opposed to general laws and theories, and this change has also manifested in a decade of a mechanistic framework of analysis in the philosophy of biology. A series of articles dating from the late 1990s and early 2000s mainly raised these discussions. The issue of what a mechanism and a mechanistic explanation exactly are in biology has given rise to various discussions. Different responses have been proposed, but scholars agree on the definition of a mechanism. A mechanism is made of parts that interact causally to produce a certain phenomenon. A mechanistic explanation of a phenomenon therefore provides a description of the mechanism responsible for the production of this phenomenon. In Glennan (2002), a mechanism is defined as interactions between parts that induce a change in another part. These relationships between parts are defined primarily in terms of laws and invariant relations. In an article that has come to represent this debate, Machamer et al. (2000) state that a concept expressed only in terms of ownership of shares and their changes is insufficient. For this reason, I introduced the concept of activity. One of its purposes is to avoid resorting to the concept of law because these mechanistic explanations would be meaningless. According to their definition, a mechanism consists of entities and activities that produce regular changes. The entities and their activities must be capable of explaining how one moves from an initial to a final state through a succession of stages with no break in continuity. The purpose is to explain how the changes are made. The notion of activity is better than that of interactions for explaining this presumed ability to make changes. We will not compare these two concepts in detail, but it should be noted that, as Tabery (2004) showed, they are not only compatible but complementary. On the one hand, the notion of activity can give a better account of the ways in which interactions produce changes in ownership. On the other hand, the idea of productivity becomes less abstract when it is approached in terms of change of ownership. It should be noted that, in the mechanistic framework, the relationships between levels are not designed in a strictly reductive manner in the sense that these are complex relationships that must be thought of in two ways. As noted, in a mechanistic explanation, there is a double constraint: on the one hand, lower-level entities create the phenomena in question; on the other hand, the system in which the mechanism is built must be fully examined. The explanation, therefore, cannot be reduced to a single level (Darden 2005). It is in this sense that we often talk about inter-level models and integration between sectors (interfield integration). However, a mechanistic explanation may move in a reductionist direction to the extent that a phenomenon is produced by interactions between entities at lower levels.

At present, references to economics as “applied biology” (or vice versa) must assume a set of “pure” concepts in biology that find applications in empirical economics. This is the case of a characteristic that has marked the recent development of both sciences, namely the use of computational methods. The idea of computation was first used in biology, assuming biological mechanisms of information processing. This idea was introduced in the mid-twentieth century by two Nobel laureates, Max Delbruck and Linus Pauling. They argued that the main element of the interaction of biological reactions was complementary molecular structures (like a lock and key). This intuition proved correct, and today it constitutes the perspective on which on the entire field of molecular biology is based: structure and function. The shape of a molecule determines with which other bioactive molecules it can interact, i.e., those that have a complementary shape. As written by the philosopher of science Telmo Pievani: “While the explanation is based on the physical consequences of the timeless laws, the biological explanation has to do with <Functions>: we do not say that the Moon has the function to raise the tides or the explosion of supernovae is <per> produce heavy materials which make up life. Instead, we say that the eyes see and the ears are <per> <per> heard that one and the others are products of natural selection and are <built> , in one way or another during the development of each individual, from the information carried by genes responsible” (Pievani 2005, p. X).

The idea of computing today adds a further element of discussion. It can be summarized in the motto “Structure = Function = Computation”. The appealing idea, placed in the center of numerous philosophical debates, is that the mind is not identical to the brain or the interaction of its parts, but the mind is equivalent to information processing (computation). The latter is the intuition that Turing applied to our thoughts, viewed as the equivalent of computational steps that “run” on a particular device, in this case, our brains. All the “stuff” of thought is nothing but configurations of the information gathered, processed and transformed by physical mechanisms of the brain. In this way, abstract and intangible things like thoughts are the basic operations of our physical brain. Even Turing’s intuition found fertile ground in biology, so that we can speak of its application to living cells. For example, consider DNA. The deoxyribonucleic acid (DNA) has physical properties that are quite complex. This is a double-stranded polymer formed from bases, equipped with a series of structural and chemical properties that contain nitrogen. But the really important thing is that DNA stores and processes information, which means that, above all its thousands of physical properties, DNA is a nanoscale computational device. And what is even more interesting is that DNA carries information for the construction of small machines (proteins) that can transfer more information, also encoded in its sequence (Montague 2006). Genes are transcribed from DNA, genes form proteins, proteins form tissues and organs, such as the heart, lungs, bones and brain, which operate to control bodily functions and psychological experiences. The structures (patterns) are, in turn, new structures in which each small part is engaged in computation. The whole organism can therefore be broken down into discrete parts, each of which will find an optimally adaptive function. The latter, namely the creation of good function, is dependent on a specific function, almost like a car that has a good "construction of the gear". In this economic rationale, algorithmic or computational models are particularly relevant. In fact, an algorithm is a mechanism (because it can be broken down into routines and subroutines) that is subjected to constraints of tractability. To immediately dispel any contradictory interpretation, it is important to specify that you are not speaking of what has been described in the classic functionalism (or computational) literature. This literature, in fact, reduced cognition to only one level of organization (the software), which was made of a material (the brain), justifying this reduction through a simplistic analogy with the computer. In the new view, however, both computers and the brain are seen as hierarchical structures that can be decomposed into many levels of information processing: the system of exploitation, the essential system (shell), assembly, machine language, and application. Rather than simply a structural and a functional level, there is a hierarchy in which each function makes up a higher-level structure, in the same way that the brain is described as a hierarchy of mechanisms: molecules, neurons, networks, and circuits.

An interpretation of the body plan of this type therefore requires a particular kind of explanation: namely, an explanation by mechanistic models. A "mechanism", as we have seen, is in fact a collection of entities and operations (or activities) organized to produce regular changes in a sequence in time, which can be decomposed into the initial conditions (the parameters that allow the conduct of the mechanism), intermediate (causal chains: cycles, joints, and networks) and terminals (production, disposal and balance). Let us investigate an example. X1 is an entity that causes Φ1, and "because Φ1 X1" is a phenomenon observed in an attempt to describe X1's constitutive, contextual and etiological properties. In the constitutive explanation, an "explanandum" type "X1 Φ1 because", a phenomenon that one tries to understand the nature of X1 (such as X1 due Φ1) is coupled to an explanans type "P1, P2, P3… Pn performance σ1, σ2, σ3… σn that allows X1 to cause Φ1". This explains, for example, that the pancreas (X1) controls blood glucose (Φ1) because the alpha cells of the islets of Langerhans (P1) produce insulin (σ1) and the beta cells (P2) produce glucose (σ2), and so on, according to a certain causal sequence that flows from some initial conditions to the ultimate conditions. A more detailed description could include the mechanisms by which cells produce the hormone.

This type of analysis, however convincing it may seem, has encountered many challenges since its appearance (Lewontin 1974). It is difficult to decide how one should (or would) split a body into its "parts", or discrete components (if they exist), that also play functional roles in specific adaptations, including single parts and single functions (often because the condition of an adaptive character, or an organ, is represented by a plurality of functions or "interweaving of causes"). In fact, there may be a multitude of levels and a multitude of sub-elements at every level. It should be noted, however, that the explanation is not a constitutive redescription of the same system at a lower level, but a decomposition of a system into subsystems. "X1 is Φ1" may also include an explanation in context, knowing the interaction of X1, X2, X3… Xn that performs Φ1, Φ2, Φ3… Φn so that a system that includes them, S1, performs an Ψ1. Therefore, for example, the pancreas, which controls blood sugar levels in a mechanical explanation of the digestive system, interacts with the stomach, liver, duodenum, and other organs. In this case, the interactions of the pancreas with other entities are of interest, but the constituents of the pancreas are not. Constitutive and contextual explanations explain how an entity or activity is part of a causal network: the first shows how Φ1 X1 is due to the activities of the parts P1, P2, P3… Pn from a higher level to a lower level. The second kind of explanation, in contrast, interacts with other entities, such as X1 X2, X3… Xn, and the explanation is distributed both horizontally (interaction between entities) and vertically (moving from a lower to a higher level). A third way to explain how the entities and activities are part of the order is by virtue of their causal history. An etiological explanation identifies a causal sequence that leads to the current performance of Φ1 by X1. Adaptationist explanations are typical examples of etiological explanations, but they are not the only explanations. Development, individual history, learning and social phenomena can be represented in etiological explanations. They do not describe Φ1 X1, X1 or how X1 interacts with other entities to produce a system that includes an activity at another level, demonstrating where constitutive explanations end and contextual explanations begin (although the former can inform the latter). Thus, an etiological explanation of the regulation of blood glucose (Φ1) from the pancreas (X1) can involve both the development of the genes that control cell division and morphology and evolution (the regulation of glucose by the pancreas is a feature that is found in organisms today because their ancestors had the same feature). We note that an etiological explanation is only one kind of backward-looking explanation involving adaptationist thinking, which argues that the current presence of a characteristic is due to its earlier adaptivity. An etiological explanation is costly from an epistemological perspective as it can only illustrate the trajectory of a mechanism.

A mechanistic explanation is fundamentally different from a deductive-nomological explanation, but the former provides a scientific explanation in the same way without the intervention of the law. Rather than an idealization, as with the Newtonian laws, a mechanistic explanation postulates a model rather than abstractions. An idealization is indeed a perfect representation of an entity that has no empirical counterpart, whereas an abstraction is a simplification of an existing entity (Cartwright 1989). The first case, therefore, describes the formal relationship between fictions, whereas the second case describes the relationships between entities as real items (Schick 1991). The contemporary philosophy of science recognizes that, in many cases, to explain a phenomenon is to establish “a causal connection” rather than a nomological one. In biology, economics and psychology, the causal explanation proceeds with the formulation of mechanistic models rather than a deduction-nomological model. A scientific model will prove to be so, and a simplified representation of a hypothetical system is too complex to be studied directly. Some theoretical models predict a role of these mediators (Morgan and Morrison 1999) between the theory and the data. The theories provide principles or practices to formulate a simplified model of the phenomena. In biology, natural selection, for example, is not a law of nature, but a general principle from which one can construct models of the dynamics or population genetics that will serve to make some prediction or explanation.

The goal of theory development is thus not to produce axioms or psychological models, but to put forth another kind of knowledge that, despite sharing some properties with descriptive theories, cannot be reduced to one. It aims, ultimately, to produce technical knowledge, i.e., knowledge that is not confined to the natural sciences or mathematical exercises (although the latter often use their own theories). This knowledge (which applies in various fields of human activity, from biology to aerospace engineering, genetic engineering, robotics and macro- and micro-economics) is, in fact, technical knowledge, i.e., knowledge of “engineering”.

Engineering is often defined as an applied science that requires a set of technical skills that are difficult to assess as “knowledge”. Instead, it is customary to distinguish engineering from other sciences with reference to the categories of Gilbert Ryle (1949): “know how” (knowing-how) and “know” (knowing-that), which are sometimes reformulated as procedural and declarative knowledge, respectively. Even to assimilate the engineering “know-how”, science uses “propositional knowledge”. Although practical, these distinctions better reflect uses of the word “knowledge” than they do actual differences between engineering and science. Developmental psychologists, for example, use this distinction to differentiate memory systems according to their degree of conscious accessibility or the type of representation involved. The difference lies not in the form of knowledge, but in its use. Although it uses scientific knowledge for a practical purpose and makes use of technical skills, some epistemologists consider engineering expertise to constitute domain knowledge in the proper sense of the term, that is, distinct from knowledge of the natural or social sciences. Engineering knowledge is, in fact, the designing and building of artifacts that are destined to change the physical or social environment to meet certain needs or desires (Vincenti 1990). This form of knowledge, then, is distinct from scientific knowledge, which is pragmatic and goal-oriented. So, whereas the scientist describes “what is the case”, the engineer describes “what might be the case”, or what, based on some assumptions and according to some objectives, “should be the case” (Auyang 2004).

In fact, when we deal with building bridges, railroads, skyscrapers, rockets or automated systems, conditions can affect the outcome of the project. Good projects achieve their goal: a “normatively correct” bridge does not correspond to an idealized model of the bridge, but it ensures the safe transport of vehicles by meeting the criteria of simplicity of construction, durability, cost, and other parameters. In this case, therefore, engineering projects are considered for their practical value rather than being evaluated in terms of their suitability to theoretical knowledge. Therefore, knowledge of “engineering”, and technological knowledge more generally, is essentially prescriptive, that is, composed by procedures or rules that describe actions that should be useful for achieving practical purposes (Kroes 1998). The criteria for evaluating technical knowledge, such as effectiveness, cost, safety and utility, are more comprehensive than the criteria for evaluating scientific knowledge. The latter is governed by criteria such as truth, empirical adequacy and explanatory power. However, even if their projects are different, science and technology are in a balanced relationship (Goodman 1984): science uses technology during research (e.g., measurement and analysis), and technology uses science for project development (design, construction, and modeling). The links between science and engineering were highlighted primarily in an epistemological study by Quine. Against the normativist objection to naturalism, Quine’s naturalized epistemology stated that it was not inconsistent with rules and regulations: “(The) normative epistemology is a branch of engineering. It is a technology of the search for truth or, epistemologically more cautious in terms of prediction. Like any technology, using freely any scientific discovery that it suits her purpose. […] There is no question here of ultimate value, (but) its efficacy for an ulterior end, truth or prediction. What is normative, as in engineering, becomes descriptive when the terminal parameter is expressed” (Quine 1986, pp. 664–665).

Quine’s view can also be applied to economic rationality through a proposed engineering concept of rationality. This proposal can be broken down into three different theses:

(1) Economic rationality is a branch of engineering.

Under this thesis, the rationality of action is a type of control because it explains how a system should regulate itself using the information available so that, by satisfying certain criteria, it can reach a certain goal. More precisely, the definition of rationality belongs to “intelligent” control research, a research area that bridges industrial engineering, operations research and artificial intelligence (Antsaklis 1994; Saridis 1985). To define a rational agent, a search will be conducted within the concepts offered in control theory (in engineering), the theory of agents (artificial intelligence), and the optimization theory (in operations research). The rules for these control systems are actually conditions: they must reach their goal, keep the system in a certain state, minimize the energy expended, and perform calculations in a reasonable time. The engineering approach thus offers the advantage of working with endogenous constraints to formulate rational decision problems: rather than representing a rational agent perfectly, a “frictionless” rational agent must take into account the various limitations and constraints inherent to agents.

(2) Economic rationality is a technology for research utility.

Ensuring the control of a system is a way to solve the problem of optimization in which one needs to determine the best solution to a given objective function that must be maximized, minimized or stabilized or to estimate the solution as closely as possible. Thus, to construct a theory of rationality is to create a model of a system optimizer that maximizes its usefulness, taking into account some limitations. In particular, the treatment prevents the control algorithm from formulating models that are incalculable or intractable.

(3) Like any other technology, the economic rationale freely uses any scientific discovery that suits its purpose.

In engineering and in artificial intelligence, several control structures make use of knowledge from other fields: biology (adaptation, evolution, and genetic algorithms), neuroscience (neural networks), psychology (learning, models, planning, and memory), social sciences (multi-agent control), decision theory and philosophy (logic and knowledge). Therefore, behavioral ecology, evolutionary and experimental economics, neuroscience, the psychology of decision making, population genetics, game theory and robotics can be valuable resources for the construction of models of rationality. Theoretically, adopting these views in the study of economics could broaden the field beyond the ideas generally accepted by economists and philosophers, according to which the concept of rationality has "a face on both sides" (Bermudez 2000), which shares properties with many skills. In this way, decision theory, game theory or market behavior maintains the same relationships both with theories and regulations and with the descriptive sciences, chemical engineering and theoretical and analytical (laboratory) chemistry. Thus, beyond the obvious differences between the systems studied, the computer simulations collect entities according to criteria of hierarchical organization or complexity, thus providing scientific knowledge of other classes of entities in which the properties of one system can be used to make predictions about the properties of another. In this new way of grouping areas of research, microeconomics, ecology, robotics, artificial intelligence (AI) and artificial life (AL) share the same methods of calculation, i.e., simulation-based agents.

3.4 Complexity

When explaining a phenomenon, event or observation, it is typical to favor the theory that most adequately describes its occurrence and, simultaneously, to remove the rival hypotheses by means of sophisticated reasoning showing how they are false, insufficient, less likely or less satisfactory. This approach, which is typical of philosophy, seems to be very different from what happens in science, where assumptions, theories and ideas are compared with reality as directly observed with the senses or perhaps observed using a particular instrument. By virtue of a constant dialogue between theory and empirical observations, this approach allows scientists to obtain a large number of detailed, objective and accurate predictions that may be replicable by others. These differences in the approaches used by these fields created a sort of boundary between social sciences and natural sciences, with each of the disciplines involved in studying the phenomena that best suited their respective instruments, thereby neglecting the possible relationships between fields of study on either side of the divide. However, in recent years, many cognitive scientists (but also physicists, philosophers of science, epistemologists and economists) have tried to bridge this divide, convinced that the connections between phenomena in nature can be explained by appealing to many different disciplines. Based on these considerations, many researchers considered it necessary to find new ways of unifying research through what the physicist Snow (1993) called the "third culture" in the hopes that the new knowledge resulting from such an approach would be able to reconcile the humanistic and technical-scientific perspectives.

In this way, we tried to overcome the most negative consequence of disciplinarity in science, namely the idea that one of the most effective, precise and concise ways to describe the reality of a phenomenon is to break it up into small parts, describe the dynamics of the individual parts, and then partially combine the descriptions to understand the phenomenon as a whole. This method of investigation (which is known as reductionism), provides an idealized view of a phenomenon or class of phenomena capable of giving rise to an effective mathematical description that expresses the alleged regularities of the phenomena through strict formal rules. The reductionist approach, therefore, is characterized by the belief that to understand a certain part of the world, it is useful to examine in detail its constituent elements and the laws that govern them, with the conviction that what is observed at the macroscopic level depends on relationships at the microscopic level.

Since the 1970s, there has been increasing interest in new research areas that do not lend themselves to reductionist approaches. Their objects of study are systems that are often made up of easily identifiable sections. These parts, or components, cannot be isolated from the system in which they operate without the system breaking down completely. This fact justifies the existence of the science of complexity, the primary mission of which is "to overcome the simplifications and idealizations that lead to unrealistic points of view" (Chu et al. 2003, p. 19). The science of complexity arose from interactions of different fields of science, including physics, mathematics, biology, economics, industrial engineering, and computer science, and now encompasses even more important fields of scientific research. The use of the term "complexity" as a label to indicate an object of study is relatively recent. Before the mid-twentieth century, it appeared almost exclusively as the opposite of "simplicity" and with a slightly negative connotation, referring to the properties leading a researcher, supported by experimental data, to choose a particular theory over competing ones. This kind of preference for simplicity can be traced back to Occam's razor, which states that we should not multiply entities beyond the strictly necessary (*Entia non sunt multiplicanda praeter necessitatem*), i.e., one should not multiply entities if such a natural multiplication is not logically necessary. Besides simplicity, another key feature of science until the late nineteenth century, which was linked to classical mechanics and the reductionist and mechanistic view of the world, was a deep-rooted concept of linear causality. Classical determinism assumed, in fact, that identical effects had identical causes, and also that the basic principles of simple equations or simple motion in themselves led to equally simple forms of dynamic behavior. These assumptions, in fact, provide an effective model for only a relatively small number of isolated phenomena in linear systems. In this approach, it is assumed that the causes and the effects are linked by linear laws that, at the microscopic level, can be reversed in time and are therefore, in principle, sufficient to allow us to trace back the initial conditions of a phenomenon that has developed over time. Linearity was then the central paradigm of classical mechanics, mathematically formalized with algebraic and analytical methods. However, physicists (who of course knew of the existence of nonlinear systems, but had observed exceptions to their rules for a number of years) began to find it increasingly difficult to bend many natural phenomena to a description adhering to rigid determinism, with a linear relationship between cause and effect, using the reductionist methods of the classical paradigm. The classical paradigm in particular proved inadequate for describing processes related to biological life, in particular the phenomena due to interactions between individuals in human societies, such as those relating to communication, decision-making and the intricate set of relationships between man and ecological systems (Bertuglia and Vaio 2011). One example among many comes from protein folding. Proteins fold in a particular way, like twisting a shoelace. Although, from a reductionist point of view, researchers understand all of the elements and forces in play, not only does each protein fold in a different way, but its final form is never the same and is highly unpredictable (Licata 2011). The mass of accumulated data on the elements of biological systems has indirectly opened a Pandora's box in that it is no longer possible to believe that reductionism will enable scientists to understand biological systems. The organization of the constituents is not contained in the description of these same constituents, as in the protein folding example. However, the properties of the system depend entirely on the constituents' organization rather than on their characteristics.

As of the 1920s, first in the context of biological phenomena and, subsequently, in that of social phenomena, the notion of system began to be imposed: a system is seen as an entity that exceeds the sum of its parts, in which the most important factor is not the laws that govern the dynamics of single parts, but the complex interactions that bind each part to the others. The science of complexity suggests that systems are characterized by nonlinear dynamics and can transform a simple behavior into a complex behavior and vice versa (Limburg et al. 2002). These systems are characterized by linear systems in the sense that to understand the state of a system at a specific time, a nonlinear function of the system state at an earlier time must be used (McDaniel and Driebe 2001). A complex system is structured by feedback circuits in interaction (Forrester 1975), and more precisely, by strengthening circuits that amplify and reinforce the phenomena in the system and balancing circuits that resist and oppose the change, maintaining the equilibrium of the system (Sterman 2000). As a result, a complex system is dynamic because of its internal structure and fundamentally causal due to the presence of feedback throughout the system (Meadows and Robinson 1985). In addition, these systems have the unlimited capacity to adapt based on their past and current behavior. The context in which the system evolves, then, has a strong and important influence on behavior (Chiva-Gomez 2004). Ultimately, these complex systems are characterized by the multiplicity of their components (natural, technical, economic and social) and their interactions, but also by the diversity of their dynamic behavior. The behavior of complex systems supports the notions of feedback and homeostasis, and analysis of adaptation is generally based on a selective information theory in which information returned from the environment, or linked to past experiences, is incorporated into future behavior (Simon 1962).

Although the concept of system has been defined in multiple ways (e.g., Simon 1962; von Bertalanffy 1968; Morin 1990), it usually includes the following characteristics: (1) the system involves a large number of elements; (2) these elements interact dynamically; (3) the interactions are numerous in the sense that every element of the system may influence or be influenced by another element; (4) the interactions are nonlinear, (5) interactions generally occur in the short term; (6) the interactions include positive and negative feedback loops; (7) the system is open; (8) the system works under certain conditions that depart from its equilibrium; (9) it has a history; and (10) its individual elements generally ignore the behavior of the global system in which they operate. Generally, the complexity of a system can thus be described in three dimensions: (1) the number of elements that constitute it; (2) connectivity, i.e., the number of links between these elements; and (3) functionality, i.e., the functional interconnections between the elements (Sterman 2000). In addition, uncertainty is usually present in these systems, making it difficult to anticipate future behavior (Glouberman and Zimmerman 2002; McDaniel and Driebe 2001; Morçöl 2005). The most important theories in the field of complexity that have advanced from these common principles are: (1) general systems theory and cybernetics, (2) the theory of catastrophes and chaos theory, and (3) complex adaptive systems theory. Whereas the first two theories concern deterministic dynamic systems (i.e., systems in which the state at time t determines the state at time t + 1), we concentrate on the third theory from this point forward because it is more useful to our purposes of focusing on the study of regularities that emerge from interactions among interconnected individuals in complex adaptive systems (Anderson 1999).

The systems of interest are complex adaptive systems composed of heterogeneous agents interacting with each other and with their environment (Chiva-Gomez 2004). "An agent" is a generic term referring to "semi-autonomous entities that contains a complex system, such as atoms, molecules, […] methods, individuals, groups, businesses, etc…" (Maguire et al. 2006, p. 204). The theory of complex adaptive systems presents the feedback loops in systems, such as emerging connections between the agents involved in the system and the connections between different variables (Anderson 1999), as do the other two theories presented above. McDaniel and Driebe (2001) describe five common characteristics of complex adaptive systems:

(1) The agents. The systems include a large number of agents, who develop and exchange information and react to changes in information. They are constantly in action and interacting with each other. However, all agents are different; they do not have access to the same information, and no one agent can have a complete view of the system as a whole.
(2) The interconnections. The essence of a suitable system is captured in the relationships between agents rather than by the agents themselves. The connections that the agent has with other agents in the system and with other environmental agents are agents of the environment. These relationships are not linear in nature.
(3) Self-organization. This concept emphasizes the spontaneous emergence of new structures and new forms of behavior in open systems, characterized by internal feedback loops. Self-organization is emerging as changing behaviors related to correlations in the system. In other words, any change is connected to a reaction and interactions between agents.
(4) Emergency. Emergency is the result of nonlinear dynamics that generate new properties. In fact, the agents interact, self-organize and generate emergent properties of the system. However, the behavior of a global system cannot be obtained by summing the behaviors of its constituent parts; the agents cannot predict the behavior of the system as a whole and can barely control these emergent properties.
(5) Co-evolution. There is co-evolution between the right system and its environment such that each influences the development of the other.

The science of complexity, therefore, offers a new vision of the decision-making process. The focus of this field is mainly on the unpredictable and ambiguous world, the importance of non-linear relationships and the roles of self-organization, emergence and co-evolution in organizational dynamics, without any reference to reductionist ideas of prediction and linearity. The traditional command, control, prediction and planning activities give way to the attribution of meaning, learning, improvisation and reflection (McDaniel and Driebe 2001). In fact, in complex systems, the issue is no longer the bounded rationality of decision makers, but rather the lack of knowledge of the dynamics inherent to the systems. In this context, individuals must develop a collective vision of the situation, its "raison d'être" and its consequences. Defining a situation is thus a social act that requires interaction between agents. The ability to process information, follow rules and connect to other agents is necessary for an organization to understand complex situations. Also, agents constantly create and recreate sense and meaning because the behavior and order in the system are continuously changing (McDaniel and Driebe 2001). Therefore, the main functionality of an action in a complex situation is the creation of connections and correlations, and the quality and type of connections between the actors are more important than the actors' individual quality. Participation in decision-making is therefore a strategy of action that can be used to improve the system (Ashmos et al. 1998). In fact, participation is "a mechanism designed to increase social exchange information" (Ashmos et al. 1998, p. 27). This concept implies the importance of the agents during the decision process, not only in terms of the number of individuals involved or the different types of people represented, but also in terms of the depth and frequency of their involvement and the mechanisms by which they become involved (Ashmos et al. 1998). Therefore, given the complexity of a complete system, actions should support holistic approaches (Jackson 2006) and focus on small changes that could provide positive feedback in the system (McDaniel and Driebe 2001).

A central theme touched on in the theories of complexity concerns the representation and modeling of systems using tools and techniques (Sharif and Irani 2006). Modeling is in fact inseparable from science and systems thinking. It is a traditional methodology that involves the use of formal models or simulations to analyze complex systems and, consequently, to better understand them and their actions (Trochim et al. 2006). A model is a simplified representation of reality that allows individuals to better understand the key aspects of a complex situation (Lyons et al. 2003). Models have two main roles: integrating knowledge from other sources and acting as sources of knowledge themselves. In fact, models are a representation of knowledge that promote understanding and learning and prove indispensable for improving decision-making processes. Although different modeling techniques exist by which one can represent and study aspects of complex and dynamic systems, the technique that has received the most attention in recent years is agent-based simulation (ABM). The basic unit of this type of simulation is the agent, which operates on the basis of knowledge localized to a specific environment, and although each agent is equipped with only limited skills, together the agents ensure that the system that they form shows a certain kind of general behavior. Therefore, the model used in this study, which uses technical concepts and tools borrowed from game theory, biology and AI, aims to describe a system's operation in terms of the micro-level behavior of individual agents. Multi-agent systems are designed, therefore, to break down the emergent properties of the system into their elementary constituents, for example, an economic system into individual economic agents. The remote origins of multi-agent modeling can be traced back to Forrester (1968), a professor of computer science at MIT and the author of the famous book Industrial Dynamics, in which he applied a dynamic systems analysis to economic cycles in industry for the first time (thus concluding the research of Marshall described above). The basic principle is that the Forrester behavior of a system is characterized, at some point, by the level of resources that are independent and their rate of change. Eighty years after Forrester, more impetus for the spread and development of multi-agent systems came from researchers at the Santa Fe Institute through the software platform SWARM. "Swarm intelligence" is another term for collective intelligence. Swarm intelligence manifests itself in large groups of agents, which individually show limited cognitive abilities, but that act in a coordinated manner within a structured group, yielding a collective intelligence.

Once again, biology led the way in this field of study, incorporating research tools from the physical sciences and their applications in robotics, and researchers in this area have developed a series of useful applications for the study of the dynamics of biological systems. For example, the study of the allocation of tasks within a colony of wasps has aroused much curiosity in the scientific community because it was interpreted as resulting not from a hierarchical structure, but from the internal self-organization of the system. Each insect follows a few rules and continually exchanges information with its group through direct and indirect contact. In this way, some collective properties emerge that make the colony capable of organizing its activities in an efficient, flexible and robust manner. In fact, the organization of these systems is efficient because complex collective properties emerge from a few simple rules that are widely distributed, rather than using complicated centralized rules. In addition, the system is flexible because it can adapt to environmental changes and is robust because it would work even if some agents were not able to perform their tasks properly. This approach proves effective even with very simple individual agents in which the amount of intelligence is minimal, as in the case of wasps, ants, bees and all social insects in general.

Starting with biological systems, this type of analysis has shifted to the study of social systems, including the actions of economic agents, and has been used to simulate the consequences of actions and interactions between economic agents. Unlike classical economics research, in this case we conceive of the economy as a complex system of which no agent has full knowledge or an adequate representation. In this perspective, the agents are to the ants as the economy is to the nest. Consequently, the complexity of the market cannot be explained by analyzing consumers as individual players, let alone by studying the application as an aggregate phenomenon; it can only begin to be understood by shifting the focus to the actions and interactions between consumers or to the behaviors and interactions of individual companies. The economic sectors in which simulations and agents are commonly used are public economics, industrial economics and finance. Simulation techniques differ between sectors; for example, environmental economics (or even in agriculture) mostly uses the dynamic systems approach, whereas micro-simulations are used primarily in economics for predicting the effects of economic policy measures on public attitudes. The agent-based simulations lie at the boundary between the use of simulations to make predictions and calculations and their use for the purposes of understanding and explanation. The Environment Agents Rules are a typical example, which take into account different levels of explanation in the construction of the simulation and agents. The most important level is the environment, namely the context in which events occur, the rules of interaction are clarified, and managers decide which agents should change their behavior based on available data (Terna 2000). Action is different from simple behavior by virtue of the fact that in the former, unlike the latter, some mental causes can be isolated. In fact, in this case we are dealing with so-called cognitive agents, or BDIs (Beliefs, Desires, and Intentions), that agents build in a way that expresses intentional actions. The BDI model is based on three fundamental components: (1) the beliefs that represent the agent's knowledge about the world; (2) the agent's desire to reach its desired state; and (3) intentions, which represent persistence in achieving goals. Intentions play a major role because they drive the actions of the agents and the choice of when to pursue a goal and when to quit the action, thus dictating the agent's strategy. As demonstrated by Cohen and Levesque (1990), an agent will abandon an intention when it has been satisfied, i.e., when the target has been reached or, conversely, when it became too difficult to satisfy. In this model, therefore, agents are able to construct hierarchies of objectives and sub-goals, in part through translatable skills, in the process of instrumental rationality that identified the goal and set in motion a strategy for seeking out the best means for achieving it. The cognitive representation typical of an agent can therefore be defined as an internal state that allows the agent to act on informing the structures of entities or activities located internally or externally. A representation has a semantic content that can play a causal role in the action of the agent. The semantic content, or the mode of presentation of a concept, gives the agent information about this concept that it can use for an inference or action. If an agent has two representations—(1) the effect that if x is F, and (2) the effect that if x is F—then it can produce the action A. If the agent is motivated and actually performs A, then its performance at that time is explained by Fx.

In addition, consider the case of Fx Fx, where the agent has a representation of x defined as G, and the agent believes that gx implies that A is not possible; then, the mode of presentation (F or G) is one of the constituents of the agent's causal action. As demonstrated by Dretske (2006), the action of a system, in addition to being caused by the representation, has the advantage of providing an explanation for the action. However, it is one thing for an action to be caused by an event that has a certain content, but quite another for it to be explained by the fact that the event has a certain content. The problem, according to Dretske, is that another agent may, in theory or in practice, construct an argument whose premises are causally effective and terminate the action. Ultimately, it seems imperative that there be a "portraying" criterion. But a causal model and mechanical properties should not intervene as "being explained by another agent". For this reason, we will explore one other approach. The causal theory of action (Davidson 2004; Goldman 1976; Searle 1983; Velleman 2000) postulates that these representations, which include both causes and reasons, are normally propositional attitudes (AP). An AP is the intentional relationship between an agent and a proposition (or statement). It is available from a certain epistemic agent in relation to a proposition. This provision falls under two categories: the provision for a proposition to hold true (in the case of beliefs) and to maintain the truth of propositions or utterances (in the case of desires). Thus, an agent, A, can articulate a predicate, p, (e.g., a thought, wish or hope) and a propositional content, Fx. The general form of the AP is thus: p (AFX). These representations fall under the categories of doxastic representations (beliefs) and volitional representations (desires), which are classically characterized by their directions of adjustment: the world into cognition (for beliefs) and cognition to the world (for desires). One of the characteristics of the AP is that the transitive property does not hold; if we suppose $a = b$ and p (AFA), it does not necessarily follow that p (AFB) is true. You can have the belief that Paris is a beautiful city without this having the effect that you believe that Paris is the capital of France (if you do not know), although Paris is the capital of France. So, Fa and Fb can have the same reference (Paris), but different meanings, and agents may have different thoughts but appear intentionally and extensionally identical. Propositional attitudes are semantically evanescent, but their intent makes the truth of their conditions difficult to clarify, and one must appeal to extension. The intentionality attributions of propositional attitudes signal the intentionality of mental states. APs, according to the standard causal theory, play an important role in parametrically, strategically and socially rational actions. At the parametric level, they are causes of actions; at the strategic level, they are involved in the allocation of rationality to another agent (A believes that B wishes x): they refer to other APs. Finally, at the social level, beliefs are socially shared language exchanges and economic conditions.

To say that beliefs and desires represent the AP is a philosophical thesis that Russell (1985) put forth as one of the dogmas of rationalism. The corollary is that an agent has beliefs and desires only if he can have a relationship with a proposition (or statement), which implies that language is a necessary condition for rationality. If it is undeniable that the APs are beliefs and desires, it is less certain that the APs are the causally effective set of representations that guide rational action. In fact, other types of representations used in control engineering and robotics have directions of adjustment and causal roles that cannot be considered APs. Newell (1990), for example, identifies two general classes of representations: propositions and models. A proposition is a way of presenting information that is extremely rich in expression (mathematics and logic can be expressed in this way) and size: the symbol VEHICLES (x) refers to all vehicles and collects, in one category, different objects, such as boats, space shuttles and elevators. Propositional representations are formed by a set of discrete and arbitrary symbols (which have no similarities with their reference) according to the rules of composition. They involve quantifiers and logical connectors to derive a representation of another representation (Evans 1982). The components must be able to generate propositional representations that can be generally recombined for other uses. Thus, an agent, by representing F (a) must be capable of, or willing to, generate the set of propositions where F (x): (F (a), F (b), F (c)…) (G (a), G (b), G (c)…). In conclusion, the propositional representations refer to reality by virtue of their truth-conditions: the reference is what makes the proposition true.

The agents and cognitive architectures of control of the simulation models are simplified representations of external entities. To build a model of an entity, E, the control device must have information on the variables, and it must be possible to measure the values of these variables, the evolution of these variables and the structure of E (Zeigler et al. 2000). In a system (1) modeling the evolution of another system (2), when some variables and relationships from system 1 are coupled with variables and relationships from system 2, the pattern "follows" the evolution of the system 2. The model refers to E according to a principle of structural correspondence: each part of the model is an aspect of E (provided that the model is appropriate). Thus, a robot can explore its environment and synthesize a spatial representation of the geometry and topology of the environment. Later, when it chooses a route, it should consult the map it has created (Stein 1994). If an obstacle presents itself, the robot reconfigures its route. By contrast, in a sentence like "$\forall$ (x) {($\neg$ F (x) $\vee$ G (x)) $\supset$ H (x)}", the symbols $\forall$, $\neg$, $\vee$ and $\supset$ do not refer to properties or entities. According to Newell, the propositions are distinguished by their size and the efficient treatment of information (Newell 1990). In fact, a model makes it more difficult to abstract concepts such as denial, separation or conditional and presents only some dimensions of the represented entity, whereas the power expression of the propositions is virtually unlimited. The representation model does not assume, furthermore, either an infinity of recombination nor a constriction of generality. A model is constructed from a synthetic representation of the sensors by a flood of information, bias and learning contained in the modules. When a model is built, the system can be simulated: the variables of the model are coupled to a time variable and, according to the information available, the model can calculate a future state of E. Despite their lack of precision, control architectures based on the models have shown great effectiveness in real-time control in the field of robotics. These architectures have proven particularly effective because of their ability to anticipate disturbances and recalibrate the system in real time. Industries where error is very expensive or the margin for error is very small require accurate, stable models that can tolerate uncertainty. Rather than following a reactive pattern (action $\rightarrow$ effect), the control-based model follows a premonition or projective pattern (desired effect $\rightarrow$ action to be applied). Verification within the model is a good compromise between a purely reactive control that reacts to a detected perturbation only once and control based on propositional knowledge. The former does not have predictive ability, and the latter must encode information in the form of propositional inferences to generate an appropriate action and then re-encode the information to implement the control. The latter approach is more effective in a deterministic and limited environment.

In conclusion, simulations represent a promising way to represent reality to provide researchers with a starting point of great value and scientific rigor. What distinguishes simulation from other media (language, mathematics, and graphics) is their different degrees of intelligibility. In fact, math and language must be understood by another human mind and therefore must be shared. Programming languages used in simulations do not need to be understood, as they are intended for computer models and are not the means by which the theory is intended to circulate. Unlike traditional theories and methods, scientists do not need to interpret the meanings and assumptions because the simulations generate the phenomena, leaving the researcher with the sole task of observing the phenomena produced by the theory. However, the agent-based methodology is relatively young, and some difficulties with this methodology remain to be resolved. It probably will not be able to create models that are sophisticated enough to explain some decisions that can have enormous significance for the individual. The development of the agent-based methodology is in fact still in its early stages, but it appears to be a promising addition to the theories and methods already in use.

References

Anderson, P. (1999). Complexity theory and organization science. *Organization Science, 10*(3), 216–232.

Antsaklis, P. J. (1994). Defining intelligent control. Report of the task force on intelligent control. *IEEE Control Systems Magazine,* 14, 4–66.

Ashmos, D. P., Duchon, D., & McDaniel, R. R. (1998). Participation in strategic decision making: the role of organizational predisposition and issue interpretation. *Decision Sciences, 29*(1), 25–51.

Auyang, S. Y. (2004). *Engineering: An endless frontier*. Cambridge: Harvard University Press.

Bergson, H. (1996). *L'évolution créatrice*. Paris: PUF.

Bermúdez, J. L. (2000). Rationality, logic and fast and frugal heuristics: Review of Gigerenzer et al. *Behavioral and Brain Sciences, 23*, 744–745.

Bertuglia, C. S., & Vaio, F. (2011). *Complessità e modelli*. Bollati Boringhieri: Torino.

Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Oxford University Press.

Chiva-Gomez, R. (2004). Repercussions of complex adaptive systems on product design management. *Technovation, 24*, 707–711.

Chu, D., Strand, R., & Fjelland, R. (2003). Theories of complexity: Common denominators of complex systems. *Wiley Periodicals, 8*(3), 19–30.

Clark, N., & Juma, C. (1988). Evolutionary theories in economic development thought. In C. Dosi, R. Freeman, G. Nelson, & L. Soete (Eds.), *Technical change and economic theory* (pp. 197–218). New York: Pinter Publishers.

Cohen, P. R., & Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence, 42*, 213–261.

Cournot, A. (1838). *Recherches sur les principes mathematiques de la theorie des richesses*. Paris: L. Hachette.

Darden, L. (2005). Relations among fields: mendelian, cytological and molecular mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences, 36*, 349–371.

Darwin, C. (1967). *L'origine della specie*. Torino: Bollati Boringhieri.

Davidson, D. (2004). *Problems of rationality*. Oxford: Oxford University Press.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher, 35*(3), 125–129.

Dretske, F. I. (2006). Minimal rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 107–115). Oxford: Oxford University Press.

Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.

Forrester, J. W. (1968). *Principles of systems*. Cambridge: Wright-Allen Press.

Forrester, J. W. (1975). *Collected papers of Jay W. Forrester*. Portland: Oregon Productivity Press.

Foucault, M. (1998). *Le parole e le cose*. Rizzoli: Milano.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*(3), 342–353.

Glouberman, G., Zimmerman, B. (2002). Complicated and complex systems: What would successful reform of medicare look like? *Commission on the future of health care in Canada*. Discussion Paper N 8.

Goldman, A. I. (1976). *A theory of human action*. Princeton: Princeton University Press.

Goodman, N. (1984). *Faits, fictions et prédictions*. Paris: Minuit.

Hodgson, G. M. (1993). *Economics and evolution, bringing life back into economics*. Cambridge: Polity Press.

Hull, D. L. (1974). *Philosophy of biological science*. New York: Prentice-Hall.

Jackson, M. C. (2006). Creative holism: a critical systems approach to complex problem situations. *Systems Research and Behavioral Science, 23*(5), 647–657.

Keynes, J. M. (1924). Alfred Marshall, 1842–1924. In G. Keynes (Ed.), *Essays in biography* (pp. 125–217). New York: Norton & Company.

Kitcher, P. (1984). 1953 and all that: a tale of two sciences. *Philosophical Review, 93*(3), 335–373.

Kroes, P. (1998). Philosophy of technology. In E. Craig (Ed.), *Routledge encyclopedia of philosophy* (http://www.Rep.Routledge.Com/article/q103sect3). London: Routledge.

Lewontin, R. (1974). *The genetic basis of evolutionary change*. New York: Columbia University Press.

Licata, I. (2011). *Complessità. Un'introduzione semplice*. Palermo: Duepunti Edizioni.

Limburg, K. E., O'Neill, R., Costanza, R., & Farber, S. (2002). Complex systems and valuation. *Ecological Economics, 41*, 409–420.

Lyons, M.H., Adjali, I., Collings, D., Jensen, K.O. (2003). Complex systems models for strategic decision making. *BT Technology Journal*, *21*(2), 11–27.

Machamer, P., Darden, L., & Carver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*(1), 1–25.

Maguire, S., Mckelvey, B., Mirabeau, L., & Oztas, N. (2006). Complexity science and organization studies. In S. R. Clegg, C. Hardy, T. B. Lawrence & W. R. Nord (Eds.), *The sage handbook of organization studies* (pp. 165–214). New York: Sage Publications.

Marshall, A. (1890). *Principles of economics*. London: Macmillan and Co.

Marshall, A. (1919). *Industry and trade*. London: Macmillan and Co.

Marshall, A. (1991). Analogies mécaniques et biologiques en économie. *Revue Française d'Economie, 6*(1), 103–113.

Mayr, E. (1982). *Storia del pensiero biologico. Diversità, evoluzione, eredità*. Torino: Bollati Boringhieri.

McDaniel, R. R., & Driebe, D. J. (2001). Complexity science and health care management. *Advances in Health Care Management, 2*, 11–36.

Meadows, D. H., & Robinson, J. (1985). *The electronic oracle: Computer models and social decisions*. New York: Wiley.

Ménard, C. L. (1979). Equilibre, déséquilibre, temps: Un peu d'histoire. *Economie Appliquée, 32*(2–3), 229–251.

Montague, R. P. (2006). *Why choose this book? How we make decision*. New York: Penguin.

Morçöl, G. (2005). A new systems thinking: implications of the science of complexity for public policy and administration. *Public Administration Quarterly, 29*(3), 297–320.

Morgan, M. S., & Morrison, M. (1999). *Models as mediators: Perspectives on natural and social sciences*. Cambridge: Cambridge University Press.

Morin, E. (1990). *Introduction à la pensée complexe*. Paris: Éditions E.S.F.

Nagel, E. (1961). *The structure of science*. London: Routledge & Kegan Paul.

Nelson, R. (1995). Recent evolutionary theorizing about economic change. *Journal of Economic Literature, 33*(1), 48–90.

Newell, A. (1990). *Unified theories of cognition*. Cambridge: Harvard University Press.

Pievani, T. (2005). *Introduzione alla filosofia della biologia*. Laterza: Bari-Roma.

Quine, W. V. (1986). Reply to Morton White. In L. E. Hahn & P. A. Schilpp (Eds.), *The philosophy of W. V. Quine* (pp. 663–665). Illinois: Open Court.

Robbins, L. C. (1935). *An essay on the nature and significance of economic science*. London: Macmillan and Co.

Ruse, M. (1971). Reduction, replacement, and molecular biology. *Dialectica, 25*(1), 39–72.

Russell, B. (1985). *The philosophy of logical atomism*. Illinois: Open Court.

Ryle, G. (1949). *The concept of mind*. London: Penguin Books.

Saridis, G. N. (1985). Foundations of the theory of intelligent controls. In G. Saridis & A. Meystel (Eds.), *Proceedings of the IEEE workshop on intelligent control* (pp. 23–28). New-York: Troy.

Schaffner, K. (1967). Approaches to reduction. *Philosophy of Science, 34*(2), 137–147.

Schaffner, K. (1969). The Watson-Crick model and reductionism. *British Journal for the Philosophy of Science, 20*(4), 325–348.

Schick, F. (1991). *Understanding action: An essay on reasons*. Cambridge: Cambridge University Press.

Schumpeter, J. A. (1954). *History of economic analysis*. New York: Oxford University Press.

Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.

Sharif, A. M., & Irani, Z. (2006). Applying a fuzzy-morphological approach to complexity within management decision-making. *Management Decision, 44*(7), 930–961.

Simon, H. A. (1962). The architecture of complexity. In H. A. Simon (Ed.), *The sciences of the artificial* (pp. 192–229). Cambridge: MIT Press.

Snow, C. P. (1993). *The two cultures*. Cambridge: Cambridge University Press.

Stein, L. A. (1994). Imagination and situated cognition. *Journal of Experimental & Theoretical Artificial Intelligence, 6*, 393–407.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Boston: Irwin-McGraw-Hill.

Tabery, J. G. (2004). Synthesizing activities and interactions in the concept of a mechanism. *Philosophy of Science, 71*, 1–15.

Terna, P. (2000). Economic experiments with swarm. A neural network approach to the self-development of consistency in agents' behavior. In F. Luna & B. Stefansson (Eds.), *Economic simulation in swarm. Agent-based modelling and object oriented programming* (pp. 73–104). Dordrecht: Kluwer.

Trochim, W. M., Derek, A. C., Milstein, B., Gallagher, R. S., & Leischow, S. J. (2006). Practical challenges of system thinking and modeling in public health. *American Journal of Public Health, 96*(3), 538–546.

Velleman, J. D. (2000). *The possibility of practical reason*. Oxford: Oxford University Press.

Vincenti, W. G. (1990). *What engineers know and how they know it: Analytical studies from aeronautical history*. Baltimore: Johns Hopkins University Press.

von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. New York: George Braziller.

Waters, K. (1990). Why the anti-reductionist consensus won't survive: the case of classical mendelian genetics. *PSA, 1*, 125–139.

Wimsatt, W. (2000). Emergence as non-aggregativity and the biases of reductionisms. *Foundations of Science, 5*, 269–297.

Zeigler, B. P., Praehofer, H., & Kim, T. G. (2000). *Theory of modeling and simulation: Integrating discrete event and continuous complex dynamic systems*. San Diego: Academic Press.