Bernard Walliser

Cognitive Economics

Springer

Cognitive Economics

Bernard Walliser

Cognitive Economics

Springer

Professor Bernard Walliser
Paris School of Economics PSE
48, boulevard Jourdan
75014 Paris
France

ISBN 978-3-540-71346-3 e-ISBN 978-3-540-71347-0

DOI 10.1007/978-3-540-71347-0

Library of Congress Control Number: 2007937634

Original version in French language published as “L’Economie Cognitive”, © Odile Jacob, 2000

Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Coverdesign: WMX Design GmbH, Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Introduction

> *We need a theory of knowledge and of the ways human beings process knowledge as the foundations of our theory of economics.*
> H. Simon (1999)

During the last half-century, economic science has collided with two major criticisms leveled against the two pillars on which it is founded, namely individual rationality and collective equilibrium. On the one hand, individual behavior appears to be too idealized, since it attributes perfectly rational behavior to an actor without making explicit the process of mental deliberation on which it is grounded. On the other hand, collective equilibrium appears to be much too virtual, since it assumes that the actors coordinate on some equilibrium state without exhibiting the process of dynamic adjustment through which that state is attained. Responding to those criticisms, two research programs have been jointly developed, introducing the mental and temporal dimensions of individual and collective entities more deliberately into economic models.

Over the last thirty years, the "epistemic program" has emphasized the actors' cognitive achievements as a major explanatory factor for their behavior, and consequently for the economic phenomena resulting from their conjunction. Individually, each actor is endowed with personal 'beliefs' concerning his environment, which act as intercessors between theinformation he receives and the expectations he builds up. Collectively, all actors are involved in communication 'networks' structuring their permanent relations, which evidently act as supports for their material encounters, but essentially as vehicles for their exchanges ofinformation. Such (bounded) cognitive rationality has been highlighted in economics with regard to financial speculation involving crossed expectations, but it is also at work in price bargaining involving mutual beliefs.

Over the last twenty years, the “evolutionist program” has stressed the actors’ dynamic learning processes as an essential component of their interactions, and consequently of the economic changes stemming from them. Individually, each actor performs certain mental ‘inferences’ transforming his various beliefs, which act as intercessors between the assumptions he holds and the actions he intends to take. Collectively, all actors are animated by adaptive ‘processes’ conditioned by their environment, which allow them to co-evolve and may globally induce certain emergent structures. Such (procedural) interactive temporality has been adopted in economics for job seeking involving random search, but it is also at work in the diffusion of (technological or institutional) innovations involving neighborhood imitation.

Over the last ten years, in keeping with the ‘cognitive turn’ influencing all social sciences, these two programs have tended to join in one single stream labeled “cognitive economics”, allying the mental and dynamic points of view. The traditional ‘homo oeconomicus’ has been succeeded by an original ‘homo cogitans /adaptans’, for whom the recurrent work of time may compensate for his limited cognitive capabilities. The traditional ‘equilibrium state’ has been replaced by an original ‘self-organizing mechanism’, through which social structures may emerge from myopic adjustments of messages and actions. Hence, cognitive economics can be defined as the study of the reasoning operations and adaptation processes followed by economic agents in an interactive and dynamic setting.

Historically, the epistemic research program has developed in three steps, by shifting the object of study from an external view of input-output behavior to an internal investigation of deliberation supported by mental states. Firstly, the actor’sinformation is understood to summarize the data about the past history of his environment, gathered in an involuntary or in a more deliberate way and assumed to be univocally interpreted. Secondly, the actors expectations are analyzed as representations of the future path of his environment, representations that he adapts through time according to his past observations. Thirdly, the actors beliefs are structured in autonomous models of the relevant features of his environment, which he revises in keeping with the factual or structural messages he continually receives.

The emergence of the epistemic paradigm has been favored by the availability of new logical and mathematical tools, capable of giving a formal account of reasoning operations. Firstly, epistemic logic, which is a variety of modal logic, qualitatively formalizes the hierarchical belief structure an actor holds about himself and others, and specifies the inference methods he uses for transforming it. Secondly, probability calculus, extended from additive to non-additive probabilities, gives a quantitative account of an actors subjective beliefs and suggests different mechanisms for combining and transforming them. Further tools, such as neural networks, have added inductive, analogical or taxonomical operations to the usual deductive processes of an actor.

The epistemic research program has been influenced by certain other social sciences, also interested in the social consequences of an actor's beliefs and reasoning. Social psychology showed how interacting agents may be locked in pathological situations resulting from crossed expectations which are kept through time because they are locally validated (Watzlawick, 1970). Sociology demonstrated how desirable or undesirable social states can be obtained and selected by self-fulfilling prophecies about a phenomenon of common interest among agents (Merton, 1936). Philosophy of language brought to light the way in which cooperating agents may be coordinated on some salient social configurations by common beliefs acting as shared conventions (Lewis, 1969).

But the epistemic paradigm entertains particularly close relations with cognitive science, which, even if only recently appeared and institutionalized, has tackled head-on the question of actors' modes of reasoning. Bounded rationality was modeled by introducing limited computational capabilities faced with the complexity of the problem an agent has to solve.information management was underlined by making explicit the trade-off between exploration (in order to get moreinformation for later choice) and exploitation of available information (for computing the present choice). Distribution of (crossed) knowledge among agents was duly categorized, by making a sharp distinction between distributed knowledge, shared knowledge and finally common knowledge.

The evolutionist[1] research program also developed historically in three steps, by shifting the object of study from the actor's simulated time to the system's concrete time. Firstly, in crushed time, with regard to the information he holds about his fixed environment, the agent takes all current decisions at the present time then implements them instantaneously. Secondly, in spread out time, after predicting the future

---

[1] The term 'evolutionist' is used to qualify any dynamic process while the term 'evolutionary' is used to qualify a dynamic process inspired by biology

state of his (evolving) environment, the agent makes all his decisions at the initial time and implements them afterwards, without calling them into question again. Thirdly, in sequential time, the agent adapts his decisions in each period to his current beliefs, which are themselves revised according to the evolution of his perceived environment.

The emergence of the evolutionist paradigm has benefited from the availability of new mathematical instruments, which allow a better consideration of the evolution of the social system. Firstly, non linear dynamics, popularized in meteorology, justifies sequences of behavior that are very sensitive to initial conditions and that converge towards punctual, cyclical or chaotic equilibria. Secondly, stochastic processes, originating in statistical physics, describe opinion or action trajectories which are strongly unpredictable and subject to bifurcations depending on both context and history. Finally, more applied tools like genetic algorithms enable the step-by-step simulation of more complex learning processes of interacting agents combining selection and mutation.

The evolutionist research program has been sporadically enriched by other social sciences also interested in dynamic organizational processes. Psychology has drawn attention to the role of reinforcement mechanisms based on an agents past performance, and the possibility of their convergence toward suboptimal situations (Bush, Mosteller). Sociology has underlined the unexpected effects resulting from the composition of human actions, especially the emergence of original global structures from merely local and plain interactions (Boudon, 1981). Political science has brought to the fore the existence of various nested institutions regulating actors' behaviors, generated from the progressive amplification of germs and hazards (North, 1990).

The evolutionist research program has also taken advantage of the development of cognitive science, which has examined from its very beginning adjustment processes among actors, even if crudely. It has formalized adaptive rationality, proposing heuristics adopted by the agents, transforming progressively interpreted past experience into actions directed towards progressively expressed goals. It has conceptualized the diffusion ofinformation, relating agents through mimetism or contagion, with the possible emergence of common knowledge if the communication network is sufficiently dense and reliable. It has also explored the coordination of action, achieved by decentralized reasoning and learning of agents, with a possible convergence due to the work of positive and negative feedbacks.

The epistemic and evolutionist programs lead in fact to inter-related developments in the fields of both individual behavior and collective coordination. They act as substitutes when an agent compensates for limited reasoning with a time-improving learning process or, conversely, when he uses a more active and conscious reasoning system to bypass a 'lock-in' caused by a myopic adaptation process. They act as complements when an agent behaves rather reactively to others' actions over the short term in simple settings and by sophisticated expectations about others' behavior over the long term in complex settings. In fact, ever since A. Smith, the 'official' founder of economics, compared the specialization process acting for material production and for acquisition of knowledge, many economists have stressed the cognitive foundations of economics and the economic constraints on cognition.

In one direction, cognitive principles were soon applied to the basic economic entities in order to explain phenomena for which they appear to be essential factors. For instance, J. M. Keynes stressed the influence of differential states ofinformation and degrees of reasoning on expectations, and their eventual stabilization on conventional values in the case of radical uncertainty. F. von Hayek called attention to the distributed character of price signals in the exchanges on markets, and to the influence of evolutionary processes in the emergence of markets. T. Schelling highlighted the importance of crossed expectations among the players in a strategic context and R. Aumann emphasized the coordinating role played by these beliefs in the existence and selection of equilibrium states.

In the other direction, economic principles have been applied to actors' cognition, considered as a scarce resource, to show how this activity is efficiently managed. For instance, A. Marshall suggested that firms should use certain production routines progressively adapted through time in order to economize on their cognitive resources. H. Simon analyzed the internal cognitive constraints faced by an actor when gathering and treatinginformation, and the implications of his bounded rationality for the design of multi-agent organizations. R. Stigler studied the problem of the acquisition ofinformation from external sources in terms of a trade-off between its cost and expected utility and J. Muth studied the problem of the optimal use ofinformation in the formation of expectations.

As far as theory is concerned, Cognitive economics applies to both levels of generality usually considered in the economic literature, namely game theory on a general level and exchange economics on a more specific one. It enriches and softens the traditional vision of game theory, shared by other social sciences, i.e. strategic bilateral interactions between undifferentiated players in a non-institutional context. It contributes to and corrects the usual point of view of exchange theory, specific to economics, i.e. passive exchanges of goods between specialized agents mediated by a special institution, the market. Moreover, the empirical dimension of cognition is growing in line with the fast development of experimentation game theory and observation of the 'knowledge economy'.

Cognitive economics, through the purely epistemic and evolutionary streams which are components of it, is strongly linked to several other research programs. For instance, it is associated with the "behavioral program", which studies how heterogeneous agents think and adapt in various interactive situations, with a special emphasis on experimentation. Likewise, it intersects the "institutionalist program" which studies how institutions behave over the short term and emerge over the long term through individually-driven epistemic and evolutionary factors. Finally, it is related to "social cognition" which extends the influence of cognitive factors to cover the whole social domain, but which concentrates on phenomena of collective communication rather than individual reasoning.

Several recently-published books already deal with cognitive economics, but are mere collections of articles. One gathers many theoretically oriented articles written in the last century and considered as classics of the movement (Egidi, Rizzello). Four others are concerned with technical developments and more empirical works, following the very first conferences (Rizzello, 2003); (Kokinov, 2005); (Topol, Walliser) and summer schools (Bourgine, Nadal) hold in the field. The present book aims at a structured presentation of the main messages of cognitive economics by stressing its constructive aspects rather than its critical ones. It cannot claim to be a synthesis, since the domain is still under development and has fuzzy frontiers, but it may be a benchmark for economists, cognitive scientists as well as philosophers.

The book is organized into four parts, over the course of which the actor's behavior becomes more and more precisely and completely specified in an ever-enriched social context. Each part is itself developed into two chapters, the first of which introduces the main concepts in a static perspective, while the second extends these concepts into a more dynamic perspective. Each chapter reinterprets, in an explicitly cognitive perspective, classical achievements as well as more original developments, either formalized or not, and mentions existing experimentation works in support of them. Moreover, every section is illustrated by a recurrent example (presented in italics) giving the actor a different role in each part, successively soothsayer, surgeon, driver and worker.

Chapters 1 and 2 consider the actor as a reasoner, who holds imprecise beliefs and bounded reasoning capacities and tries nevertheless to decipher his surrounding environment as best he can. Chapters 3 and 4 consider the actor as a decision-maker, whose beliefs reflect his uncertainty about his environment and are associated with preferences in order to make an appropriate choice. Chapters 5 and 6 consider the actor as a player who uses his beliefs in order to simulate his opponents' behavior and to achieve certain coordinated actions with them. Chapters 7 and 8 consider the actor as an economic agent, acting in an institutional environment and able to buyinformation on the market and to accumulate knowledge as an immaterial capital[2].

---

[2] I thank R. Crabtree for his accurate revision of the translation of the book

1

Structure of individual beliefs

Knowing ignorance is strength
Ignoring knowledge is sickness.
Lao Tseu

Any actor involved in a social system holds certain subjective beliefs about his physical, social and institutional environment as well as about his own characteristics. These beliefs, expressed in a linguistic form, are evaluated in terms of their internal logical consistency, but even more in terms of their external suitability to the gathered information. The actor's view of the system is evaluated by reference to the modeler's view, which includes the actors' beliefs and is considered by construction as perfect and complete. In contrast, the actor only has a noisy and partial model of the system, or even a wrong model of it, since he has to deal with limited information and bounded cognitive rationality.

In epistemic logic, any belief can be expressed according to one of two approaches, the syntactic (propositional) one (1.1) and the semantic (possible-world) one (1.2). These two contrasted approaches are logically equivalent (1.3) and can be simultaneously extended from a set-theoretic framework to a probabilistic one (1.4). beliefs are assumed to obey certain cognitive rationality axioms, which express the actors reflexive and deductive reasoning capacities and are again stated in syntax (1.5) or in semantics (1.6). An actor's hierarchical beliefs can be studied in a more systematic way, either beliefs about his own beliefs (1.7) or beliefs about other actors beliefs (1.8).

1.1 Syntactical beliefs

The modeler's aim is to represent formally a 'system' exclusively composed of a 'physical environment' and of several 'actors' holding beliefs about that environment. In a syntactic framework, the exclusive units of knowledge are logical propositions expressed in formal language. Basic propositions p, q, r simply describe what the physical environment looks like. Further propositions concern the beliefs held by an actor about that environment and are defined by a 'belief operator' B_i such that B_ip means that 'actor i believes proposition p'. Moreover, compound (well-formed) propositions ϕ are obtained by applying logical connectors to the above: $\neg$ (negation), $\bigwedge$ (conjunction), $\bigvee$ (union), B_i (belief).

Propositions can be made more precise as concerns their interpretation. 'Factual propositions' assert facts, i.e. basic properties associated with a specific system. They are generally expressed by 'categorical propositions' of form ϕ. Facts may be material (like p) or epistemic (like B_ip). 'structural propositions' assert laws, i.e. structural regularities linking several properties of a generic system, and are expressed by 'conditional propositions' of type 'if ϕ, then ψ'. The conditional may be of type '$\phi \rightarrow \psi$' where $\rightarrow$ is a material implication (true as soon as ϕ is false or ψ is true) or of type '$\phi > \psi$' where $>$ is a counterfactual operator (studied later, see 2.5). Laws may relate physical facts to each other (like 'if p, then q'), actors' beliefs to each other ('if B_ip, then B_iq') or facts to beliefs ('if p, then B_iq'), but they seldom relate beliefs to facts ('if B_ip, then q').

Such a language allows us to express not only physical properties and basic beliefs, but even beliefs about beliefs. beliefs about beliefs are either 'homo-hierarchical' beliefs (beliefs of an actor about his own beliefs, such as B_iB_ip) or 'hetero-hierarchical' beliefs (beliefs of an actor about others' beliefs, such as B_iB_jp). A k-level proposition is a proposition where the belief operator appears at most k times. Since, all actors are assumed to hold beliefs about the same basic propositions as the modeler does, the interpretation of a proposition has to be the same for the modeler and every actor. When writing B_{ip}, one already assumes that the meaning of p is the same for the modeler and for all the actors. Moreover, when writing B_iB_jp, one assumes that actor i considers that actor j attributes the same meaning to p as he does.

According to belief operator B_i, $B_i\phi$ expresses the necessity for actor i of proposition ϕ, i.e. the highest degree of belief assimilated to 'endorsed belief'. If we now introduce an 'acceptance operator' L_i, $L_i\phi$ expresses the possibility for actor i of a proposition ϕ, i.e. the lowest degree of belief assimilated to 'accepted belief'. However, the two operators are considered as dual since they are defined jointly: an actor i accepts a proposition p if he does not believe in its negation: $L_i\phi = \neg B_i \neg \phi$. It is easy to prove that belief entails acceptance. Hence, for actor i, any proposition ϕ can be classified into three categories : $B_i\phi$ means that it is believed, $B_i\neg\phi$ means that its negation is believed, $N_i\phi$ means that neither the proposition nor its negation is believed.

Among all well-formed propositions of the language, the modeler considers as true a subset of propositions constituting a 'syntactical structure' $\mathbf{K}$. This contains a set of basic propositions (physical structure) and sets of propositions believed by each actor (individual belief structures $\mathbf{K}_i$). The structure is assumed to be formed of propositions which satisfy the following properties. It is 'non-contradictory' in the sense that if a proposition is true, its negation is not true. It is 'deductively closed' in the sense that if some propositions are true, their logical consequences are true too. It is 'complete' in the sense that either a proposition or its negation is true. Consequently, the modeler appears to play the role of God, since he knows exactly what is true as concerns the environment and the actors' beliefs. He can even be endowed with a special belief operator B_M such that $B_M\phi = \phi$. Conversely, each actor may hold false beliefs and his individual belief structure may be contradictory, unclosed or incomplete (see 1.5).

It is possible to compare two belief structures based on the same set of basic propositions. By definition, the structure $\mathbf{K}'$ is more 'accurate' than the belief structure $\mathbf{K}''$ if, for each proposition of the second, there is a corresponding, less precise proposition of the first. A proposition ϕ is said to be less precise than a proposition ψ under some technical conditions. Of course, such an order is only a partial one since two structures generally share a common set of beliefs, and that set is then completed in each structure by specific beliefs. For a single actor, one polar case (certainty) is obtained when he believes all the propositions that the modeler considers. The other polar case (complete uncertainty) is obtained when he believes no proposition at all.

As a standard example, consider a pack of double-sided cards, each card having a color (blue, yellow or red) on one face and a number (0 or 1) on the other. A card is taken at random from the pack and a soothsayer has to guess which card it is. Categorical propositions are, for instance, b = *'the card is blue',* y = *'the card is yellow',* r = *'the card is red',* e = *'the card is even'. A conditional proposition specifies that a blue card is even:* $b \rightarrow e$, *while a red card is odd and a yellow card can be either. The soothsayer* i *is color-blind and, if the true card* *is red, he believes that the card is yellow or red, but he cannot be more precise:* $B_i(y \bigvee r)$*, but* $\neg B_i y$ *and* $\neg B_i r$*. An observer j may know that the soothsayer is color-blind:* $B_j B_i(y \bigvee r)$*, but conversely he may also (wrongly) think that the soothsayer has perfect sight:* $B_j B_i r$*.*

1.2 Semantical beliefs

The modeler's aim is again to represent the physical environment and the actors' beliefs about it. In a semantic approach, he considers 'possible worlds', assumed to be mutually exclusive and globally exhaustive, as basic units of knowledge. Each world represents all the relevant features of a given situation, including the actors' beliefs. One of the possible worlds is the 'actual world' denoted $w*$. The physical environment is symbolized by an 'assignment domain' $H_0(w)$ indicating which 'state' s (in a set of states) is associated with which world w. The beliefs of actor i are represented by an 'accessibility domain' $H_i(w)$ which groups together all the worlds that he cannot distinguish when the actual world is w. An event being defined as a subset of worlds, it is possible to combine the events by applying the usual set operators ^{c} (complementation), $\cap$ (intersection) and $\cup$ (union).

The worlds can again be detailed by inserting 'factual features' reflecting properties of a given system and 'structural features' reflecting laws imposed on all systems. By definition, 'potential worlds' are worlds for which the factual features may differ from the actual world while the structural laws are maintained. More drastically, 'virtual worlds' are worlds for which factual properties as well as structural laws are modified in relation to those of the actual world. If considering only potential worlds, the laws become satisfied in all worlds. Of course, the actor may consider not only potential worlds, but certain virtual worlds which are known by the modeler to be out of reach. Conversely, the actor may not be aware of worlds considered as potential worlds by the modeler.

The accessibility domains of an actor can be generated by an accessibility relation R_i, which relates a world to another when the second cannot be distinguished from the first by the actor: wR_iw' iff $w' \in H_i(w)$. A multi-agent graph can be defined by the combination of all agents' relations. Of course, it is possible to construct intertwined accessibility relations from the individual ones. For a single actor, $H_iH_i(w)$ defines the worlds that the actor i considers to be accessible to him. For two actors, $H_iH_j(w)$ defines the worlds that the actor i considers to be accessible to actor j. In both cases, the number of possible worlds increases compared with $H_i(w)$, which simply expresses an increase in uncertainty as we climb up the levels of (self or crossed) beliefs. Here again, it is assumed that each possible world is similarly interpreted by the modeler and any actor.

More precisely, every possible world w is a complete description of the state of the environment and of the actors' beliefs considered at any hierarchical level. It is perfectly characterized by the assignment and accessibility relations defined in this world (and known only to the model). Since the accessibility relations involve other similarly-defined worlds, the definition of the worlds is global and self-referent. Abstractly, it is always possible to consider that all worlds the modeler can think of are already defined in an infinite network. But practically, one considers only a finite set of worlds assumed to represent a specific situation, an assumption which introduces certain restrictions. As a consequence, if the system changes, the set of worlds may change drastically.

Finally, a 'semantic structure' $\mathbf{H}$ is defined by a set W of worlds, a set S of states, an actual world $w*$, an assignment relation H_0 and several accessibility relations H_i. The assignment relation defines the 'physical structure' and the accessibility relation defines the 'individual belief structure' of each actor. It can be observed that $\mathbf{H}$ gives rise to another structure just by changing the actual world $w*$. The structure is only submitted to some weak constraints: the assignment relation $H_0(w)$ defines one and only one state; the accessibility relations $H_i(w)$ are never empty.

It is again possible to compare two semantic structures defined on the same set of worlds (or even on different worlds). By definition, the belief structure $\mathbf{H}'$ is 'collectively more accurate' than the belief structure $\mathbf{H}''$ if, in each world w, and for each actor i, the accessibility domain of the first is included in the accessibility domain of the second: $H_i(w) \subseteq H_i''(w)$. The structure $\mathbf{H}'$ is 'individually more accurate (for player i)' than the structure $\mathbf{H}''$ under the following conditions: $H_i'(w) \subseteq H_i''(w), H_j'(w) = H_j''(w)$. Both orders are only partial orders, with given extremal elements. For a single agent, one polar case is obtained when $H_i(w)$ is a singleton for any w, i.e. when the actor always considers as accessible a unique world (which may however be false). The other polar case is obtained when $H_i(w)$ is the whole set of worlds for nay w, i.e. when the actor cannot discriminate between any worlds.

In the standard example, the possible worlds for a selected card are the four possible cards: $w_1 = (b, 0), w_2 = (y, 0), w_3 = (y, 1), w_4 = (r, 1)$. *These are in fact the virtual worlds, a potential world being* $w_5 = (r, 0)$

expressing a red card which is odd, or $w_6 = (g, 0)$ *expressing a green and even card. The event 'the card is even' is represented by the subset* $\{w_1, w_2\}$ *while the event 'the card is yellow' is represented by the subset* $\{w_2, w_3\}$*; a yellow and even card is precisely at the intersection of the two events. Conversely, the event* $\{w_1, w_3\}$ *cannot simply be interpreted in terms of colour and number on the card. If the soothsayer i observes only the colour face and is colour-blind, his accessibility relation is:* $H_i(w_1) = w_1, H_i(w_2) = H_i(w_3) = H_i(w_4) = \{w_2, w_3, w_4\}$*. If another observer sees the real colour and knows that the soothsayer is colour-blind, his accessibility relation is:* $H_j(w_k) = \{w_k\}$ *for any k.*

1.3 Link between syntax and semantics

The syntactical framework expresses an actor's beliefs in a linguistic way which stays close to the natural language formed of 'separate' sentences. In accordance with an 'individualistic theory of beliefs', any belief is defined independently and can be added or subtracted by keeping the others unchanged. Moreover, the belief set is 'segmented' since the beliefs may belong to separate 'fields' which are not linked together and which are treated independently. However, the syntactical framework is hard to handle in its logical form, since propositions cannot easily be combined. In order to make it nevertheless easier, the logical propositions may be given more precise formulations such as logical predicates (expressing a qualitative link between properties of objects in a given class) or functional relations (expressing a quantitative correspondence between magnitudes of objects in some class).

The semantic framework expresses beliefs in a way which is more abstract, but closer to the usual representation of scientists, either physicists (material states in mechanics) or psychologists (Johnson-Laird's mental states). In accordance with the 'holistic theory of beliefs', beliefs are the result of a whole structure which changes globally when new beliefs are received. The framework is 'geometric' because it is based on a space of comparable worlds. The semantic approach is therefore easily tractable, since it is based on operations on sets. In particular, it is possible to give more structure to the world space by defining weight distributions on it as well as distances.

Nevertheless, the syntactical and semantic approaches share some ontological assumptions about beliefs. Firstly, the beliefs are 'intentional' (rather than imaginary) since they refer to some actual physical system rather than to a virtual one, especially a physical or social artifact. Secondly, the beliefs are 'epistemic' (rather than praxeological) since they are defined independently of their further function, especially their use in decision-making. Thirdly, the beliefs are 'context-free' (rather than situated) since they refer to the system independently on contextual factors, especially the social environment of their design. It follows that the actors' beliefs are stated according to a common reference, hence can be rightfully compared.

In fact, the syntactical and semantic approaches are formally equivalent, since they are related by certain 'transcription principles'. Firstly, it is possible to associate a unique event with each proposition. This is called its 'field', and is composed of the set of worlds where the proposition is true. Secondly, the syntactic (logical) connectors of propositions are transformed into semantic (geometric) operations on events. With the negation, conjunction and disjunction of propositions are associated the complementation, intersection and union of events. Thirdly, the material states are just defined by the truth value of the basic propositions. In a given state, for each basic proposition, either that proposition or its negation is true.

For the transposition from semantics to syntax, the following transcription rules define the truth value of any proposition in a given world. Firstly, a basic proposition is true or false according to the assignment relation associated with the world. Secondly, a proposition is believed by an actor if the accessibility domain in that world is included in the field of the proposition, i.e. if the proposition in true in all worlds the actor considers as accessible in that world. It follows that the syntactical system associated with the semantic one has the required properties. Non-contradiction is obtained by the non-emptiness of the accessibility domains. Completeness is obtained since any event is true or false in some world. Deductive closure is obtained since modus ponens is automatically realized in semantics.

For the transposition from syntax to semantics, the construction of a set of worlds associated with a syntactical belief structure is harder since multivocal. This is because it is necessary to construct simultaneously the structured set of all possible worlds imagined by the actors. However, such a construction is possible, and some transcription rules apply simply when the worlds are already given. Firstly, in each world, the assignment domain groups either the basic propositions or their negations which are true in that world. Secondly, in each world, the accessibility domain is the intersection of all fields of propositions which are believed in that world. In fact, there exist several semantic structures associated with each syntactical one, all 'bisimilar' (Benthem, ter

Meulen). Other correspondences between syntax and semantics hold as well, especially the accuracy order between two belief structures.

Let us return to the standard example, in order to illustrate the comparison between two belief structures. In syntax, a soothsayer knows a card when he observes both faces and knows nothing when he observes neither face. An intermediate situation is obtained when he only observes one face. However, the belief structures when the soothsayer only observes the color and when he only observes the number cannot be compared in terms of their accuracy. In semantics, a soothsayer knowing nothing has an accessibility function: $H_i(w_1) = H_i(w_2) = H_i(w_3) = H_i(w_4) = \{w_1, w_2, w_3, w_4\}$. *A soothsayer knowing everything has an accessibility function:* $H_i(w_1) = \{w_1\}, H_i(w_2) = \{w_2\}, H_i(w_3) = \{w_3\}, H_i(w_4) = \{w_4\}$. *A soothsayer only observing the color has the accessibility function:* $H_i(w_1) = \{w_1\}, H_i(w_2) = H_i(w_3) = \{w_2, w_3\}, H_i(w_4) = \{w_4\}$. *A soothsayer only observing the number has the accessibility function:* $H_i(w_1) = H_i(w_2) = \{w_1, w_2\}, H_i(w_3) = H_i(w_4) = \{w_3, w_4\}$. *No general inclusion between accessibility domains can be observed when comparing the last two cases.*

1.4 Probabilistic beliefs

In the preceding approaches, any belief was defined either logically (in syntax) or set-theoretically (in semantics). Such a belief is only roughly expressed in an all-or-nothing form, since a proposition is simply believed or not (and accepted or not). But the belief can be more subtly stated in a quantitative form and especially in a probabilistic one. Probability then expresses a 'degree of belief' in a proposition and is subjective in the sense that it may differ for different actors. It may be linked to the reliability of the source of the belief or to the confidence of the reasoning leading to it. Such a probabilistic view was spontaneously introduced in decision and game theory in order to reflect the uncertainty of a player (see 3.5 and 5.5).

In syntax, each actor has a probabilistic belief operator such that $B_i^\alpha p$ means that 'actor i believes proposition p with at least probability α'. Such an operator is defined on both basic and compound propositions. There is no dual operator associated with a probabilistic belief operator or, more precisely, it coincides with the primal one. However, it is possible to consider the operator M_i^α such that $M_i^\alpha p$ means that 'actor i believes p with exact probability α; it is then natural to state that $M_i^\alpha p = M_i^{1-\alpha} \neg p$. In semantics, each actor has a probabilistic accessibility relation which associates, with each world w, a probability distribution on all worlds: $P_i(w';w)$. Since a world comprises physical features as well as actors' beliefs, the actor simultaneously forms a probability on both components.

The link between syntax and semantics is again defined by a formal transcription condition. An actor believes proposition p with probability α in some world w if and only if the probability of the field of p measured in world w is greater than α. Such related frameworks again enable the representation of homo-hierarchical as well as hetero-hierarchical beliefs. In syntax, the individual probabilistic belief operators can be repeated in order to obtain crossed probabilistic beliefs: $B_i^\alpha B_j^\beta p$. In semantics, the individual probability distributions of several players can be combined in order to construct entangled ones: $P_{ij}(w'';w) = P_i(w'';w').P_j(w';w)$.

For the model M, a probability distribution may be objective in the sense that it refers to known proportions (in a population of objects) or to observed frequencies (in a sequence of events): B_M. According to the 'Miller principle', it is usually assumed that the actor i accepts the objective probability when given by the modeler and transforms it into a subjective one: $B_i B_M^\alpha p = B_i^\alpha p$. However, the actor may directly adopt a subjective probability distribution when he gets no objective information. In any case, it is important to distinguish between an (objective) probability related to the physical environment and a (subjective) probability related to the belief about the environment. For instance, $B_i^\alpha(p \rightarrow q)$ signifies that actor i believes (with subjective probability α) that proposition p entails proposition q and $B_i(p \rightarrow_\alpha q)$ signifies that actor i believes that proposition p entails stochastically (with objective probability α) proposition q.

The probabilistic framework appears as a generalization of the set-theoretic framework in the following sense. It is always possible to associate a set-theoretic framework with a probabilistic one. In syntax, an actor believes a proposition if and only if its probability is equal to one; likewise, an actor accepts a proposition when its probability is strictly positive. In semantics, in each world, the accessibility domain is assimilated to the support of the corresponding probability distribution. However, with an infinite state of nature, such a reduction is more ambiguous, since an event of probability one may well not be certain or an event of probability zero not be impossible. Conversely, it is not possible to associate univocally a probabilistic framework with a set-theoretical one, since indiscernible worlds cannot be considered as equiprobable.

Within a probabilistic framework, it is harder to compare the accuracy of two belief structures possessed by the same player. In syntax, a system structure is more accurate than another if the probability an actor attributes to each proposition is higher. In semantics, the counterpart of such a property is not at all clear. However, a probability distribution is said to be less 'informative' than another by application of certain conventional criteria of different kinds. For instance, this happens when one probability distribution is a 'mixture' of another, i.e. is obtained by combining the second with another probability distribution. It is also the case when one probability distribution has lower 'entropy' than the other.

In the cards example, a first source of uncertainty is related to the distribution of cards. Consider, for instance, that there are only four cards in the pack, one of each sort. They are therefore equiprobable. If the soothsayer sees an even card, he knows with probability 1/2 that it is blue and with probability 1/2 that it is yellow. If he sees a yellow card, he knows with probability 1/2 that it is even and with probability 1/2 that it is odd. A second source of uncertainty comes from the perception of the cards. If the soothsayer is color-blind, he may see a yellow card to be yellow with probability 2/3 and red with probability 1/3 and he may see a red card to be red with probability 4/5 and yellow with probability 1/5.

1.5 Syntactic properties of beliefs

Coming back to a set-theoretic framework, several additional axioms may be imposed onto the individual belief structure of each actor (Rubinstein, 1994). They characterize his 'cognitive rationality', which is 'informational' when it concerns the empirical relevance of his beliefs and 'computational' when it concerns the formal validity of his reasoning. An axiom is 'external' when it concerns the relation of beliefs to the physical environment and 'internal' when it only concerns the actor's reasoning. These axioms are rather heroic and define a 'strong rationality', but they can be weakened in several ways. Moreover, they need not be independent; for instance, in the following, the second and fourth entail the third.

A first axiom is 'logical omniscience', which states that an actor believes all consequences of what he believes. Such an internal axiom means that the actor's belief structure is deductively closed in the same way as that of the modeler (he uses the same inference rules, in fact the modus ponens rule). This axiom assumes that the actor is able to implement a mental reasoning with very strong computational capacities, be it applied to true or false beliefs. According to ‘foundationalism’, the actor holds a set of ‘primary beliefs’ gathered from different external sources and infers ‘secondary beliefs’ by logical deduction. logical omniscience is weakened in different ways. For instance, one can consider that an actor has separate beliefs in different areas and only uses logical omniscience in each area.

A second axiom is ‘veridicity’, which states that what an actor believes is true. It is external, since it relates the actor’s beliefs to his environment and ensures that they are consistent with the modeler’s beliefs; they are therefore at least non-contradictory. This allows a distinction to be made between ‘knowledge’, considered always to be true, and ‘belief’, which may be false. In fact, this axiom may be applied to three kinds of propositions. Firstly, it may concern the physical environment, hence a basic proposition, in which case it is very exacting. Secondly, it may concern another actor‘s belief, hence a proposition B_jp, referring to the other’s mental states, about which errors are frequent. Thirdly, it may concern his own belief, hence a proposition B_ip, which is far less problematic because an actor is assumed to make no mistake about his own beliefs. This axiom can be weakened simply by assuming that the actor’s belief structure is logically consistent, excluding any ‘cognitive dissonance’.

A third axiom is ‘positive introspection’, which states that an actor believes what he believes. Such an internal axiom just reflects the self-specularity of the actor, which is a privilege of human beings. It is generally considered to be non-problematic (although it is called into question by ‘self-deception’), and it therefore has no reason to be weakened. According to another interpretation, however, positive introspection means that the actor is conscious of his beliefs. The axiom then appears more questionable, and a distinction must be made between ‘explicit beliefs’ the actor is able to express and ‘implicit (or tacit) beliefs’ he cannot enounce. To be more precise, a ‘consideration operator’ E_i is defined such that E_ip signifies ‘actor i considers explicitly proposition p’, before assessing it as acceptable or endorsed. It may be reduced to the fact that ‘actor i believes p or believes $\neg p$ or believes that he does not believe in p or $\neg p$’. In particular, primary beliefs are generally considered as explicit, while secondary ones are either explicit or implicit.

A fourth axiom is ‘negative introspection’, which states that an actor believes what he does not believe. It is internal and reflects the capacity of a player to know what he is unaware of. This is very problematic, since an actor may not even be aware of the propositions about himself considered by the modeler. It can at best be satisfied in a 'small world' where all possible propositions are exhaustively considered. Its weakening leads to the definition of an 'unawareness operator' A_i such that $A_i p$ signifies 'actor i is not aware of the proposition p'. This may be reduced to the fact that 'actor i does not believe p and does not believe that he does not believe p'. If unawareness implies that the actor's belief structure is not complete, the converse is not true, since an actor may ignore whether a proposition is true or not even when he is aware of it.

Within a probabilistic framework, the preceding axioms can be more or less easily extended. logical omniscience qualitatively expresses that the actor applies correctly the probability calculus. But it has no clear quantitative counterpart since an actor cannot deduce anything from the conjunction of $B_i^{\alpha}(p \rightarrow q)$ and $B_i^{\beta}(p)$. However, the 'weakest link assumption' asserts that any chain of reasoning has the strength of its weakest link and allows to state that $B_i^{\gamma}(q)$, where $\gamma = min(\alpha, \beta)$. The veridicity axiom just says that a true proposition has a positive probability, hence it is easily realized. But it does not mean that the actor knows the objective probability when it exists. As well tested, an actor frequently under-estimates events with high probability and over-estimates events with weak probability. The positive and negative introspection axioms coincide and say that if an actor knows a proposition with some probability, he is sure about that probability, and is therefore not uncertain about himself. This can be linked to the fact that an actor is often overconfident about his probabilistic judgment.

Let us return to our pack of cards to test these axioms. logical omniscience is rather easily satisfied: knowing that blue cards are even and red cards are odd, and that there are only three colors, the soothsayer may deduce that an odd card is blue or yellow (by contraposition and combination). Negative introspection is violated when the soothsayer sees a blue card as blue, a yellow card as yellow and a red card as yellow or red. When the selected card is actually red, he does not believe it is yellow and does not believe that he does not believe it is yellow. Positive introspection (together with negative introspection) is violated when the soothsayer sees a blue card as blue, a yellow card as blue or yellow and a red card as yellow or red. When the selected card is red, he believes it is not blue, but he does not believe that it is not blue. Veridicity (together with positive introspection and negative introspection) is violated when the actor sees a red card as yellow and conversely.

1.6 Semantic properties of beliefs

When considering the transcription rules from syntax to semantics, it appears that logical omniscience is automatically satisfied. For any material implication considered in some world, if the accessibility domain is contained in the field of the antecedent, it is included in the field of the consequent. Only weaker forms of accessibility relations between worlds can block the 'subjective' modus ponens. For instance, the semantics of Scott-Montague associates with each world not a single event, but a whole set of events (reflecting the propositions believed by the actor, which are not deductively closed). In particular, when beliefs are segmented into separate areas, the deductive reasoning applied in each area limits the consequences obtained and prevents the actor from being aware of eventual contradictions.

All other axioms have a semantic counterpart expressed as a property of the accessibility relation. Veridicity corresponds to the 'reflexivity' property of the accessibility relation. It states that each world is accessible from itself. Positive introspection corresponds to the 'transitivity' property of the accessibility relation. It states that if a second world is accessible from a first world and if a third world is accessible from the second world, the third world is directly accessible from the first one. Negative introspection corresponds to the 'Euclidianity' property of the accessibility relation. It states that if two worlds are simultaneously accessible from a given world, each of them is accessible from the other.

When all axioms are simultaneously satisfied, the accessibility domains of some actor (obtained when the actual world varies) form a 'partition' on the set of worlds (Aumann, 1999). A partition divides the world space into separate cells which cover the whole set. When the real world belongs to a certain cell, the actor just considers that the accessible worlds are those of that cell. Consequently, the uncertainty of the actor appears 'modular'. He cannot distinguish what happens inside a group of worlds, but he discriminates perfectly between the groups themselves. Such a structure makes it easier to compare the accuracy of different belief structures. One partition has to be 'finer' than another, in the sense that any of its cells is contained in a cell of the other one.

Obviously, with weaker axioms, weaker belief structures are obtained. For instance, when veridicity alone is not satisfied, we obtain a 'pseudo-partition', i.e. a partition over a subset of the world space. Some worlds are just ignored by the actor and the remaining ones are again treated in a modular way. Likewise, when negative introspection alone is not satisfied (but replaced by a weaker axiom), we obtain a 'nesting', i.e. a set of embedded coronas. The actor classifies the worlds in some order of plausibility and in each world, he cannot discriminate between the worlds which are more or equally plausible. This is exactly what happens in the game where an actor looks for an object and gets the answer 'cold, warm, hot' according to his distance from the target.

Now, when we consider probabilistic beliefs, deductive reasoning is replaced by probability calculus. logical omniscience in its weakest form is again automatically satisfied: in any world, if some event is included in another one, its probability is naturally smaller. Veridicity corresponds to the 'positivity' property: in each world, the probability attributed to itself is positive. Introspection (both positive and negative) corresponds to the 'uniformity' property: in each world, the same probability distribution applies over all worlds. Hence, when all axioms are simultaneously satisfied, there exists a unique probability distribution over the worlds with the whole world set as support.

In many applications, the belief structure of the actor is formed of two items. Firstly, a prior probability distribution common to all actors reflects some public (objective) knowledge about the physical environment. Consequently, the actors' beliefs are pre-coordinated by this prior probability distribution. Secondly, a partition (or more generally any accessibility relation) reflects the private (subjective) beliefs of the actor about the physical environment and the other actors' beliefs. Such a partition is specific to each actor and induces a degree of heterogeneity among them. By combination of the two sources, each actor defines, in each world, a subjective probability distribution on all worlds.

Consider the cards example, with only one card of each type. The soothsayer therefore knows that there is a uniform (objective) probability distribution on the worlds: (1/4, 1/4, 1/4, 1/4). In addition, he is endowed with a partition when he is color-blind, so that he does not distinguish between yellow and red cards. When the card is red, he believes that it is red with probability 1/3 and yellow with probability 2/3; he believes with probability 1 that he believes this. Likewise, he is endowed with a pseudo-partition when he believes that a red card is blue and a yellow or red card is yellow. If the card is red, he considers that it is yellow with probability 1/2. Finally, he is endowed with a nesting when he believes that a blue card is blue, a yellow card is blue or yellow and a red card is of any color. If the card is red, he believes with probability 1/4 that it is blue, with probability 1/2 that it is yellow and with probability 1/4 that it is red. But he believes that he believes that it is blue *with probability 7/16, that it is yellow with probability 6/16 and that it is red with probability 3/16.*

1.7 Homo-hierarchical beliefs

The usual Kripkean semantic framework already makes it possible to represent beliefs about beliefs. An alternative semantic framework (Fagin et alii, 1995) makes things more precise by segmenting worlds into successive layers: 0-worlds, 1-worlds, k-worlds, culminating in a unique ω-world. Each $(k+1)$-world is composed of k-worlds; the worlds can thus be put on the nodes of a tree (or rather a lattice). A general consistency condition asserts that k-beliefs and $(k+1)$-beliefs lead to the same beliefs about events of a lower order. Moreover, the relation between two levels can be set-theoretic (a $(k+1)$-world is a set of k-worlds) or probabilistic (a $(k+1)$-world is a probability distribution over k-worlds). Hence, one can construct pure set-theoretic hierarchies (sets on sets), pure probabilistic hierarchies (probability distributions on probability distributions) or mixed hierarchies (alternating set-theoretic and probabilistic relations).

Such a hierarchical structure is made even more precise by distinguishing between physical and psychical layers. The r first layers are physical: the 0-layer is composed of objects, the 1-layer of populations of objects, and so on. The relation between two physical layers expresses the (objective) relation which is believed between a population and its elements. The layers from r to ω are psychical: the $(r+1)$-layer is a belief on the r-metapopulation, the $(r+2)$ layer is a belief on the previous belief, and so on. The relation between two psychical layers expresses the (subjective) relation between successive beliefs, in fact an assessment of the lower beliefs. However, the psychical layers are reduced to one $(r+1=\omega)$ in accordance with the ‘no schizophrenia axiom’. Deduced from positive and negative introspection, it asserts that an actor cannot simultaneously hold alternative subjective beliefs about lower beliefs.

The simplest case is a three-layered structure formed of a belief about a population of objects $(r=1, \omega=2)$. The relation linking a 1-world to 0-worlds describes the objective belief about the composition of the population and expresses some ‘uncertainty’. The relation linking the unique 2-world to the 1-worlds deals with the subjective belief concerning the population and expresses some ‘ambiguity’. Especially, when the last is probabilistic, it expresses a degree of reliability of the basic belief. Four formal combinations are possible according to the structure of each layer. A ‘bi-set-theoretic belief’ is formed by a set-theoretic belief about set-theoretic beliefs and was already studied. A ‘bi-probabilistic belief’ is formed by a probabilistic belief about probabilistic beliefs and can be reduced to a unique probability distribution. A ‘family of distributions’ is defined as a set-theoretical belief acting on probabilistic beliefs. A ‘distribution of events’ is obtained as a probability distribution on set-theoretical beliefs (the basic events for which the probabilities are strictly positive are called the ‘focal events’).

In such a structure, any ‘material’ event can be weighted using two values obtained by a given procedure, which is applied layer by layer from the basic worlds up to the highest layer. The primal value is always higher than the dual one, and together they define an interval containing the assumed value of the event. Moreover, the primal value of some event is the complement to 1 of the dual value of the complementary event; hence, one value defined on all events is enough to define the other value. Finally, two belief hierarchies of different types are said to be ‘equivalent’ if they give the same values to each event. It can be shown that a distribution of events is always equivalent to a family of probability distributions, the opposite being true only under certain regularity conditions.

For a bi-probabilistic belief, the interval is always nil, since a probability distribution is ‘precisely’ defined. For a family of distributions, the dual values of an event correspond to its ‘upper probability’ and its ‘lower probability’. For a distribution of events, the dual values of an event correspond to its ‘credibility’ and to its ‘plausibility’ (the first being the so-called ‘Dempster-Shafer belief function’). In particular, when the focal events are separated, any event again has a unique probability. But when the focal events are nested, the dual values correspond to the ‘necessity’ and ‘possibility’ of the event. As far as the latter is concerned, the necessity of the conjunction of two events is the minimum of the necessity of each event, signifying that the weight of a chain of arguments is equal to the weight of its weakest link.

Moreover, a belief hierarchy (of an actor) can be said to be more accurate than another if, for any event, the primal (dual) value is higher (lower) in the first than in the second. This means that the value interval determined by the first structure is included in the interval determined by the second. For instance, one bi-probabilistic belief is never comparable to another one. A family of distributions is more accurate than another one if the first is contained in the second. A distribution of events is more accurate than another if the second is obtained from the first by translating weight from some focal event to another one including it. Contrary to usual probability distributions, the two last structures, called generalized probability distributions', are able to express total ignorance as well as perfect knowledge of an event.

Returning to our cards, let us first consider the 'Ellsberg pack' constituted of one blue (even-numbered) card and two yellow (even or odd) cards, the soothsayer drawing a card and seeing only the color side. His uncertainty can be formalized by a family of probabilities on the three first worlds: (1/3, 2/3, 0), (1/3, 1/3, 1/3), (1/3, 0, 2/3). This can be translated into an equivalent distribution of events: b(1/3), y(2/3). Now let us consider the usual pack, in which the soothsayer draws a card which is placed color-side up or number-side up with equiprobability. He sees the number precisely if the card is number-side up and is color-blind if the card is color-side up. The uncertainty is then represented by a distribution of events: {w1} (1/8), {w1, w2} (2/8), {w3, w4} (2/8), {w2, w3, w4} (3/8). He can infer the interval value of some interesting events: {w1} (1/8, 3/8), {w2} (0, 5/8), {w1, w2} (3/8, 6/8). Moreover, the distribution of events is now equivalent to the following family of probability distributions: (3/8, 0, 0, 5/8), (3/8, 0, 5/8, 0), (1/8, 5/8, 0, 2/8), (1/8, 5/8, 2/8, 0).

1.8 Hetero-hierarchical beliefs

The above hierarchical semantics also applies to crossed beliefs between players. Each $(k+1)$-world of an actor is formed of k-worlds of another actor, forbidding any fusion between layers. The relation between two layers can again be set-theoretic (set of an actor over the other's worlds) or probabilistic (probability distribution of an actor over the other's worlds). However, it is the pure set-theoretic (sets on sets) and the pure probabilistic (probabilities on probabilities) hierarchies that are particularly studied. The hierarchical structure may be self-contained in that it summarizes all crossed beliefs between agents. More precisely, one actor's knowledge of the hierarchical beliefs of the other does not give him more information in the probabilistic case, contrary to the set-theoretic case. However, such a hierarchical structure is equivalent to a Kripke structure, where each world contains information about the material environment as well as about the crossed beliefs of each actor at any level.

Inside a syntactic belief system, no specific axioms are defined as concerns crossed beliefs. For instance, it is perfectly possible each actor believes some fact and believes that the other believes the opposite: $B_i p$ and $B_i B_j \neg p$. In fact, individual belief structures can be more or less heterogeneous, due to different mental structures or experiences. However, it is possible to consider increasing degrees of correlation between actors' beliefs about a basic proposition p. Proposition p is a '*distributed belief*' if the actors are able to deduce p by associating their respective beliefs. It is a '*localized belief*' if one actor at least believes p. It is a '*shared belief*' if all actors believe p. It is a '*common belief*' if all actors believe p, believe that the others believe p, and so on ad infinitum.

The crossed belief operators we have just introduced satisfy the usual axioms, particularly veridicity and positive introspection, as soon as all the individual operators do so. Thus, when veridicity is satisfied for all actors, a shared belief or common belief is called shared knowledge or common knowledge. Satisfying these rationality properties, crossed beliefs may be conventionally attributed to some collective entity. However, they do not really reflect a 'collective belief' because they are obtained exclusively through the association of individual operators. Nevertheless, each actor may consider that some collective belief, denoted $B_G p$, exists, hence $B_i B_G p$. More subtly, each actor may believe that p and at the same time believe that the group believes $\neg p$: $B_i p$ and $B_i B_G \neg p$.

In Kripke semantics, in a set-theoretic framework, the crossed operators may be represented by specific accessibility relations (not attributed to an actor). For instance, if all the actors are endowed with belief partitions, the common belief operator is represented by the finest partition among the partitions coarser than those of the actors (the 'meeting' of the partitions). It follows that, if one system structure is collectively more accurate than another, more propositions become shared beliefs or common beliefs. In other respects, the crossed operators can be extended to probabilistic beliefs. For instance, 'α-common belief' is obtained with 'beliefs with probability α' for each actor at each level.

More profoundly, the fundamental concept of common belief, introduced by the philosopher D. Lewis (Lewis, 1969), may be the object of three alternative definitions. The 'hierarchical' definition states that a proposition p is common belief if it is a shared belief that it is a shared belief... that p; it is based on an infinite conjunction of propositions. The 'circular' definition says that a proposition p is a common belief if it is a shared belief that p is true and that p is a common belief; it remains on a fixed point of the shared belief operator. The 'operational' definition says that a proposition p is a common belief when it had been exemplified in the presence of all actors; it is based on a public announcement, hence on a practical protocol. As concerns the two formal definitions, the second implies the first, but the converse is not true.

Common knowledge of a given phenomenon in a group is not unattainable, since it is sufficient that each actor observes the phenomenon in the presence of the others. If he is aware of the mutual observation, he is able to proceed inductively because the reasoning step is always the same from one level to the next. However, taking into account the limited computational capacities of the actors, certain weakenings of common belief have been proposed. The concept of 'k-shared belief' (shared belief up to level k) expresses the idea that actors are limited in their crossed reasoning. Many writers such as M. Proust (Proust, 1954) or psychoanalysts such as J. Laing (Laing, 1970) study such a crossed reasoning and stop in fact at level 4. The concept of 'α-common belief' expresses the idea that the players are never sure about each other's beliefs. Many authors have pointed out that the degree of confidence in crossed beliefs decreases from level to level.

In the cards example, consider that one soothsayer observes the color side and a second one the number side of a selected card. The first soothsayer is then endowed with the accessibility partition $\{w1, w2\}; \{w3, w4\}$ *while the second has the accessibility partition* $\{w1\}; \{w2, w3\}; \{w4\}$. *The soothsayers have a distributed belief about each possible card: by pooling their information, they can know the card, whatever it is. They have a localized belief, whether the card is blue or red, since the first soothsayer knows this. Some events are shared beliefs between the soothsayers at level 2: when the card is red, each believes that the other believes it is not blue. No event (except the whole pack of cards) is a common belief (the 'meeting' of the accessibility partitions is the complete set of worlds). However, a common belief may be obtained by an operational procedure: a referee shows both sides of the card to each soothsayer in the presence of the other.*

2

Change of individual beliefs

People will come to believe what they want to believe only to the extent that reason permits.
Z. Kunda

Initial beliefs are transformed into final beliefs each time an actor receives and duly codifies new pieces of information provided by environmental sources. belief change ensures both the internal coherence between beliefs and message and the external appropriateness of beliefs to the represented system, since the message is generally considered as true. A fruitful analogy can be drawn between the actor's rules of belief change and the epistemological rules followed by the modeler when he revises his theoretical models to fit new observations. These change rules capture the dynamic aspects of the actor's cognitive rationality, especially the economy principle which states that he changes his beliefs as little as possible.

Belief change is considered in two different contexts, corresponding to messages about either a fixed or an evolving environment (2.1), and is naturally expressed by syntactic axioms (2.2). For both contexts, explicit rules of semantic belief change can be derived from the axioms, and expressed in either a set-theoretic framework (2.3) or a probabilistic one (2.4). The belief change process can further be iterated for successive messages (2.5) and can also be applied to homo-hierarchical or hetero-hierarchical beliefs (2.6). Furthermore, several reasoning operations performed by the actor, such as nonmonotonic, conditional or abductive reasoning (2.7), can be shown to be isomorphic to belief change (2.8).

2.1 Contexts of belief change

Belief change is always analyzed in terms of the same scheme: an actor is endowed with an 'initial belief'; he receives a 'message' from outside; he combines both in order to coin a 'final belief'. The initial belief is considered as already given, implicitly obtained by former reasoning operations. The message about his environment or about himself is obtained by direct observation or through the testimony of others. The final belief is of the same nature than the initial belief. All three elements are considered as beliefs related to the same system, and they are interpreted univocally. However, if the initial belief is generally a structural belief, the message is more a factual one, and the final belief is again structural. Likewise, if the initial belief may be explicit or implicit, the message is generally codified (as a 'signal') and the final belief is of any kind.

Formally, although the message may be consistent with the initial belief, it may also contradict it, hence creating a 'cognitive dissonance'. Two main 'contexts' of belief change are generally considered, depending on the interpretation given to the message. The context of '*revising*' assumes that the actor's environment is fixed and that the message gives further information about such a fixed system. The context of '*updating*' assumes that the environment is evolving and that the message gives some information about the system's direction of evolution. In the first case, a non-contradictory message makes the initial belief more precise (specification message) while a contradictory message helps to modify the possibly wrong initial belief (rectification message). In the second case, the initial belief and the message can always be interpreted as actually non-contradictory, since they correspond respectively to the initial and ongoing states of the system.

Several attempts have been made to formally reduce one context to the other by simply modifying the description of the system. The aim is to transform the axioms relevant to one context into the axioms relevant to the other one. One of these consists in considering a path of the system through time as an inter-temporal world. In that case, updating appears as revising, since it gives information about one step in the path. Another consists in considering a message as always giving information about some change in the system. In that case, revising appears as a limit case of updating when the change is seen as zero. In either case, it is necessary to give more structure to the propositions or to the possible worlds by explicitly modeling the evolution of the system, for instance by an appeal to temporal or dynamic logic.

A third context called '*focusing*' is involved when the system is a population of objects and the message gives some information about an object drawn from the population. In fact, such a context deals with two distinct (physical) levels. The initial belief concerns the population of objects with its statistical properties while the message (and the final belief) concerns the object drawn. As a consequence, the initial belief is usually not affected by the message, unless the two are contradictory. However, the focusing context can be reduced to a revising one, thanks to the conjunction of two formal operations. The (physical) operation of 'projection' states that the drawn object is picked according to the same probability distribution as the distribution of the objects in the population. The (psychical) operation of 'revising' is applied to the projected belief after reception of an original message concerning the selected object.

An actor's revising process can naturally be compared with the learning process of a scientist. The scientist possesses an initial theory which is formally expressed as a set of assumptions and which admits some testable consequences. He receives a message in the form of results from experimentation or field observations, generally new observations. The theory is refuted when the observations contradict the consequences of the theory. The scientist must then determine which assumptions he should modify in order to restore the coherence between theory and observations. The Duhem-Quine problem states precisely that this operation cannot be carried out univocally on a pure logical basis; it requires the introduction of further principles.

In any case, belief change is again assumed to be a purely epistemic operation. The evolution of the belief is not a result of the actor's decision, but of a specific reasoning mode. Like other reasoning modes, it is a calculus performed on beliefs as stated by the 'computational theory of mind'. It is governed by general rules that the actor cannot manipulate. In particular, he cannot believe that a proposition is true and decide to believe that it is not true (the Moore paradox), except conventionally, in special forms of reasoning such as *reductio ad absurdum* or counterfactual reasoning (see 2.7). Moreover, belief change generally assumes that the actor is endowed with strong cognitive rationality. He is able to apply the rules of belief change without confronting strict cognitive constraints. Such constraints may, however, be considered, an actor revising for instance the beliefs held in a reduced field without taking into account the impact outside the field.

Firstly, consider that the soothsayer thinks he has put a pack of cards in his pocket and that, at some moment, he observes that his pocket is *actually empty. He may interpret this fact in a revising way, by assuming that he forgot in fact to put the pack in his pocket. He may interpret it in an updating way, by assuming that the pack has been lost or stolen. Now, consider the pack with its usual but uncertain content, and assume that the soothsayer receives a certain message about it. A revising message says that there is no red card in the pack; it gives some information about the pack by contradicting the initial belief, which assumed a red card. An updating message says that there is no longer a red card in the pack; it states that a physical operation has been carried out on the pack, thus maintaining the truth of both the initial and the final beliefs. A focusing message says that a card drawn from the pack is not red; it gives information about a selected card which does not contradict the initial belief (but may do so if the message says the card is green).*

2.2 Syntactic change

In syntax, belief change is first studied in a restricted framework where a unique actor i faces a physical environment. The whole transformation takes place in the same language defined by basic propositions and logical connectors. From an initial belief operator B_i and a message m, the actor infers a final belief operator B_i^*, where $B_i^*\phi$ means that 'actor i believes ϕ after receiving message m'. In order to be more succinct, but without restriction, the modeler considers the belief structure of the actor directly. The initial belief set K_i is formed of a subset of propositions about the physical environment. The m is a unique proposition about the physical environment, and is considered to be true. The final belief K_i^* is again a set of propositions.

Belief change axioms give necessary conditions that relate the initial belief to the final belief. They make no difference between structural and factual beliefs. Most axioms rely on two qualitative principles assumed to be adopted by any actor. The 'priority principle' states that the message is more reliable (for revising) or more recent (for updating) than the initial belief and is treated as more relevant. The 'economy principle' states that the actor modifies his belief as little as possible. More precisely, the 'weak economy principle' asserts that, if some part of the initial belief can be preserved because it is consistent with the message, it is kept. The 'strong economy principle' asserts that, if some part of the initial belief must be modified because it contradicts the message, it is changed minimally.

A first set of axioms is common to both the revising and the updating contexts. According to the priority principle, the ‘success axiom’ states that the final belief always validates the message, in the sense that it entails it. It assumes that the message is always true while the initial belief may be false, and that their combination is therefore asymmetrical. According to the economy principle, the ‘conservation axiom’ states that if the initial belief validates the message, it is conserved. Likewise, the ‘inclusion axiom’ (generalized, for two messages, into the ‘sub-expansion axiom’) asserts that the final belief keeps at least the part of the initial belief which is consistent with the message: a piece of belief which is not refuted is conserved.

Two specific axioms are then added in a revising context, stressing the independence of messages (Alchourron, Gärdenfors). The ‘preservation axiom’ (generalized, for two messages, into the ‘super-expansion axiom’) states that the final belief keeps at most the part of the initial belief which is consistent with the message: no new piece of belief can enter the picture outside of the initial belief and the message. The ‘right distributivity axiom’ states that the final belief resulting from the revision of an initial belief by the union of two messages is the union of the corresponding final beliefs: if several messages become available to the actor simultaneously, they can be treated separately.

Two specific axioms are similarly added in an updating context which stress the ‘modularity’ of initial beliefs (Katzuno, Mendelzon). The ‘preservation axiom’ (as well as its extension, the super-expansion axiom) is still required, but only when the initial belief is ‘complete’ (i.e. each proposition is believed or not by the actor without ambiguity). The ‘left distributivity axiom’ asserts that the union of two initial beliefs leads to the union of the corresponding final beliefs when updated by the same message: if the initial belief can be decomposed into separate beliefs, these beliefs can be updated separately.

The axioms specific to the revising context and those specific to the updating context are incompatible (except when the initial belief is complete). More precisely, the preservation axiom is violated by the updating axiom system and the inclusion axiom by the revising axiom system. In other respects, the change axioms which are common to both the revising context and the updating context may be modified conjointly. For instance, the success axiom can be abandoned if the initial belief and the message are considered to be equally reliable, leading to their symmetrical combination (fusion of beliefs). Likewise, the sub-expansion axiom can be refined to require that the message only modifies the part of the initial belief which is specifically concerned by it.

In the cards example, the axioms can easily be interpreted. According to the success axiom, if the message says that there are no red cards in the pack, in contradiction to the initial belief, the message is considered to be true. According to the conservation axiom, if the message says that there are even cards in the pack as already known, the initial belief is unchanged. According to the inclusion axiom, if the message says that there are no red cards in the pack, the soothsayer does not change his mind about the yellow or blue cards. According to the preservation axiom, if the message says that there are no red cards in the pack, the soothsayer does not suddenly believe that there may be green cards in the pack.

2.3 Semantic change

In semantics, belief change is geometrically expressed with reference to a set of possible worlds. The initial belief is represented by some event K_i, corresponding to the accessibility domain of actor i associated with the (implicit) actual world; this actual world may be outside K_i since the initial belief may be false. The message is expressed without ambiguity by another event M; since the message is considered to be true, the actual world belongs to it. The final belief is again an event denoted K_i^*, considered to be univocally defined. The initial belief and the message are non-contradictory if and only if they admit a non-empty intersection, i.e. some worlds belonging to both sets.

The syntactic axioms on propositions can be translated into semantic properties on events. For instance, the success axiom simply says that the final belief is included in the field of the message. The rules of belief change then specify precisely how to select the ‘new’ worlds (in K_i^*) from the set of ‘old’ ones (in K_i) for some message M. In each context of belief change, a representation theorem links the change axiom system to a family of change rules. The change rules are always characterized by a specific order on the possible worlds that the actor is assumed to hold, expressing some kind of ‘plausibility’ he attributes subjectively to them.

In a revising context, a representation theorem introduces a ‘global preorder’ of the worlds, relating to the initial belief. It is expressed by a set of coronas around the initial belief (which constitutes the first corona), a world being more distant from the initial belief if it is in a more outlying corona. The final belief is nothing other than the intersection of the message with the first corona to intersect it. Thus, the old worlds are replaced by the new worlds which are nearest to them and at the same time contained in the message. When the message does not contradict the initial belief, the final belief is nothing other than their intersection, hence is more accurate than the initial belief, as the actor gains some information. When the initial belief and the message contradict each other, the final belief cannot be compared to the initial one, but the actor has nevertheless gained something in the change.

Transcribed into syntax, the preceding order on worlds induces an ‘epistemic entrenchment’ order on propositions. One proposition is more entrenched than another if the actor, on receiving the message that one of them at most is true, keeps the first rather than the second. The change rule then operates in the following way. The message is added to the initial belief and completed by their common logical consequences. In the case of contradiction in the completed set of propositions, the less entrenched propositions are removed one by one until the consistency of the belief system is restored. The epistemic entrenchment order is again specific to the actor’s prior view of the world; it may be revealed by the modeler observing several belief changes on the part of the actor.

In the case of the scientist considering a theory, he can likewise be endowed with a prior entrenchment order on the assumptions forming the theory under consideration. Such an order is in principle specific to each scientist, but there may be a consensus among scientists about it. When the theory is refuted by some experience (considered as non-controversial), the scientist removes the less entrenched assumptions until the remaining ones are once again consistent with the observations. In fact, he may replace the abandoned assumptions by others which are consistent with the experience. The most entrenched assumptions, only abandoned as a last resort, form the ‘hard core’ of the scientist (logical and mathematical principles, well- tested and accepted assumptions).

In an updating context, a representation theorem introduces a ‘local preorder’ over the worlds, relative to each world (especially those in the initial belief). Such an order is expressed by a set of coronas around a given world (which constitutes itself the first corona). The final belief is nothing other than the union, for all worlds in the initial belief, of the intersections of the message with the first corona to intersect it. Hence, each old world is replaced by the new worlds which are the nearest to it while contained in the message. When the message intersects the initial belief, the worlds in common are kept while the worlds in the initial belief and not in the message are transferred to the nearest worlds in the (Lewis change rule). In general, the final belief cannot be compared with the initial one in terms of accuracy.

In the cards example, consider a pack which contains at most two cards: it may or may not contain a blue card and it may or may not contain a red card. The initial belief indicates that there is at least one card in the pack. The message indicates that there is no red card in the pack. The belief change depends on the interpretation of the message. In a revising context, the message says that in fact the pack never contained a red card. The final belief then asserts obviously that there is a blue card in the pack. In an updating context, the message says that the red card, if initially present in the pack, has been removed. The final belief then asserts either that the pack only contains a blue card (if it initially contained two cards or only a blue card) or that it contains no card at all (if it initially contained only a red card).

2.4 Probabilistic change

Syntactic belief change can be transcribed to a situation where the initial belief is probabilistic (*prior* probability distribution), the message keeps set-theoretic and the final belief is probabilistic too (*posterior* probability distribution). The transcription of the change axioms can be more or less demanding. The transcription is ‘weak’ when the set-theoretic axioms are simply transposed (without added value) in terms of the support of the probability distribution. The transcription is ‘strong’ when original axioms are introduced, although in the spirit of the set-theoretic axioms, which take into account inequalities between probabilities. The transcription is ‘very strong’ when axioms deal directly with equalities between probabilities.

For instance, the ‘probabilistic weak success axiom’ states that the posterior probability of the message is equal to one. The ‘probabilistic strong inclusion axiom’ states that the posterior probability of an event is greater than the prior probability of the intersection of that event with the message. The ‘probabilistic very strong right-distributivity axiom’ (or ‘linear mixing axiom’) states that the posterior probability when receiving a union of two messages is a linear combination of the posterior probabilities obtained for each message. The ‘probabilistic very strong left distributivity axiom’ (or ‘homomorphism axiom’) states that the posterior probability when considering a weighted sum of prior probabilities is the weighted sum of the corresponding posterior probabilities. The axiom systems which correspond respectively to a revising and an updating context are again contradictory. In particular, the homomorphism axiom is violated by the revising axiom system.

Both axiom systems can again be transposed into probabilistic belief change rules through representation theorems. The link of new worlds to old ones is broken down into two parts. A 'selection rule' specifies the worlds which are selected after reception of the message. It coincides with the change rules introduced for set-theoretic beliefs. An 'allocation rule' shows how the weights of the old worlds are transferred to the new worlds. It characterizes the way in which the axioms have been more or less strengthened. For instance, the strong revising axiom system gives rise to 'conditioning rules' (including the Bayes rule) and the strong updating axiom system gives rise to 'imaging rules' (including the Lewis rule). Consequently, the Bayes rule must violate the homomorphism axiom, a contradiction well exemplified by the 'Simpson paradox' (illustrated in the example of 2.8).

A 'generalized Bayes rule' (already axiomatized by Popper-Miller) is univocally obtained by applying the super-strong linear mixing axiom in a revising context. The usual Bayes rule states that the posterior probability of some event is equal to the prior probability of the intersection of that event with the message divided by the probability of the message. Hence, its selection rule states that the new worlds are the old worlds of the support which are consistent with the . Its allocation rule states that the weight of an old world excluded by the message is transferred to the new worlds proportionally to their own prior weight. Its main weakness is that it is not defined when the message contradicts the prior belief. Conversely, the generalized Bayes rule applies even in the case of a 'surprising' message, i.e. a message of null prior probability.

It follows that, when founded on purely epistemic principles, the Bayes rule is relevant only in a revising context. It appears natural for objective probabilities defined on a population of objects. Here, it simply reflects how the proportions of objects are modified when considering a subpopulation of objects. But it is far less natural when considering subjective probabilities on specific events. Since it combines in a specific numerical way a prior probability and a message, the Bayes rule can easily be refuted and has indeed been refuted in many laboratory experiments (Kahneman, Slovic). For instance, with reference to the Bayes rule, 'conservative' actors attach more importance to the prior belief by dismissing the message, while 'reformist' actors attach more importance to the message by dismissing the prior belief. As a limit case, some actors may ignore a contradicting the initial belief (confirmationist bias) or may ignore the prior in favor of the perceived as unavoidable since it happened (retrospective bias).

Consider now an extended situation in which the initial belief is a probability distribution and the (revising) message is itself probabilistic. More precisely, the actor knows a conditional probability of the message (given the actual world). Such a probability results from the fact that the message is really fuzzy or is not perfectly perceived. It is objective when the actor receives a signal which is correlated to the actual state of the environment through a known law. It is subjective when the actor gives some reliability to the sources of the message. In semantics, different change rules were proposed for that situation without axiomatic foundation. The 'Jeffrey rule' combines the initial belief and the message in an asymmetric way, the message being considered as true in its probabilistic form. The 'Dempster rule' combines the initial belief and the in a symmetric way, the two sources of information being considered equally reliable.

Consider again the pack constituted of at most two different (red or blue) cards. The initial belief is that there is at least one card in the pack, more precisely that it contains two cards with probability 1/2, a blue card only with probability 1/3 and a red card only with probability 1/6. The message says that there is at most one card in the pack. In a revising context, this means that there was at most one card in the pack, and in an updating context, that there remains at most one card in the pack. The final belief is the same in both contexts and asserts that the posterior probability is 2/3 for a blue card and 1/3 for a red card. Now consider the more general pack of cards with the following prior distribution: 5/6 of blue cards and 1/6 of red cards. A card is selected at random and the soothsayer sees it with a 4/5 probability of being right and with a 1/5 probability of being wrong. The posterior probability for the card to be red is then 4/9, less than 1/ 2.

2.5 Iterated change

The belief change process does not specify where the initial belief comes from. It is interested in the transformation of beliefs rather than in their formation. Certainly, belief formation can be reduced to belief change by considering a completely uninformative belief as the initial belief, progressively revised by successive messages. Conversely, belief formation may be based on a specific process, founded on inference operations such as induction from all available information (see 2.7). More profoundly, belief change is based on a 'fundamentalist' point of view: the messages form a privileged belief base from which the other beliefs (except eventually the initial one) are deduced. But belief change also includes a ‘coherentist’ point of view: the final belief has to be logically consistent.

The actor may receive successive messages about the same environment from diverse sources, more or less independent. The sources are considered of equal weight, even if it is possible to take into account their relative reliability. The actor has two procedures for the sequential processing of these possibly contradictory messages. On one hand, he can change his belief with each new message. Hence, he has neither to remember past messages nor to group successive messages together into one unique message. On the other hand, he can wait until he has enough messages to perform a global belief change. An equivalence condition can be imposed, stating that the present belief has to be the same at each time for both procedures. An open problem is whether the final beliefs tend, for some regular sequence of (true) messages, towards a limit more or less independent from the initial belief.

In syntax, iterated change may satisfy two basic axioms. The ‘commutativity axiom’ states that for two successive messages, the final belief is independent of the order in which the messages were received. This axiom is natural in the revising context, since all messages inform about the same world, but not in the updating context, where the last received message is pre-eminent because it concerns the most recent situation. It results from the conjunction of the sub-expansion and super-expansion axioms, but may be psychologically irrelevant, for instance when exchanging good and bad news. The ‘idempotence axiom’ states that, for two successive identical messages, the final belief is the same as the intermediate belief. This axiom is again natural for revising, but less so for updating, since it assumes some stability of the environment. It results from the conjunction of the success and conservation axioms, but may be psychologically irrelevant, for instance when a second message confirms the first.

Restricting to a revising context, iterated change is submitted to certain additional axioms whatever the messages. For instance, the ‘filtering axiom’ states that if two messages arrive successively, the second being more restrictive than the first, the second alone gives the same final belief as both. The ‘predominance axiom’ states that if two contradictory messages arrive successively, the second alone gives the same final belief as both. These axioms are more controversial than those proposed in the case of a single revision, since they make assumptions about the memory the actor retains about the sequence of messages.

As experimentally observed, when observing several occurrences of a random event, an actor may think that it has now less chance to happen again (gambler's fallacy) or conversely, that it has more chance to happen again (hot hand fallacy). In fact, several axiom systems are competing without a consensus about them.

In semantics, the change rules are then generalized from the case of a single message to the case of successive messages. A representation theorem states how the order on worlds itself evolves for each new message. Generally, the relative order on the subset of worlds which are consistent with the message is preserved and they form the first new coronas. Likewise, the relative order of the subset of worlds which are refuted by the message is also preserved (or gathered in a single category) and they form the last new coronas. However, the order of the confirmed worlds and the order of the refuted worlds can be entangled in several ways. With such simple rules, when receiving an infinite sequence of regular messages, the asymptotic order of worlds may be directly computed.

In a probabilistic framework, the doctrine called 'epistemic Bayesianism' assumes both that the prior beliefs are probabilistic and that they are revised according to the Bayes rule. The prior probability distribution is exogenously given and is often a uniform distribution, without precise justification. The Bayes rule satisfies both the commutativity and idempotence axioms. It leads to continuous changes of the probability distributions in case of non contradicting messages. For specific sequences of messages, the posterior probability distribution tends toward an asymptotic one, according to the central limit theorems. When extended to a probabilistic message, the change rules are more contrasted. The Jeffrey rule is idempotent and non-commutative since the last message imposes its change; the Dempster rule is commutative and not idempotent, since each message reinforces the preceding one.

Consider three cards showing on one side a blue, yellow and red color respectively (example adapted from the Monty Hall problem). On the other side, only one card is even and the soothsayer has to guess which it is. When seeing only the color, he chooses one card at random since he attributes an equal probability of 1/3 to each. However, a referee, knowing the winning card, turns at random one of the two cards not chosen in order to show that it is not even. He then asks the soothsayer if he maintains his initial guess. Curiously, it is always in the soothsayer's interest to change his mind, since the posterior probability of his first chosen card remains 1/3 while the posterior probability of the other card is now 2/3. Contrary to intuition, the message given *by the turned card is, on average, informative. If his chosen card is the even card, the referee may turn either of the other cards and the message effectively has no informative value; but if his chosen card is not the even one, the referee is constrained to turn the remaining odd card, and the message reflects that constraint.*

2.6 Change of hierarchical structures

Consider first that the initial belief is a homo-hierarchical belief structure and that the message is a (true) event. The most interesting case is again the case of a two-layer initial belief where the first layer consists in a population of objects and the second in uncertainty about the composition of the population. It applies to the case where belief change concerns both the objective belief of an event and the degree of reliability attached to it. The contexts of belief change are subject to a clear interpretation. revising arises when the message gives some further information about the population of objects. updating arises when the message gives some information about the transformation of the population of objects. Focusing arises when the message gives some information about one object drawn from the population.

In semantics, the change rules apply to the two dual values previously defined for each event. They are obtained by adapting the rules obtained for one-layer initial beliefs in different contexts. For a bi-probabilistic belief, revising follows the 'minimal rule' (Bayesian revision of the upper level probabilities according to the message), updating follows the 'maximal rule' (Bayesian revision of the lower level probabilities according to the message) and focusing follows the 'synthetic rule' (Bayesian revision of the synthetic probability distribution). For a distribution of events, revising follows the 'geometric rule' (consideration of the only focal sets included in the message), updating follows the 'Dempster rule' (consideration of the intersections of the focal sets with the message) and focusing follows the 'Fagin-Halpern rule' (Bayesian revision of all distributions consistent with the initial belief).

Consider now that the initial belief is a hetero-hierarchical belief structure and that the is again a (true) event. In a revising context, it becomes necessary to distinguish between two dimensions of any message, its content and its status. The content of the message indicates the information given to an actor. More precisely, the message may be material (about the physical environment) or epistemic (about another's belief at any level). The status of the message concerns the diffusion of the message among the agents. For instance, a message may be 'secret' (one agent receives it and the other does not know he receives it), 'private' (one agent receives it, the other knows that he receives it without knowing its content and this is common knowledge) or 'public' (each actor receives the message and this is common knowledge). But a message may adopt a great variety of statuses (quasi-secret message, private message believed to be public, etc).

In syntax, even if one keeps to a 'specification message', i.e. a message which does not 'surprise' the actor, the usual axioms are no longer automatically satisfied. In particular, the success axiom is no longer relevant since, when an actor learns the other's initial belief and the message he receives, he knows that he will change his beliefs. For instance, if a message is common belief and asserts simultaneously that p and $B_i \neg p$, the final belief must assert that $B_i p$. An 'inference axiom' is however satisfied, corresponding to a weak form of 'modus ponens' applied by the actor. It asserts that, if an actor learns a message, a proposition is believed in the final belief if and only if it is entailed by the message in the initial belief. Especially, if a proposition is believed in the initial belief and does not contradict the message, it is kept.

In semantics, the final belief structure resulting from a specification message is obtained in a rather simple way. It is a combination of the initial belief structure and of an auxiliary belief structure formally expressing the status of the message (for a given content). Each possible world of the final structure is a pair formed by a possible world of the initial structure and a possible world of the auxiliary structure. The accessibility domains of the final structure are obtained by considering the (non-void) intersections of the accessibility domains of the initial structure and the auxiliary structure. The state associated with each new world is the same as the state associated with the old one. Note that, even when the message is a specification message, the final belief of an actor may be wrong, especially when the other has received a secret message.

A fundamental property of multi-agent belief revision is that the actors have more accurate beliefs after receiving the specification message than before. More precisely, for a given initial belief, if one message is more accurate than another (collectively, individually or intimately), the final belief obtained with the first is more accurate (in the same sense) than the final belief obtained with the second. For instance, since a public message is collectively more accurate than a null message, it leads to a final belief which is collectively more accurate than the initial one. Likewise, a private message leads to a final belief which is individually more accurate than the initial belief. The actor's beliefs are progressively refined as long as the messages do not contradict them.

In the cards example, consider the 'Ellsberg pack' formed of a blue card and two yellow cards (each either even or odd). The initial distribution on (blue, yellow even, yellow odd) is (1/3,1/3,1/3). A revising message tells a soothsayer that there is no odd card in the pack; it leads to a family of distributions reduced to a single distribution: (1/3, 2/3, 0). A focusing message says that one card was drawn from the pack and that it is not odd; it leads to the family of probabilities: (1/3, 2/3, 0), (1/2, 1/2, 0), (1, 0, 0), where blue is evaluated between 1/3 and 1, yellow and even between 0 and 2/3. Now consider the usual pack where one soothsayer sees the color and another sees the number. A public message tells both that the card is not red, hence transforming the beliefs. A private message tells the second that the card is yellow. The second then knows the card, while the first knows that the second knows. A secret message tells the first that the card is even. The first then knows the card while the second keeps his (false) belief.

2.7 Reasoning operations

The basic reasoning operation for the modeler is deduction and it is quite naturally attributed to the actor too. Likewise, the explanation mechanism followed by the modeler is attributed to the actor too. In syntax, explanation is represented, at least for its formal part, by the deductive-nomological model (Hempel-Oppenheim). In order to be explained, a factual phenomenon k (conclusion) is deduced from a structural law H (major premise) and a factual condition h (minor premise). Hence, explanation associates two reasoning operations, deduction when linking assumptions and conclusions, particularization when associating structural and factual propositions. In semantics, it is not easy to find a counterpart to explanation since, if deduction has a clear counterpart (inclusion of events), this is not the case for particularization.

This basic scheme can be used by the scientist or the actor in three different modes, according to what is already known to him (Peirce, 1978). The 'projection' mode consists in deducing the phenomenon k from law H and condition h. The 'attribution' (or 'abduction') mode consists in revealing the condition h from phenomenon k and law H. The 'induction' mode consists in inferring the law H from pairs of condition h and phenomenon k. If projection is simply deduction, applied to past conditions (postdiction) or future conditions (prediction), attribution and induction are more complex reasoning operations which are not precisely and unambiguously defined by the above scheme.

In fact, the modeler or the actor implements mentally various reasoning operations in order to structure his beliefs. There is no recognized taxonomy of these reasoning operations, even if a short list exists which is neither exhaustive nor exclusive. These operations are studied by modal logic, directly for the scientist, by removing the belief operator for the actor. They are generally expressed as contextual inferences of the form ‘from antecedent A, infer consequent C in some context K’. In syntax, they are defined by an axiom system acting on propositions. In semantics, they are defined by rules acting on possible worlds. These operations are generally not logically valid (but pragmatically acceptable) because they are ampliative (they infer propositions by generalizing from the true ones). Three of them are especially interesting, and they are now examined in greater detail.

Firstly, ‘nonmonotonic reasoning’ is a weakened form of deduction which states that ‘fact A normally entails fact C’, since it admits some exceptions which are given by an incomplete list of provisos. Verified by classical deduction, the ‘monotony axiom’ states that if ‘A entails C’, then ‘A and B entails C’. It is violated by nonmonotonic deduction, which provides a conclusion which is defeasible when new information arises. For instance, from ‘Birds fly’ and ‘Carlo is a bird’, one normally infers that ‘Carlo flies’, unless Carlo is a penguin (or an ostrich). Among the axioms concerned, the ‘reflexivity axiom’ states that ‘A normally entails A’. The ‘cut axiom’ states that if ‘A normally entails B’ and if ‘A and B normally entail C’, then ‘A normally entails C’. The ‘cautious monotony axiom’ states that if ‘A normally entails B’ and if ‘A normally entails C’, then ‘A and B normally entail C’.

Secondly, ‘abduction’ is a weakened form of reverse deduction (the latter being called ‘classical abduction’) which reads ‘from facts A, abduce assumptions C’ capable of explaining them. It is especially used in diagnosis analysis, either for medical diagnosis (abducing an illness from symptoms) or criminal investigation (abducing a criminal from clues). For instance, according to the fact that ‘this observed object flies’, one may abduce that ‘it is a bird’. Likely, from the facts that ‘Arthur is a black falcon’ and ‘Oscar is a black eagle’, one may abduce that ‘birds of prey are black’. Among the axioms sustaining the main form of abduction, i.e. ordered abduction, the ‘reflexivity axiom’ states that ‘A abduces A’ and the ‘transitivity axiom’ asserts that if ‘A abduces B’ and ‘B abduces C’, then ‘A abduces C’.

Thirdly, 'conditional reasoning' is a restricted form of 'material implication' which enounces that 'if fact A happens, then fact C is realized'. Contrary to it, it is not automatically considered as true whenever its antecedent is false. Either retrospective or prospective, it is called 'profactual' when its antecedent is true, 'counterfactual' when its antecedent is false and 'afactual' when its antecedent has no truth value (especially when it concerns the future). For instance, the conditional 'if Arthur is a falcon, he has a hooked beak' is profactual while 'if Arthur were a cock, he would sing' is counterfactual. A further distinction is made between two forms of counterfactual propositions: the indicative ('if Gutenberg did not invent printing, somebody else invented it') and the subjunctive ('if Gutenberg had not invented printing, somebody else would have invented it'). Among the axioms satisfied by conditionals are the 'reflexivity axiom' and the 'cautious monotony axiom'.

For instance, consider a pack of cards with a lot of yellow even cards, a few yellow and odd cards, and a few blue and even cards. When seeing a yellow card, the soothsayer may nonmonotonically deduce that it is even (yellow and even cards are exceptions). When just seeing an even card, he may abduce that it is yellow (even and odd are observed signals, blue and yellow act as hidden properties). When examining a lot of cards, he may (correctly) abduce that blue cards are even or (incorrectly) abduce that yellow cards are odd. When observing the yellow cards on both sides, he may induce (in an approximate way) that the proportion of odd cards is equal to their frequency.

2.8 Reasoning and belief revision

Many reasoning operations can be related to belief change by means of some well-adapted transcription principle. More precisely, the inference 'from A, infer C in context K' is transcribed into the belief change 'C results from initial belief K changed by message A'. The transcription principle works both ways, but requires that the reasoning operation takes place in some context K, acting as background knowledge and corresponding to the initial belief. In syntax, the axioms proposed for the reasoning operations can then be transcribed into belief change axioms (either in a revising or an updating context). In semantics, the reasoning operations can consequently be expressed as reasoning rules related to belief change rules.

For non monotonic reasoning, the transcription principle states that 'A normally entails C in context K' is equivalent to 'revising K by A validates C'. In syntax, the reflexivity axiom, the cut axiom and the cautious monotony axiom correspond respectively to the success axiom, the sub-expansion axiom and the super-expansion axiom. In semantics, each event E contains a 'normal part' denoted E', obtained by the intersection of E with the nearest corona from K (the coronas being defined in belief revision). Hence, A normally entails B if and only if its normal part A' is included in B. Expressed in a more quantitative way, A normally entails B if the 'necessity' (previously defined as an extension of probability) of C conditional on A is strictly positive.

For non-reflexive abduction, the transcription principle states that 'facts A abduce assumptions C in context K' is equivalent, for non-reflexive abduction, to 'C entails initial belief K revised by message A', and for ordered abduction, to 'initial belief K revised by message C entails initial belief K revised by message A'. In syntax, the axiom system of abduction (in both forms) is equivalent to the axiom system of revising, but the equivalence does not hold axiom by axiom. In semantics, abduction is again transcribed in rather complex rules acting on possible worlds. Of course, abduction is treated here in a given space of possible worlds, excluding any act of creativity that could imagine new worlds.

For conditional reasoning, the transcription principle, called the 'Ramsey test', states that a conditional 'if A, then C' is accepted with regard to some background knowledge K if 'initial belief K updated by message A validates C'. In syntax, a one-to-one correspondence between axioms of (subjunctive) conditional reasoning and belief change can again be made explicit, but the relevant belief change context here is an updating context and no longer a revising context. The reason is that the antecedent A of the conditional refers to a real transformation of the environment. In semantics, in some world w (acting as context or initial belief), the conditional 'if A, then C' is valid when, in the nearest world to w where A is true, C is true too. Such an intuition really justifies 'possible worlds' as worlds which are counterfactually accessible (Lewis, 1973).

It is not easy to find a probabilistic counterpart to the first two reasoning operations. For nonmonotonic reasoning, it is not enough to assume that A normally entails B if the probability of A conditional to B is greater than some threshold, as illustrated by the 'lottery paradox' (each lottery ticket does not ensure an actor will win the jackpot, but their disjunction does). For abductive reasoning, it is not enough to assume that B is abduced from A if the probability of A conditional on B is greater than the probability of A. But for the third operation, counterfactual reasoning, a transcription principle called the 'Stalnaker test' asserts that the probability of a conditional, with regard to some prior probability, is the posterior probability of the consequent with regard to the antecedent. However, since the relevant context is updating, the probabilistic) belief change has to be performed by Lewisian imaging and not by Bayesian conditionalization.

Since several reasoning operations were successfully reduced to belief change, the last appears to be the fundamental reasoning scheme. Some more recent trials have attempted to reduce two other reasoning operations to belief change. Firstly, 'analogic reasoning', which transfers structures from one domain to another, appears to be some sort of conditional reasoning, arguing that 'if two domains share similar primary properties, they must share similar secondary properties linked to the first'. Secondly, 'taxonomical reasoning' (especially 'pattern recognition'), which groups objects into homogeneous classes, appears to be some sort of abduction, arguing that 'if two objects share similar properties, they must obey the same laws'. But these two operations fundamentally rely on a similarity relation on worlds rather than a preorder relation on worlds.

Relevant for several types of reasoning, the Bayes rule is not adapted to conditional reasoning which involves an updating context. In particular, it does not satisfy the 'homomorphism axiom' (see 2.4), as can be illustrated by the 'Simpson paradox'. Consider two packs of two-sided cards. In the first pack, there are 12 cards, 8 blue ones and 4 red ones. Among the blue cards, 5 are even and 3 are odd, while among the red cards, 3 are even and 1 is odd. In the second pack, there are 8 cards, 2 blue and 6 red. Among the blue cards, 0 is even and 2 are odd, while among the red cards, 1 is even and 5 are odd. Consequently, in each pack, the blue cards are less often even than the red ones (in probabilistic terms). Now unite the two packs to make a single pack of 20 cards, of which 10 are blue and 10 are red. Among the blue cards, 5 are even and 5 are odd while among the red cards, 4 are even and 6 are odd. Somewhat paradoxically, the blue cards are now more often even than the red cards.

3

Decision-making as reasoning

> *The analysis of what people will do can only*
> *start from what is known to them.*
> F. von Hayek

According to a ‘minimal psychology’, a decision-maker defines a plan of action through a mental deliberation process based on three determinants acting as mental states. Besides his objective opportunities and preferences, he holds subjective beliefs about his physical environment, and moreover about his own opportunities and preferences. He combines them according to two forms of rationality, cognitive rationality linking beliefs to information and instrumental rationality linking opportunities and preferences. Each form of rationality is moreover bounded by cognitive constraints, both informational and computational, which are activated when he gathers and treats the relevant information.

In decision theory, the decision-maker follows a rational deliberation process broken down into several steps (3.1) and exemplified by the strong rationality model (3.2). Since he faces increasing forms of uncertainty (3.3), he reacts by implementing choice rules based on more and more sophisticated beliefs about his environment (3.4). Moreover, he has to frame his decision problem in relation to a reference situation (3.5) and he adapts his choice rules by categorizing the main features of his environment (3.6). Finally, he is constrained by his own bounded computational capabilities (3.7) and takes them into account by adopting simpler choice rules which appear simply as reasonable heuristics (3.8).

3.1 Rational choice models

Decision-making is first founded on a '*decomposability principle*' which asserts the existence of autonomous entities able to perform actions on a physical environment. It can be made more precise by means of two postulates. The 'actorialist postulate' states that the social fabric can be broken down into specific entities displaying independent behavior. The 'actionalist postulate' states that the actor's behavior can be broken down into parallel and sequential actions. The decomposability principle is at the heart of 'methodological individualism', which states that any social phenomenon can be reduced to the combination of actions of decision-makers. However, the decomposability principle does not assume that decision-making is only relevant at the level of individuals, but that it can also be considered at the level of collective entities (firms, governments, etc.).

The decision process is assumed to take place in three phases, relating the decision-maker reciprocally to his environment. The 'information phase' assumes that the decision-maker gathers certain signals about his environment, through sensors and filters, and then categorizes and organizes these signals into pieces of information. The 'deliberation phase' assumes that the decision-maker mentally processes his information in order to form an intention (or plan of action). The 'implementation phase' assumes that the decision-maker breaks down and schedules the plans of action, then transforms them into actions on the environment, through effectors and instruments. The three phases are assumed to take place sequentially without feedback from one phase to a preceding one.

Decision-making is essentially founded on a '*rationality principle*' which asserts that the deliberation phase is a kind of reasoning supported by certain mental states that are specific to the decision-maker. It can be made more precise thanks to two postulates. The 'consequentialist postulate' states that the decision is arrived at by considering only the consequences of each action. The 'utilitarian postulate' states that the consequences of an action are evaluated by trading off their costs and benefits. The rationality principle is founded on an 'intentionalist view' which asserts, contrary to a 'behaviorist view', that an actor's behavior can only be explained by reference to the way the decision is computed. But it does not assume that the mental states or the mental process combining them are conscious.

The deliberation phase can be broken down into three steps, each involved with a specific 'determinant' acting as a mental state. The 'generation step' assumes that the decision-maker, given his '*opportuni-*

ties', delineates a set of available actions. The 'prediction step' assumes that the decision-maker, given his '*beliefs*', predicts the consequences of each possible action on his environment. The 'evaluation step' assumes that the decision-maker, given his '*preferences*', weights the expected consequences in a synthetic utility index. Moreover, the mental states are considered as independent, especially as concerns beliefs and preferences, which are supposed not to influence each other.

Two types of rationality can be considered in relation to the different steps in the deliberation process (Walliser, 1989). 'Instrumental rationality' expresses the fit realized by the decision-maker between the means at his disposal and the objectives he pursues. 'Cognitive rationality' expresses the fit realized by the decision-maker between the representations he adopts and the information he possesses. On a first level, cognitive rationality concerns the formation of the decision-maker's beliefs about his environment and himself. On a second level, it concerns reasoning supported by all mental states and leading to a choice. Cognitive rationality may be reduced to instrumental rationality by assuming that the actor uses his information optimally to form his beliefs. But instrumental rationality is more naturally reduced to cognitive rationality, since opportunities and preferences are supported by beliefs and their combination is achieved through a specific form of reasoning.

The framework provided by the rational choice model is too general to be applied as such to a given decision problem. It needs a more precise specification of the determinants involved and of the principle which articulates them. For specific 'contexts' relating to the perceived environment, the rational model induces some well-defined 'choice rules'. Each choice rule combines the three determinants through formal structures constrained by analytical properties and including free parameters. Any choice rule can, moreover, be justified (or conversely criticized) by one of three types of argument. A 'theoretical justification' supports the rule by an explanation scheme often expressed as an axiom system. A 'praxeological justification' shows the rule's robustness when applied to specific environments. An 'empirical justification' simply demonstrates that the rule leads to actions consistent with given observations.

The decision-making process of a surgeon, faced with a patient, illustrates the three phases described above. The information phase consists in establishing a diagnosis of the patient's possible illness on the basis of clinical observations realized either directly (pulse check) or indirectly (blood analysis). The deliberation phase consists in choosing a treat-

ment appropriate to the probable illness, taking into account the state of the art of the surgical techniques involved and their expected cost and efficiently. The implementation phase consists in carrying out the operation, after defining its place, date and concrete modalities and taking into account certain hazards which may occur during the operation.

3.2 Strong rationality models

The most usual deliberation model states that the decision-maker is endowed with 'strong rationality'. Assuming that he has infinite computational capabilities, this model can be applied to cognitive as well as to instrumental rationality. Strong cognitive rationality asserts that the decision-maker is able to form perfect expectations. He gathers all relevant information about past history, he is endowed with a complete representation of his environment, he reasons like a perfect statistician who minimizes the prediction error on the expected variable. Strong instrumental rationality asserts that the decision-maker is able to find the best action for each environmental condition. Moreover, as he has perfect knowledge of his opportunities and preferences, he combines them to maximize his well-being, for any given beliefs.

The corresponding formal choice rule is the 'optimizing model'. The environment is represented by alternative 'states' generated by an 'environment law'. A decision matrix expresses the consequences of any combination of an action and a state, hence manifests a 'consequence law'. The environment law and the consequence law, and even the realized state, are well-known by the decision-maker. Hence, the preferences can be directly defined on actions and states rather than on consequences. Preferences are summarized in a synthetic 'utility function' which integrates the costs and benefits of an action on a unique scale. Opportunities are summarized in several analytical constraints defining an action set. Finally, the optimizing model asserts that the decision-maker chooses an action by maximizing his utility function under constraints. Such a deliberation process may, however, be explicit or implicit.

By solving the maximizing program for a given state, the decision-maker adopts a 'behavior rule' relating directly the action to the (known) state. A behavior rule is expressed by a conditional statement asserting how a decision-maker reacts to his environment: 'if state S, then action A'. It is of a factual type, since no state is presently realized and only one state will be realized in the future. Having eliminated the mental states, a given behavior rule is open to several interpretations. From a behaviorist view, it appears just as a stimulus-response rule, associating a purely mechanical action with a given situation. From an intentionalist view, it appears as the reduced form of a sophisticated choice rule, obtained by getting into the 'black box' of the decision-maker's deliberation. In fact, in some contexts and under some conditions, knowing the behavior rule, it is possible to reveal the utility function (and the opportunity constraints) of the decision-maker. Such an 'attribution process' is an abductive reasoning inferring mental states from the observed actions under the rationality principle.

The optimizing model is theoretically justified by an axiom system defined on a preference relation on actions (Debreu, 1954). The completeness axiom states that the decision-maker is able to choose an action in any pair of actions. The transitivity axiom states that if one action is preferred to a second and the second is preferred to a third, then the first is preferred to the third. The continuity axiom states that the preferences do not jump by discrete variations. Under these main axioms, a representation theorem states that the preference relation can be represented by a continuous utility function. More precisely, an action is preferred to another if and only if its utility is higher.

The optimizing model is praxeologically justified in two ways. An 'evolutionary argument' states that, if non-optimizing actors are confronted with optimizing actors, the former are eliminated. Such a confrontation is modelized in game theory (see 6.8) or in economic theory (see 8.8) and involves a learning or evolutionary process. The announced result can be obtained only in specific contexts and under drastic conditions: non-optimizing actors may survive among optimizing ones in a complex or fluctuating environment. A 'defeating argument' states that a non-optimizing actor, when confronted with a suitable sequence of choices, is condemned always to lose. More precisely, the 'money pump argument' shows that an actor, endowed with cyclical (non-transitive) preferences, can be ruined after a well-adapted sequence of choices. In fact, the transitivity axiom is fundamental for rationality and less can be said when abandoning it.

The optimizing model is empirically justified when the actions taken in given choice problems coincide with the optimizing ones, whatever his interpretation. It is empirically validated as 'substantive' when the chosen action coincides with the optimizing one, whatever the deliberation process actually implemented. An illustration is provided by the billiards player who plays as if optimizing the rebounds of his ball without explicit or even implicit calculation (Friedman, 1953). It is empirically validated as 'procedural' when the chosen action is obtained by an explicit deliberation process based on an algorithm (like the 'algorithm of gradient'). One example of this is the chess player who tries to optimize by using research and selection heuristics, even if he is not able to do so due to the complexity of the game (Simon, 1982). In any case, it is generally considered as empirically valid in specific contexts: environment not too complex, full information, clear consequences and stakes.

In the surgery example, according to the optimizing model, the surgeon computes the best treatment for the patient. He is able to make a list of all acceptable treatments, to predict their precise consequences and to attribute a synthetic utility to each one. He then merely chooses the treatment with the highest utility. Utility may be purely material (medical effect, costs) or include symbolic aspects (esthetics of the scar, competition with other surgeons). It may be purely individualistic (the surgeon's pay and learning) or include social aspects (prestige of the hospital, ethical concerns). Some axioms can be directly tested on the surgeon's preferences. Completeness is testable and generally confirmed. Transitivity can be tested by offering the surgeon successive choices between two treatments. Conversely, continuity is harder to test, although it cannot be considered as a purely technical axiom.

3.3 Sources of uncertainty

The decision-maker faces several sources of uncertainty, but they may formally be reduced to uncertainty about the states of the environment. Firstly, he may be uncertain about the states themselves or about the law governing their generation. Secondly, he may be uncertain about the relation linking the consequences to a pair of action and state. In this case, it is considered that he is in fact uncertain about some factor acting on that relation, and this factor is treated as a state. Thirdly, he may be uncertain about his own opportunities and preferences. But these determinants are considered to define the decision-maker's 'type' and that type is again treated as a state. Further, uncertainty about states is usually expressed in a semantic approach, where the worlds reflect the states and the beliefs about them. In fact, since the decision-maker is unique and assumed to satisfy positive and negative introspection, the states alone reflect perfectly his uncertainty. However, the decision-maker is not assumed to satisfy veridicity and may be wrong about the states.

The modeler generally considers that the environment behaves mechanically. Firstly, the environment does not exhibit rational behavior;

it is therefore neither benevolent nor malevolent towards the decision-maker. Secondly, the environment is not influenced by the decision-maker. However, the modeler may consider that the states of the environment are correlated to the decision-maker's actions. This may be due to the direct influence of actions on states or to a third factor influencing both states and actions. Thirdly, the modeler generally considers that the environment behaves in a probabilistic way. Its law of generation can then be reduced to an exogenous probability distribution on the states. However, the modeler may well consider that the states are generated more erratically.

As concerns the decision-maker, his uncertainty about the states of nature is analyzed in four contexts, corresponding to a progressive weakening in certainty. In 'certainty', the decision-maker knows the state that has already been produced (whatever its law of generation). In 'Bernoullian uncertainty', the decision-maker knows the (true) probability distribution from which the state is drawn. In 'Knightian uncertainty', the decision-maker only knows the set of possible states, and does not know the occurrence of the actual world. In 'radical uncertainty', the decision-maker does not even know a list of states. Of course, some intermediate situations are possible. For instance, the decision-maker may only have access to partial information about the probability distribution, such as an order on the occurrences of the states.

When the decision-maker is endowed with probabilistic beliefs, the probabilities involved originate in different ways. Objective probabilities may be stated by the modeler and result from the computation of proportions in a population or from the calculation of frequencies in a sequence of states. These objective probabilities are given by the modeler to the decision-maker, who may accept them as subjective probabilities. Alternatively, no objective probabilities may be known or at least given by the modeler. In that case, subjective probabilities may directly be formed by the decision-maker from his past experience. In other cases, the decision-maker may even form subjective probabilities about the objective ones. By introspection, subjective probabilities may be enounced by the decision-maker and accepted by the modeler if considered as unbiased and sincere.

Under uncertainty, the concept of an action has to be extended, but it is always defined by its conditional consequences. For Bernoullian uncertainty, the decision-maker has to choose between 'lotteries', a lottery being a set of consequences weighted by the probability of the corresponding states. For Knightian uncertainty, the decision-maker has to choose between 'acts', an act being a set of consequences conditional on states. Moreover, in some frameworks, a decision-maker may have to choose between 'menus', a menu being a subset of actions. But in any case, the decision-maker's behavior is deterministic, since he chooses one and only one action (except in the case of ties exemplified by the Buridan donkey).

For each context of uncertainty, decision theory proposes specific choice rules. In fact, only Bernoullian and Knightian contexts are really concerned, since it is very difficult to deal with radical uncertainty. The choice rules associate a utility function on consequences with various forms of belief. Strong instrumental rationality is always represented by an overall utility function which is maximized, but cognitive rationality is expressed by specific beliefs. Now, by observing the decision-maker's actions, the modeler can reveal both his utility function and his beliefs under some conditions. Practically, given the utility function, the modeler may just reveal the decision-maker's beliefs by observing several choices made. But such an attribution process is generally not univocal. Moreover, the modeler frequently faces a discrepancy between revealed beliefs and enounced beliefs.

In the surgery example, uncertainty seldom concerns the set of possible treatments, which is generally well-known, but it may concern the utility of each treatment, which is fuzzier. As concerns uncertainty on the environment, it is summarized into two relations. Firstly, a scientific relation associates (observable) symptoms with an (unobservable) illness and (poorly-known) environmental states. Secondly, a technical relation associates (observable) results with an (observable) treatment, an (unobservable) illness and (poorly-known) environmental states. Hence, the states reflect the random factors influencing the relation between a given treatment and its results, for a presumed illness. It is assumed that the states are not influenced by the treatment itself. However, it is possible to consider the illnesses themselves directly as states, assumed to influence the symptoms randomly. The modeler can associate these illnesses with objective probabilities or simply list them. Further, the occurrence of illnesses can be considered as independent of the treatment or influenced by it.

3.4 Choice rules under uncertainty

In a Bernoullian context, the original choice rule proposed by B. Pascal is the 'expected payoff rule'. It asserts that the decision-maker is endowed with a (true) objective probability distribution on states and selects the lottery with maximum expected payoff. Empirically criticized by N. Bernoulli (Bernoulli, 1738), it was extended as the ‘expected utility rule’. Outside the probability distribution on states, the decision-maker is endowed with a utility function on (sure) consequences, leading him to select the lottery with greatest expected utility. Empirically refuted by Allais (Allais, 1953), it was again extended by Quiggin (Quiggin, 1982) as the ‘rank dependent utility rule’. Two functions are introduced, a utility function on consequences and a (cumulative) probability deformation function. The latter states that probabilities are subjectively modified by underestimating low probabilities and overestimating high probabilities. In the two last choice rules, the decision-maker’s risk-aversion is expressed by the utility function and the probability deformation function respectively.

In a Knightian context, the usual choice rule is the ‘minimax rule’. It asserts that the decision-maker selects the act with the highest minimum payoff with regard to states. It reflects very cautious behavior on the part of the decision-maker and corresponds in fact to an infinite aversion to risk. A quite different choice rule is the ‘subjective expected utility rule’, which asserts that the decision-maker forms subjective probabilities on states, is endowed with a utility function on consequences and retains the lottery with the highest expected utility. An extension of the last rule is the ‘Choquet expected utility rule’ (Gilboa, Schmeidler). The states are no longer weighted by subjective probabilities, but by non-additive probabilities of a hierarchical nature. The decision-maker considers not only a (set-theoretic) uncertainty on states, but also a (probabilistic) ambiguity concerning the uncertainty on states. He manifests not only uncertainty aversion, but ambiguity aversion.

Finally, when the action is assumed to influence the state, the expected utility rule is generalized in two ways, depending on how the decision-maker is assumed to predict the state. The ‘evidential rule’ considers the probability of the state conditional on action; it is therefore based on the Bayes change rule (and is in fact analogous to the ‘expected utility’ rule). The ‘causal rule’ considers the probability of the conditional ‘if state S, then action A’, and is therefore based on the Lewis change rule (and incorporates considerations of dominance between actions). If the second applies in an updating context where the action really modifies the state, the first applies when there is just a correlation between action and state.

The main rules have been theoretically justified by axiom systems defined by preferences on lotteries or actions. They add certain specific axioms to the traditional axioms of choice under certainty, i.e. completeness, transitivity and continuity. The 'expected utility rule' was axiomatized by Neumann-Morgenstern (von Neumann, Morgenstern). The 'reduction of lotteries axiom' implies that the decision-maker reacts in the same way when facing a compound lottery (lottery on lotteries) as he does with the reduced lottery. The 'independence axiom' asserts that if one lottery is preferred to another, it remains preferred when they are each combined in identical fashion with a third lottery. The 'subjective expected utility rule' was axiomatized by Savage (Savage, 1954). The main axiom is the 'sure thing principle', which states that a decision-maker comparing two actions does not take into account their common consequences.

These rules are also justified or refuted by empirical arguments, which lead precisely to their extension. The 'expected utility rule' was empirically refuted by the 'Allais paradox', an experiment involving carefully selected pairs of lotteries. The axiom violated by the decision-maker in the experiment can be isolated and is nothing other than the independence axiom. By weakening the independence axiom into a 'comonotonic independence axiom', the axiom system leads precisely to the 'rank dependent utility rule'. The 'subjective expected utility rule' is empirically refuted by the 'Ellsberg paradox', an experiment involving carefully selected pairs of acts. Here again, the incriminated axiom is the 'sure-thing principle'. Its replacement by a weaker axiom leads precisely to the 'Choquet expected utility rule'.

Finally, some rules may be justified by praxeological arguments. For instance, the subjective expected utility rule can be justified by a 'defeating argument' called the 'Dutch book argument', which states that, if the states are evaluated by other weights than probabilities, the modeler is able to exhibit a sequence of acts leading to a sure loss by the decision-maker. However, this situation is very artificial, since the actor is never involved in such a sequence of choices and may even refuse to participate in it. Less often invoked, an 'evolutionary argument' states that actors who use probabilities (and even use the Bayes rule to revise them) perform better than others when they are compared together.

In the surgery example, the surgeon applies an expected utility rule when he attributes an objective probability to each possible illness and a utility to each combination of an illness and a treatment and then selects the treatment which shows the highest expected utility. He applies the minimax rule when he concentrates on the worst consequences of each treatment and selects the treatment which has the "least bad" worst consequence. The probabilities attributed to an illness may be enounced *by the surgeon or revealed by the modeler. Many axioms (completeness, transitivity, independence) can be empirically tested by offering the surgeon several choices between two treatments.*

3.5 Cognitive effects

The decision-maker frequently violates the decomposability and rationality principles which were introduced above (see 3.1). He deals with several choice problems together in space and time and is unable to reduce his multidimensional context to simple states of nature. He perceives his environment by means of a specific mental operation which isolates its essential patterns and categorizes it in several classes. In fact, psychologists have described many phenomena which are not directly consistent with the rational process underlying all choice rules. However, these phenomena are being increasingly integrated into choice rules through the use of auxiliary assumptions.

A first phenomenon appears at the interface between the information and the deliberation phase. It concerns the way in which the decision-maker 'frames' his environment, in other words the way in which he constructs taxonomies about it. The 'invariance principle' states that his determinants are independent of the presentation of the choice problem. It can be broken down into two axioms, related respectively to decision theory and epistemic logic. The 'similarity axiom' states that the decision-maker chooses in the same way when facing two similar situations. The 'extensionality axiom' states that two formally equivalent situations are considered by the decision-maker to be alike. The invariance principle has already been partially incorporated into choice under uncertainty by way of the 'reduction of lotteries axiom', which states that the decision-maker reacts in the same way to equivalent lotteries. But numerous observations show that two formally equivalent choice problems may lead to different actions when they are presented in different forms.

A second phenomenon appears at the interface between the deliberation and the implementation phase. It concerns the 'cognitive dissonance' of the decision-maker who ex post adapts his beliefs to the chosen action (and its realized consequences) in order to justify it. Cognitive dissonance is more specific than 'wishful thinking', which reflects the general influence of preferences on representations. It is closely linked to the 'rationalization' phenomenon, which consists in justifying an already-chosen action using different determinants from those used actually to make the decision. Once more, numerous observations show that beliefs and preferences are not designed exogenously and may even interact together.

A third phenomenon deals with a sophisticated link between beliefs and preferences. It concerns the 'reference situation' considered by the decision-maker when he analyzes the consequences of his actions. Such a reference situation acts as a 'normative' belief which influences preferences. It can be associated with the consequences of a 'reference action' such as 'no action' or a 'normal' action. It can be associated with specific consequences, for instance, when the consequences are of a monetary nature, the present wealth of the decision-maker. The status quo, in particular, is of great relevance for the decision-maker. Here again, numerous observations show that the reference situation may be adapted by each decision-maker to the choice problem he faces.

A fourth phenomenon concerns the fact that the decision-maker is influenced by non-consequentialist and non-utilitarian factors. Emotional effects, in particular, are being considered to an ever greater extent, although emotions are very diverse in nature and influence and are not precisely classified. For instance, the pleasure of playing and the anguish of deciding are two emotions that have long been considered to influence the choice process. Likewise, the expectation of good or bad consequences may create certain emotions which exert a specific influence on the choice process. In some cases, emotions are just considered as acting on the determinants, especially on utility or disutility, in order to reinforce or inhibit some arguments or even to add new ones. In other cases, emotions are considered as factors which interfere with the rationality principle by modifying the trade-off between the determinants and even replace it by inducing some kind of 'shortcut' decision. Hence, the decision process is either cognitive or affective, either controlled or automatic.

In order to study these phenomena more carefully, a general 'logification' of the choice process is progressively being developed. If beliefs are already expressed (explicitly or implicitly) in an epistemic logical framework, preferences are still expressed in a classical analytical framework. Likewise, the deliberation process is expressed in terms of plain calculation implemented by the decision-maker. More recently, in an integrated logical framework, beliefs like preferences are syntactically expressed by formulas, using a belief operator, completed by a preference operator and an intention operator. The deliberation process becomes no more than a process of logical reasoning performed through these formulas. In semantics, actions just become integrated into the possible worlds. Up until now however, the choice process has just been plainly transcribed from an analytical language into a logical language, but more original achievements are on the way.

In the surgery example, the invariance axiom is violated by the 'framing effect', which can be illustrated as follows (by adapting the 'Asian disease problem'). When comparing two new treatments applied to a whole population of patients, the surgeon prefers treatment A, where everybody dies with probability 1/3 or nobody dies with probability 2/3, to treatment B where exactly half the patients die. But he prefers treatment B' where half the patients are saved to treatment A' where nobody is saved with probability 1/3 and everybody is saved with probability 2/3. In fact, the two situations are formally equivalent: in the first case, the positive consequences of the treatment are described and the negative ones are implicit (as complements); in the second case, the opposite is true.

3.6 Contextual choice rules

First, contrary to 'epistemic beliefs' founded on empirical evidence, choice models introduce 'pragmatic beliefs' defined from a choice perspective (Bratman, Cohen, Stalnaker). Pragmatic beliefs are more voluntary, since the decision-maker may influence them in certain respects. They are also more contextual, since the decision-maker adapts them to the specific choice problem facing him. More precisely, if epistemic beliefs are endorsed according to their proximity to truth, pragmatic beliefs are endorsed according to their relevance for decision. However, the two types of belief are not independent of each other. If an epistemic belief is already validated, this is a good reason to adopt it as a pragmatic belief. Conversely, if a pragmatic belief is accepted, it can be considered as a good candidate for an epistemic belief.

Further, contrary to 'absolute preferences' independent of the choice problem, choice models introduce 'contextual preferences' adapted to a specific choice situation. Contextual preferences principally introduce a reference level for consequences, possibly different for each type of consequence. They are incorporated into 'prospect theory' (Kahneman, Tversky), which considers that the decision-maker treats gains and losses differently, considered in relation to a reference point. Likewise, they make it possible to consider the concept of 'regret', expressed by the loss of payoff from having chosen a wrong action, relative to the best one. In other respects, preferences can be parameterized by the states of nature (state-dependent preferences), by the beliefs or even by the action set of the decision-maker.

Finally, choice models become ‘hierarchical’ in the sense that the decision-maker proceeds on two choice levels. On the first level, he chooses an action according to a given choice rule (generally under uncertainty). On the second level, he chooses a choice rule according to the decision context he perceives. However, such a hierarchical choice raises two tricky problems, at least in a static framework. Firstly, there is no set of choice rules already available to the decision-maker. Such a set can only be exogenously given, even if it may result from past experience. Secondly, the meta-choice of a choice rule obeys no precise criteria. However, the ‘same’ rule (for instance animitation rule) can be used on both levels, even if it has to be adapted on the second level.

These extended choice rules may again be supported by axiomatic justifications. This is indeed the case for contextual preferences such as those involved in prospect theory or in state-dependent utility. Pragmatic beliefs, on the contrary, have not yet been axiomatized, as regards either their structure or their revision. Likewise, the hierarchical choice model is very hard to axiomatize and even to justify less formally. In fact, the axioms which are required are always more complex in their expression and are not easy to interpret. They are adapted to a specific feature of choice and consequently look rather *ad hoc*. Moreover, only a few axioms can be empirically tested separately.

The extended choice rules are seldom supported by praxeological justifications. However, a paradox appears in the formulation of the choice rules (Arrow, 1974). When shifting from certainty to uncertainty, the information needed to apply the rule becomes drastically more refined. In particular, probability distributions of the states of environment are needed. When shifting from uncertainty to context dependence, this information is increased still further. In particular, the reference points have to be made explicit. From the modeler’s point of view, this is no problem, since he is assumed to be able to compute any problem. But it is problematic from the decision-makers point of view, since he only has bounded capacities for treating information (see 3.7).

Finally, the extended choice rules can be supported by empirical justifications. In fact, these rules were proposed precisely in order to integrate phenomena already described by psychologists. However, it is difficult to test them just by observing the actions of the decision-maker. In particular, it is hard to reveal the determinants of the decision-maker (beliefs as well as preferences) because they have become too sophisticated. It is necessary to consider the determinants as they are described by the decision-maker himself. Consequently, some introspection has to be accepted, even if it is affected by systematic biases. The biases are involuntary when the decision-maker has limited introspection facilities. They are voluntary when the decision-maker wants to give an idealized image of himself.

In the surgery example, the surgeon has subjective choice determinants. He perceives his opportunities as a set of treatments limited by technical constraints of feasibility, financial constraints of cost and ethical constraints of legitimacy. He adopts representations as structural links between each treatment and its short-term consequences for the curing of the illness or its long-term consequences for the patient's recovery. He holds preferences producing a global appraisal of each treatment, through the evaluation and aggregation of its therapeutic virtues, its psychological impact and its financial consequences. In particular, he has a reference point which may be the results he has obtained in the past, the results obtained by colleagues or a target he has set himself as a personal challenge.

3.7 Computational limitations

The decision-maker has 'bounded rationality', since he has 'limited capabilities for gathering and treating information' (Simon, 1982). This has led to a closer examination of procedural rationality as a means of bypassing these cognitive constraints. Bounded rationality is informational when it deals with the data gathered by the decision-maker (perception biases, information costs). It is computational when it deals with the reasoning process implemented by the decision-maker (memory limitations, calculation costs). However, since imperfect or incomplete information is generally considered as an independent phenomenon, bounded rationality is essentially focused on limitations to reasoning. Contrary to strong rationality, which is uniquely defined, bounded rationality has given rise to a lot of models in several directions, in terms of both cognitive and instrumental rationality.

Bounded cognitive rationality concerns the process by which the actor forms his beliefs from available information or previous ones. On the one hand, he is generally unable to frame his decision problem in a realistic way, due to the multiplicity of relevant variables. He merely describes it in a stylized way, by considering simplified categories for modeling both his environment and his own determinants. On the other hand, he is generally unable to predict perfectly the consequences of his actions on the basis of his given beliefs. He merely performs approximate expectations, based on partial information and simplified representations. Hence, bounded cognitive rationality is sometimes analyzed as a lack of 'logical omniscience', since the actor is unable to deduce all the consequences of what he knows.

Bounded instrumental rationality concerns the reasoning process by which the actor makes his choice by combining his three determinants. He is generally unable to compute an optimal action because the calculation is too complex. This is even more apparent for choice under uncertainty than for choice under certainty, especially when there are numerous local optima which are difficult to differentiate. Two main strategies are used to surmount this complexity problem (Simon, 1957). Firstly, he may perform an optimal choice, but based on simplified determinants. Secondly, he may perform a sub-optimal choice based on the original determinants. In practice, bounded rationality is expressed by choice rules which always combine the actor's determinants in a specific way, but the combination is no longer an optimizing one. It is therefore much harder for the modeler to reveal the determinants from the observed actions.

In one direction, bounded cognitive rationality may be represented by some kind of instrumental rationality. This means that the belief formation process is treated as a plain decision process. On the one hand, the actor is assumed to attribute some 'cognitive relevance' to a belief; if expressed on a unique scale of measurement, it appears as an 'epistemic utility'. On the other hand, the actor bears some 'cognitive costs' due to the effort he makes to treat and interpret his information. The choice rule is ideally a maximizing one, when the actor maximizes the cognitive relevance of his belief for a given cost (or conversely when he minimizes the cognitive costs for a given cognitive relevance). The choice rule is more realistically non-optimizing when the actor retains the first belief which gives him cognitive relevance above a certain threshold and induces a cognitive cost under a certain threshold.

In the other direction, bounded instrumental rationality may be deduced from explicit cognitive limitations which are compensated in several ways. A non-optimizing choice rule is usually justified informally by physical or epistemic considerations at a meta-level. For instance, the 'inert rule' asserts that in any given period, the decision-maker uses the same action he used in the previous period, if he can. This is attributed to high adaptation costs when switching from one action to another. Similarly, the 'random rule' states that the decision-maker chooses a random action in some more or less restricted action set. This is justified by the absence of information about the consequences of the action or the cost of reasoning to find the best action. Finally, the ' mimetism rule' asserts that a decision-maker takes the same action as his neighbor. This is related to the conviction that the neighbour knows more than he does. But it is generally difficult to give a more formal justification of a choice rule by making explicit the cognitive constraints faced by the decision-maker.

In a more sophisticated way, a logical limit to strong rationality in relation to cognitive constraints appears in the 'meta-optimization paradox' (Mongin, Walliser). The decision-maker considers an optimizing problem where he faces high computational costs. He may then use suboptimal choice rules (in some set of rules) which involve lower costs. He performs a meta-optimization where he compares different rules by trading-off their loss of utility and their cost (which are, even more paradoxically, assumed to be known without having to compute for each rule). But the meta-optimization procedure has a cost of its own and therefore requires the decision-maker to perform a meta-meta-optimization. The infinite regression cannot be stopped, since the costs grow higher with every level reached. The only solution is to optimize at some arbitrary level while passing over the cost at this level, and then to work down through the lower levels to the first.

In the surgery example, the surgeon has bounded rationality because he has a lack of information about the heterogeneous patients and has above all limited time to make his decision in the face of them. He uses the 'random rule' when he has little information about the consequences of treatments for an unknown disease and is therefore unable to rely on his beliefs and preferences. He uses the 'inert rule' when he considers that the last treatment is still acceptable, bearing in mind the high cost of switching to another treatment. He uses the 'mimetism rule' when he adopts a treatment already implemented by a colleague considered to be particularly clever or successful.

3.8 Bounded rationality models

Representative of a first class of choice rules, the 'satisficing model' (Simon, 1955) opposes the optimizing model on the basis of the adage: "its better to let well alone". The decision-maker considers a set of indexactions and pursues partial objectives not summarized in a unique index. Moreover, the actions are ranked in a given order and the partial objectives are provided with 'aspirations levels', both elements being considered as exogenous (in statics). The decision-maker finally selects the first indexaction which exceeds the aspiration levels for all objectives. As a special case, the ϵ-rationality model (Radner, 1980) considers only one objective and introduces an aspiration level equal to the maximal possible utility up to ϵ. However, the decision-maker is then assumed to know without computation the maximal utility.

Representative of a second class of choice rules, the 'stochastic choice model' (Luce, 1959), also called the 'discrete choice model', opposes the optimizing rule by introducing an element of randomness into the choice. As usual, the decision-maker is endowed with a set of actions and attributes a synthetic utility to each action. But he selects an action with a probability which is an increasing function of its utility. As a special case, the 'proportional model' considers that the choice probability is proportional to the utility. Similarly, the 'logit model' (inspired by physics) considers that the choice probability is proportional to the exponential of utility with a parameter μ.

Representative of a third class of choice rules, the 'automaton model' directly introduces certain computational constraints for the decision-maker in two forms. On one hand, the actor has a finite number of internal states and is therefore limited in the complexity of his calculations. On the other hand, the actor faces calculation costs for all the elementary operations needed in his computation, especially when applying specific algorithms. As a special case, the decision-maker is assumed to be able to deal only with certain specific analytical functions (computable, recursive). Considering a more precise constraint relative to a specific determinant, the decision-maker may be subject to computational limitations when implementing possible strategies.

The bounded rationality models seldom admit of precise theoretical justifications in terms of an axiom system. At best, they present informally the 'missing link' which relates them to cognitive rationality constraints. The satisficing model was justified by the fact that the decision-maker only has a partial preference ordering on the set of actions, even if this partial ordering is not itself justified. The stochastic model receives a first justification asserting that the decision-maker has a random utility function, but still optimizes. The uncertainty on preferences is transformed into an uncertainty on actions, each action being chosen with the probability that it is the optimizing one. When the probability distribution is correctly chosen (double exponential), one obtains precisely the 'logit model'. The stochastic model receives a second justification asserting that the decision-maker chooses an action by means of a trade-off between its utility and a control cost, with regard to a reference action. When the cost of control is correctly chosen (in an entropy form), one again obtains the logit model.

The bounded rationality models receive a weak praxeological justification by the fact that they admit the strong rational model as a limit case in some circumstances. At least, the chosen actions converge towards the optimizing ones. The satisficing model is equivalent to the optimizing model when the aspiration levels are fixed high enough. The stochastic (logit) model is identical to the optimizing model when parameter μ tends to infinity. The automaton model converges towards the optimizing model when the internal states tend to infinity or the calculation costs to zero. This property ensures that the models are immune to certain manipulations in limit circumstances, but not that they are robust in usual circumstances. Besides, the modeler is induced to specify the computational complexity of the problem faced by the decision-maker, for instance according to the classification used in computer science (polynomial, exponential, NP-complete). He has then to assume that the decision-maker cannot solve complexity problems that he cannot himself solve.

As for the empirical justifications, the bounded rationality models were constructed precisely in order to be more realistic. In fact, it is hard for the modeler to test if such and such model is really implemented by the decision-maker. It is even hard to reveal its parameters once it supposed that it applies. More and more, a given model is associated to a given context (type of environment, of objectives, of strategies) where it may be efficient. Various classes of models are then imagined outside the three former classes. For instance, Heiner (Heiner, 1983) proposes a model where the same action is used in similar environments when many of them may arise. Likely, Tversky (Tversky, 1972) proposes the 'elimination by aspects model' which treats partial preferences in a sequential way.

In the surgery example, the surgeon may consider simplified determinants. He may reduce his objectives to some short-term medical efficiently criteria and limit his opportunities to those permitted by the immediate availability of hospital personnel and material. He has certain 'mental maps' relating symptoms to illnesses and relating treatments (and illnesses) to their results. The surgeon frequently applies bounded rationality models. He uses a satisficing model when he contents himself with achieving reasonable efficiently at a reasonable cost. He uses a stochastic model when he is unsure about his own objectives and acts upon his present 'state of mind', which varies stochastically.

4

Dynamic action and belief revision

Once I have taken my decision,
I hesitate for a long time.
J. Renard

During a sequential decision process, the decision-maker either spontaneously or more deliberately gathers certain information about nature's past states and his own past results. His deliberation process is adapted to the dynamic context: instrumental rationality thanks to the 'backward induction' procedure and cognitive rationality according to the 'forward induction' principle. Since information is used by the decision-maker for revising his representations of his environment, it acquires an operational value and is therefore chosen strategic ally. However, condemned as he is to bounded cognitive rationality, the decision-maker is led to use heuristic learning rules to adapt his behavior to a complex environment.

Applied to repeated choice problems, the rational decision process has to be suitably extended (4.1), at first as concerns the strong rationality model (4.2). Dynamic uncertainty is treated by alternating actions of the decision-maker and states of nature (4.3) and various forms of uncertainty about nature lead to specific choice rules (4.4). The strategic role of information is basically summarized in the 'value of information' (4.5) and deeply embedded in a trade-off between exploration and exploitation (4.6). Bounded rationality proves to be even more relevant in a dynamic framework (4.7) and is compensated for by implementing more or less sophisticated learning rules (4.8).

4.1 Intertemporal rationality

The decision-maker is now involved in a finite or infinite sequence of related decisions faced with nature. When adapting the standard decision process previously considered in statics to dynamics, its three phases may be performed in each separate period or over several periods taken together. In fact, by definition of a sequential decision, the implementation phase has to be repeated in each period. The decision-maker regularly brings into play a certain action inducing (together with preceding actions) certain consequences for the environment. The information phase, on the contrary, can be performed in each period or globally. The decision-maker, even if gathering information regularly, may exploit it at each occurrence or only after a sequence of occurrences.

Concerning the deliberation phase, two extreme attitudes are possible. An ‘off-line’ policy consists in deciding all future actions from the start and then implementing them as time goes by. More precisely, the decision-maker chooses a ‘strategy’, defined as the action he would implement in all possible configurations he may meet. An ‘on-line’ policy consists in deciding in each period what action to take and implementing this decision immediately. More precisely, the decision-maker instantaneously chooses an action which is appropriate to the perceived present configuration. A strategy is richer than a sequence of actions, since it says, in a counterfactual way, what the decision-maker would do even in unrealized configurations. Of course, some intermediate attitudes are possible. For instance, the decision-maker may decide a whole strategy for the future, but adapts its application to non expected recent information.

In the same spirit, the choice determinants have to be adjusted to a dynamic setting. As concerns opportunities, they may be limited by constraints acting simultaneously on the actions of several periods. However, intertemporal constraints can often be dynamically decentralized into constraints acting in each period. As concerns beliefs, they have to take into account the evolution of the environment and the link connecting consequences to actions and states. Here again, representations frequently adopt a static specification of the environment law and the consequence law, but with parameters evolving through time like the system itself. As for preferences, they are defined in terms of the consequences arising jointly for successive periods. In practice, preferences defined on each period are generally aggregated into intertemporal preferences by means of a discount rate.

Moreover, the choice determinants may well evolve over time. This is always the case for the representations which are revised according to new information. Opportunities and preferences, on the contrary, were classically considered as exogenous and stable, unconditioned by social factors. However, they are increasingly being considered as subject to evolution due to past actions and states. As regards opportunities, they depend on past actions (investments) and even on past states of the environment (technological change). As regards preferences, they evolve under the influence of past actions (addiction, sensitivity to music) or past states (sensitivity to weather). It is then possible to think of the decision-maker as being broken into successive 'selves', each self having his own local choice determinants, even if the multiple selves must be more or less tightly linked by a meta-self.

As concerns the deliberation process, the rationality principle which integrates the choice determinants is itself adapted to a dynamic framework. Cognitive rationality deals with the evolution of representations in relation to observations about the environment and introspection about the decision-maker's own opportunities and preferences. It generally works forwards in time (forward reasoning), since it is based on constantly renewed information. Instrumental rationality deals with the adaptation adaptation of intended actions to opportunities and preferences, given the representations. It frequently works backwards in time (backward reasoning), since the decision-maker first considers his long-term actions and then returns to shorter term actions. Moreover, if a reference situation is considered, it may well evolve through time according to past experience.

In a broad sense, any partial task faced by the decision-maker in a given phase of his decision process may be considered as a kind of action. An 'informational action' consists in gathering information from the environment about the choice problem under consideration. A 'deliberative action' consists in treating the information gathered for the choice problem. An operational action' consists in a physical modification of the environment as a response to the choice problem. The operational action is the main subject of the decision process and has its proper choice determinants. The other actions are auxiliary and have specific but related choice determinants. In particular, informational and deliberative actions are evaluated by special preferences which assess their impact on the choice (especially consequences) of the main operational action.

In the surgery example, the operation is now realized in two steps. The first step consists in opening the body and obtaining some infor-

mation about the illness. The second step consists in applying a chosen treatment to the patient. The surgeon frequently begins the operation having already decided on the treatment, on the basis of already observed symptoms. But he may nevertheless adapt the treatment to what he finds in the first step. Of course, the two steps are related in terms of possible actions: the way the surgeon opens the body gives more or less information and influences the subsequent treatment he may apply. The surgeon's representations change in accordance with the new information. His preferences, related to the consequences of the treatment, are generally exogenous.

4.2 Strong rationality dynamic models

In the rational model, the opportunities are perfectly formalized by a 'decision tree'. The nodes correspond to the play either of the decision-maker or of nature. The vertices issuing from a node correspond respectively to possible actions of the decision-maker or to possible states of nature. This tree imposes an order of play and allows the action set (or the state set) available at any node to depend on past actions and states. A succession of moves defines a 'history' (or 'path') in the tree. It begins at the root node and may either finish at an end node (if the tree is finite) or continue without end (if the tree is infinite). The consequences of the combination of actions and states may be considered at each node or only taken into account at the end of the path.

Preferences (defined on the consequences) are expressed accordingly on the former decision tree. If consequences are obtained at the end nodes, utility is directly defined on them. If consequences are obtained at each node, a local utility function is defined on them. In that case, they are aggregated by weighting them in an intertemporal utility function. More precisely, a discount factor weakens the utility of a given period compared with the preceding period. The exponential intertemporal utility function consists in discounting successive utilities by a constant discount factor. The hyperbolic intertemporal utility function applies a high discount rate in the first period, then a lower one in all further periods. The first option is the most frequently used, but the second seems to be more realistic. Moreover, it can be observed empirically that losses are less discounted than gains.

Strong cognitive rationality entails that the representations of the decision-maker are complete and perfect. This means that the decision-maker knows perfectly the decision tree, i.e. his opportunities and s together with the possible s of . Moreover, during a given play, he also knows perfectly his position in the decision tree, i.e. his past actions and the past states of nature. Finally, he also knows the future states of nature. Since he knows his position in the decision tree, a strategy is defined as the action the decision-maker chooses at each node where he has to move. Since he knows the future states, the strategy can even be reduced to one path in the game tree.

Strong instrumental rationality is embedded in the so-called 'Bellman principle', which states that if a path is optimal, any of its subpaths beginning at one or another node is also optimal. It leads the decision-maker to apply the 'backward induction procedure', which works backwards in time, for finite game trees. The decision-maker first makes an optimal choice at the penultimate nodes; he then makes an optimal choice at the nodes preceding the penultimate ones, now knowing what he will do afterwards; he follows the same procedure back to the root node. As expected, this procedure is consequentialist, since it judges a strategy only on its future consequences, and utilitarian since it realizes a trade-off between successive utilities.

Any sequential decision process faces a 'dynamic consistency' problem. Dynamic consistency means that, if a decision-maker chooses today some action for tomorrow, he will not change his mind when tomorrow comes. More precisely, the decision-maker already decides his future actions at the root node of the tree, and does not change his plans when he arrives at a further node and considers the subtree starting at that node. In other words, dynamic consistency imposes that the successive selves are well-coordinated by the appropriate aggregated preferences. If the intertemporal utility function is exponential, the backward induction procedure effectively leads to a dynamically consistent plan of action. But this is no longer the case with a hyperbolic intertemporal utility function.

When considering evolving preferences, the backward induction procedure always applies by considering the decision-maker to be, at each node, the corresponding self with his local preferences. The plan of action obtained by the backward induction procedure is still dynamically consistent. But other common procedures are no longer dynamically consistent. The 'myopic procedure' works forwards in time since, at each node, the corresponding self takes the best present action according to his local preferences, given the preceding actions (but violating the preceding plans of action applied to the current period). The 'resolute procedure' consists, for the self at the first node, in adopting a plan of action for all further periods according to the local preferences, and keeping to this plan whatever the opposition of the future selves.

In the surgery example, one may consider that, due to tiredness, the surgeon modifies his preferences between the first and second steps of the operation. The 'backward induction procedure' consists, at the beginning of the operation, in choosing the treatment for the second step by predicting the final preferences, then choosing the opening mode of the first step in keeping with initial preferences. The 'myopic procedure' consists in choosing the best opening mode in keeping with initial preferences, then the best treatment for the second step in keeping with final preferences. The 'resolute procedure' consists in choosing, before the operation, both the best opening mode and the best treatment in accordance with initial preferences, then keeping to the first chosen treatment even if it is refuted by modified preferences. Dynamic consistency is ensured in the first case alone.

4.3 Dynamic uncertainty

In a dynamic setting, the uncertainty of the decision-maker can be assessed in terms of two criteria. As before, uncertainty is 'external' if it concerns the properties of the environment symbolized by nature and 'internal' if it concerns the characteristics of the decision-maker himself. Further, it is 'structural' if it concerns the behavior of nature or the decision-maker's choice determinants and 'factual' if it concerns past states of nature or past actions of the decision-maker. Four sources of uncertainty are obtained by crossing these criteria, but in the standard decision tree they are reduced in such a way that they can always be attributed to nature.

The behavior of nature, as seen by the modeler, now has to be made more precise. Nature is generally considered as defining its states at a node in a stochastic way, i.e. in accordance with a probability distribution. Moreover, when considering successive nodes of nature, two polar cases are possible. On the one hand, nature may act independently at each node, with a different probability distribution each time. On the other hand, nature may act in the same way at each node, according to the same probability distribution. Some intermediate cases are also possible, where the probability distributions are linked. For instance, nature may take a certain state at a given node and express a message correlated with that state at a later node.

As concerns external and structural uncertainty, the decision-maker is more or less aware of the probability distribution of states at each node. Frequently, he knows the general stochastic law governing the states, except for a certain parameter treated as a state of nature. As concerns internal and structural uncertainty, the decision-maker may not be aware of his own determinants, especially his own preferences. These determinants are summarized in a ‘type’ of the decision-maker. Moreover, the type is assumed to be determined by nature at the beginning of the decision process. Finally, the type is considered as selected according to a probability distribution in a set of possible types. Hence, in both cases, structural uncertainty is transformed into factual uncertainty.

As concerns external and factual uncertainty, the decision-maker is more or less aware of past states. However, even if he does not know what states are actually realized during the decision process, he is assumed to know them at the end. As concerns internal and factual uncertainty, the decision-maker is generally considered to know his own past actions. However, there are some counter-examples, like the ‘absent-minded driver’ who has to negotiate two crossroads in succession and cannot remember, when coming to the second, if he has already crossed the first. In both cases, the decision-maker is unable to distinguish between past nodes in the decision tree. The non-distinguished nodes are gathered into an ‘information set’. Of course, the actions available to the decision-maker at each node of his own information set have to be the same, because otherwise, he would be able to find out where he is. In this enlarged framework, a ‘strategy’ is defined as the action the decision-maker selects for each information set.

A slightly more general framework is provided by the ‘stochastic decision theory’. The system comprised of nature and the decision-maker admits a finite number of internal configurations. It shifts from one configuration to another for any pair of an action and a state. Since nature is stochastic, a transition probability is defined from one configuration to another, conditional on an action. Likewise, a transition utility is defined from one configuration to another, conditional on an action. The system can again be represented by a graph, but it no longer resembles a tree, since it is possible to retrace the same configuration many times. A strategy is now defined by the action played by the decision-maker in each configuration. The expected utility of a strategy is defined by discounting over time the expected utilities progressively obtained on the possible paths induced by that strategy.

Finally, in order to reduce the uncertainty he faces, the decision-maker gathers information through three different modes of experimentation. In ‘pure experimentation’, the decision-maker performs a purely informational action. He buys exogenous information from a specialized office, at a certain cost. In ‘passive experimentation’, the decision-maker performs a purely operational action. As a by-product of the decision process, he receives some endogenous information without any cost. In 'active experimentation', the decision-maker performs a 'mixed' action. He modifies his 'natural' operational action in order to receive original information, but suffers some disutility as a consequence.

In the surgery example, all uncertainties may be incorporated into a decision tree, even if they cannot be probabilized. Some uncertainties concern the operating theatre (means, personnel, environment) and define its reliability. Some uncertainties concern the surgeon and correspond to involuntary deviations in the operation process. Some uncertainties concern the patient and correspond to different random reactions to operations. To limit the last of these three categories, the surgeon may acquire information on the patient in three ways. Pure experimentation is achieved by means of prior blood analyses. Passive experimentation is achieved by observing the ill organ during the operation. Active experimentation is achieved by carrying out specific operating treatments which give more information than the usual one.

4.4 Dynamic choice rules under uncertainty

Under uncertainty, the rationality principle can still be expressed in its strong form. It is embedded in several classes of choice rules. These are always extensions of choice rules defined in a static setting, by adding some new principles involving reasoning in either direction of time. In particular, they associate time and uncertainty by considering dated lotteries or acts. They can once again be interpreted in a behavioral or a more intentional way. However, the explicit introduction of time into the decision-maker's deliberation process tends to favor a realistic approach, hence 'procedural rationality' rather than 'substantive rationality' (Simon, 1976).

Assume that the decision-maker knows a prior probability distribution for states and the utility obtained on each path. The 'dynamic Bayesian choice rule' then operates in two steps on the decision tree. Firstly, the decision-maker computes the posterior probability distribution at each node corresponding to nature. He revises the prior probability distribution in accordance with the messages obtained along the path leading to that node. He applies the Bayes rule because the context is a focusing one (one specific state is drawn). Secondly, the decision-maker uses the backward induction procedure from the terminal nodes to the root node. When the node corresponds to nature, he computes the average utility for all states associated with the node. When the node corresponds to the decision-maker, he selects the action with maximal (expected) utility and attributes that utility to the node.

The use of the expected utility choice rule in dynamics is supported by further theoretical justifications (Sarin, Wakker). For instance, the independence axiom can be derived from even more profound axioms. The 'separability (or consequentialist) axiom' states that the optimal strategy obtained in a subtree only depends on the elements of that subtree. The 'dynamic consistency axiom' states that if the path associated with the optimal strategy meets a given node, this strategy is again optimal in the subtree beginning at that node. Moreover, the 'reduction of lotteries axiom' states that the decision-maker is indifferent to the aggregation of two successive moves of nature into a single one according to standard probabilistic rules.

Praxeological justifications can also be given to the expected utility choice rule applied in dynamics. Defeating arguments, like the dynamic 'Dutch book argument', may justify that the decision-maker revises his beliefs according to the Bayes rule. Evolutionary arguments, applied to a repeated static or dynamic game, may also justify the Bayes rule or even the backward induction procedure. Conversely, empirical violations can be observed as concerns the rule itself or its underlying axioms. For instance, the separability axiom is violated by the 'sunk cost fallacy'. If a decision-maker has already spent some money on a project under favorable expectations, it is in his interest to stop investing when the future becomes less favorable, but he actually continues, notwithstanding.

In stochastic decision theory, knowing the transition probabilities and utilities, the optimal strategy maximizes the discounted sum of local expected utilities on an infinite horizon. It is again obtained by a backward induction procedure on related configurations. The procedure computes the maximal expected utility that the decision-maker can obtain in each configuration for each action (Bellman equations). The optimal strategy proves to be Markovian (the chosen action is independent of past states) and stationary (the chosen action is independent of time). The procedure can be generalized to structural uncertainty where transition probabilities and utilities are not well known.

A very particular model is the 'case-based model' (Gilboa, Schmeidler), which is static in nature, but based on a comparison of the present choice with past choices. By definition, a case is formed of a problem, an action and a result. The decision-maker stores a set of past cases in his memory. A similarity function provides him with an index of proximity between any two problems he may encounter. A utility function provides him with a utility index for any result (eventually normalized by an aspiration level). The overall utility of a new case is obtained by summing up, for all remembered cases, the utility of their past results weighted by their similarity to the new case. The choice rule simply consists in taking the action which maximizes the overall utility. However, the case-based model will be more naturally expressed by truly dynamic learning rules (see 4.8).

In the surgery example, the nurse prepares a soporific for the patient. It is obtained by mixing three liquids in a cup. Each liquid may be rotten or not (with a given probability) and this can only be observed after pouring it. She can proceed either by pouring each liquid directly into the cup or by first pouring it into an intermediate container. When a liquid is poured into the cup and turns out to be rotten, the whole product is lost. When a liquid is first poured into the container and turns out to be rotten it can simply be thrown out; when it is not rotten, it is transferred into the cup. The decision problem is a stochastic one, the configurations being formed by the number of liquids already poured into the cup. Its solution depends on only three costs (the cost of a liquid, the cost of transferring the liquid and the cost of losing the product).

4.5 Value of information

The decision-maker may acquire information about his environment through a process of pure experimentation. Formally, he is involved in a two-period decision process where an operational action is preceded by an informational action. In the first period, he may buy a true message (from a set of possible messages) about the real state of nature (already fixed). The message has an exogenous cost and allows him to transform his initial belief about the state of nature into a final belief. In the second period, the decision-maker chooses an action conditional on his final belief by applying a choice rule. The action provides him with an (expected) utility.

As concerns the state of nature, the decision-maker knows that it stems from a prior probability distribution. As concerns the message, the decision-maker knows that it is true and how it depends on the actual state. A set-theoretic message defines a subset of states containing the real state, the possible messages defining a partition on the set of states. A probabilistic message defines the probability of any signal received conditional on the real state. The common limit case is obtained when the message precisely indicates the real state (a certain message).

In both cases, the decision-maker uses the given elements to compute the probability of the state conditional on the message.

The '*ex post* value' of information is nothing other than the difference between the utility obtained with the message and the utility obtained without the message. It naturally depends on the precise message the decision-maker receives. It can only be computed before choice by the modeler. The '*ex ante* value' of information is the expectation of the ex post values for all possible states (or messages). It is defined before receiving the message, on average for all possible messages. It can be computed before choice by the decision-maker himself. In order to choose an informational (and operational)operational action, the decision-maker proceeds by backward induction. Concretely, he decides whether or not to buy a message by comparing the *ex ante* value of the information with its cost.

For any decision-maker, the *ex post* value associated with an (uncertain) message may be negative. When receiving some improbable signal about the real state, he may shift to an action which is worse than the initial one. Conversely, for a decision-maker endowed with strong (cognitive and instrumental) rationality, the *ex ante* value of information is always positive. When receiving a true message, he cannot, on average, find himself in a worse situation than before. The reason is that the number of conditional actions at his disposal increases, the old one still being available. Nevertheless, this fundamental result is invalidated when the decision-maker uses a weaker choice rule than the expected utility rule. Likewise, it is invalidated when the message (considered as a specific type of belief) does not satisfy the introspection axioms (non-partitional set-theoretic message, probabilistic message with non-additive probabilities).

The two-period choice model can be extended to the acquisition by the decision-maker of endogenous information about nature. For instance, an investor may construct certain equipment in two steps in order to adapt to the demand observed in the first step. In the first period, the decision-maker has to choose between a reversible and an irreversible operational action. This action provides additional information about the state of nature and leads him to revise his prior beliefs. When the first action is reversible, the decision-maker can complete it by another operational action in the second period if and only if the state is favorable. A reversible action allows the decision-maker to profit from further information, but acting in two steps is more costly than acting in one. With the usual choice rule, it can be proved that the reversible action implicitly receives a 'flexibility bonus' compared with the irreversible action. This bonus is precisely equal to the value of the information given by the message.

The choice model can also be extended to a search implemented by the decision-maker with the aim of acquiring new information about the monetary payoff of an operational action. For instance, a consumer may prospect on the market of a desired good in order to observe a sample of prices for the good. In the first period, the decision-maker can observe, in succession, the payoff associated with each possible action he might subsequently implement. He knows the cost of each observation as well as a prior distribution of payoffs over all homogeneous actions. In the second period, the decision-maker chooses (without additional cost), from among those actions of which he has observed the payoffs, the one with the highest payoff. It can be proved that the rational decision-maker searches until he finds a payoff above a certain threshold called 'reservation value'. This value decreases with the unitary cost of prospecting (he searches more when the cost is low) and increases with the variance of the payoff distribution (he searches more when the variance is high).

In the surgery example, a patient may be of three equiprobable types, a, b or c, which react more or less favorably to an operation. The operation has a utility of +6 if it succeeds (type a or b), −9 if it fails (type c) and 0 if it is not performed. Without information, the surgeon always operates and obtains an average utility of 1. With perfect information, he only operates for type a and b and obtains an average utility of 4. The value of (certain) information is then equal to 3. As an intermediate informational situation, a test may give no information for type c and fail to discriminate between the real type and type c when the real type is a or b. Consequently, the surgeon only operates for type c and obtains a utility of −1. Compared to the situation without information, the value of this information is negative. This is due to the unusual fact that the message is not partitional and violates the introspection axioms.

4.6 Exploration-exploitation dilemma

The decision-maker may acquire information about his environment through an active experimentation process. Formally, he is involved in an infinitely repeated decision process. In each period, a state of nature is drawn in a time-independent way, but according to a stationary probability distribution. The decision-maker is uncertain both about the environment law (about states) and about the consequence law (about the relation of consequences to states and actions). The standard case considers that he is uncertain about a parameter of the probability law on states. Initially, the decision-maker is endowed with a (second-order) probability distribution about that parameter. In each period, the decision-maker performs an action which gives him some utility and provides some partial information about the state of nature. He uses that endogenous information to revise his belief about the parameter.

The decision-maker has to choose between two attitudes. He 'explores' when he performs an action which gives him helpful information for further actions, at the price of some loss of immediate utility. He 'exploits' when he uses all the available information to choose and perform his best short-term action, without seeking to obtain new relevant information. Exploration is a sort of investment in information, while exploitation is simply the consumption of information. Hence, the decision-maker faces an 'exploration-exploitation dilemma', expressed by the trade-off he realizes between more exploration and more exploitation. Of course, the conditions of the trade-off change over time. This dilemma receives an optimal solution in some specific classes of decision processes.

For instance, consider a decision-maker playing in a casino with a 'two-armed bandit'. Each arm corresponds to a possible action leading to a random result conditional on a state of nature (in fact a lottery). Moreover, each arm is characterized by a fixed probability distribution over the states. The decision-maker knows the structural form of that distribution (normal, Bernoullian), but not its parameters. Nevertheless, he is endowed with a probability distribution over the unknown parameters (for each arm). He chooses one arm in each period and observes the result obtained by his action, hence the state of nature univocally associated with it. His overall choice rule is the maximization of intertemporal expected payoff with a certain discount factor.

The problem can be solved by backward induction and its optimal solution is given by a deterministic and myopic rule, the Gittins rule (Gittins, 1989). Each arm is endowed with a 'Gittins index', depending only on the sequence of its past performances. In each period, the decision-maker revises the index according to the last result and chooses the arm with the highest index. This rule leads with positive probability to the use of only one arm after a certain time, in other words the abandonment of exploration in favor of pure exploitation. But as the process is highly path-dependent, there is some (small) probability that the chosen arm is the wrong one. This probability decreases with the discount factor and tends to zero when the discount factor tends to one. In this last case, the cost of exploration becomes very low in relation to the loss of utility, and the decision-maker explores for a long time before exploiting.

The Gittins index can be computed on the basis of the probability distribution over the states, but its expression is very complex. However, it may be asymptotically approximated, for probability distributions of finite variance, by an index related to the normal law. For the normal law itself, an approximate value for each arm can be expressed by the sum of two terms. The first term, reflecting the 'exploitation value' of the arm, equals the average result already obtained. The second term, reflecting the 'exploration value' of the arm, is proportional to the standard deviation of past results and inversely proportional to the number of trials conducted; moreover, it increases with the discount factor and tends to infinity when the discount factor tends to one.

In practice, since the optimal trade-off is generally out of reach for the decision-maker (and even for the modeler), he contents himself with a reasonable and pragmatic trade-off. In fact, a trade-off between exploration and exploitation is implicit in any heuristic choice rule (especially in any learning rule) acting sequentially on an infinite horizon (see 4.8). In order to be successful, the selected choice rule has to realize a great deal of exploration at the beginning of the process and a great deal of exploitation at the end. Moreover, some degree of exploration should be maintained until the end of the process, where it should decrease asymptotically when achieving a stable (equilibrium) state.

In the surgery example, consider that the surgeon has two operation modes, A and B. Mode A (respectively B) gives a favorable result with probability p (respectively q) and an unfavorable result with the complementary probability. Even if p and q are unknown to him, the surgeon believes that they are normally distributed. When repeating the operation on a fuzzy horizon of 50 periods on average, he reasons as if his discount rate were equal to 0.98. Assume that he has already performed 26 operations with the following results. The first mode was used 20 times and gave a positive result in 15 cases, while the second mode was used 6 times and gave a positive result only once. It can be verified that the two operation modes have approximately the same Gittins index. The first mode has been tried many times and its frequency of success becomes close to its true probability (by the law of large numbers). The second mode has rarely been used and may turn out to be better than indicated by the partial results already obtained.

4.7 Bounded rationality in dynamics

Bounded rationality becomes even more relevant in a dynamic setting where the decision-maker faces intertemporal choice trade-offs. Here, he is generally considered as ‘myopic’, since he takes his decisions separately in each period according to local choice determinants. Myopic behavior is justified by the fact that the decision-maker is unable to predict the long-term effects of his actions or assumes that the long-term effects are similar to the short-term ones. Conversely, bounded rationality is compensated for in a dynamic setting by the recurrent and cumulative work of experience. More precisely, the successive periods of action play the same role in the decision-maker’s choices as the hierarchical levels of reasoning. In other respects, bounded cognitive rationality can be compensated for by structures that are designed and materialized in the environment to help in a given decision (‘situated rationality’).

As concerns instrumental rationality, the decision-maker no longer optimizes, but contents himself with weaker models. For instance, he may apply a ‘dynamic satisficing model’. He always chooses, in each period, the first action which satisfies aspiration levels on partial criteria. But the aspiration levels are now adapted from period to period according to the number of actions tested in the ongoing period. When the satisficing action is easily obtained, the aspiration levels are increased; when the satisficing action is hard to obtain, the aspiration levels are decreased. As another example, the decision-maker may follow the ‘mimetic model’ which is dynamic in nature. Facing nature in the same way as other decision-makers, he may imitate their actions if they obtain better results than he does.

Within the context of cognitive rationality, the decision-maker faces cognitive constraints in treating his past information. For instance, he may consider different choice situations as equivalent and group them together in an ‘analogy set’ by means of a ‘similarity index’. In particular, for a given decision process, the decision-maker may consider several nodes of the game tree (where he has the same actions at his disposal) to be alike. This principle is precisely applied in the case-based decision model in order to compare past situations. Alternatively, the decision-maker may borrow taxonomies from other agents facing a similar situation. Remarkably, the case-based model can be considered as a kind ofimitation by a decision-maker of his own past actions.

The main way for a decision-maker to deal with bounded rationality is by following a ‘learning process’. Learning is understood as the ability of a decision-maker to modify his behavior, in the light of past experience, so as to improve his performances. In decision theory, learning is a spontaneous process induced by natural observations, without external regulation. It differs from learning in Artificial Intelligence, which may be directed by a guide proposing chosen pieces of information (supervised learning). In any case, learning goes beyond just gathering new information on past states or past payoffs and adapting to it. It involves the structural change of some individual characteristics by a process involving abductive aspects (categorizing, analogical reasoning, pattern recognition).

For instance, the choice determinants are no longer assumed to be given, but are progressively revealed or constructed by the decision-maker in the learning process. As concerns his opportunities, he learns that new actions are available while others prove to be unfeasible. As concerns his representations, he becomes aware of certain factors acting on his environment or of certain relations governing the evolution of the environment. As concerns his preferences, he discovers that he is more sensitive to some effects on his environment and less sensitive to others. Likely, some rules followed by the decision-maker are modified in their analytical form or through a parameter according to past experience. This can be applied to the belief revision rule, the forecasting rule or the overall behavior rule.

It is possible to define several levels of learning, in which more and more profound elements are modified over longer and longer time scales. Primary learning may be concerned with the adaptation adaptation of a classical structure to a new context (simple loop). Secondary learning may be concerned with the finding of new structures which may adapt to a larger set of contexts (double loop). Many models of learning have been proposed and are, as usual, justified by theoretical, praxeological or empirical arguments. In general, learning can only be efficient if the decision-maker has sufficient degrees of freedom and if the environment is sufficiently stable. But learning may involve various cognitive capacities for the decision-maker. Especially, it is ‘cognitive’ when it concerns essentially a transformation of beliefs and ‘enactive’ when it concerns directly an adaptation of actions.

In the surgery example, consider that the surgeon has three equiprobable types of patient, a, b and c. He may perform three modes of operation, A, B and C. Mode A succeeds very well with type a (utility +8), succeeds moderately with type b (utility +5) and fails with type c (utility −9). Mode B succeeds moderately with type a (utility +5), succeeds well with type b (utility +7) and fails with type c (utility −9). In mode C, no operation is carried out and the utility obtained is 0 in all *cases. Such structural information is perfectly known by the modeler, but only partially by the decision-maker. Different learning models can be applied to this problem (see 4.8).*

4.8 Learning models

In ‘epistemic learning’, the first category of learning models, the main principle of learning is ‘belief revision’. The decision-maker observes the past states of nature. He has bounded cognitive rationality, since he computes statistics of the past states which he assumes to be stable in the future. He has strong instrumental rationality, since he optimizes his present action according to his revised beliefs about nature, but he is partially myopic. Exploitation is usually performed in keeping with the maximization principle. Exploration is not initially present, but when needed (for instance when the actor thinks that the state depends on his action), a random mechanism is introduced so as to try other actions from time to time. The process frequently converges towards the expected utility maximizing action, since the belief on nature converges to the truth, according to the law of large numbers.

In particular, the ‘fictitious play model’ considers that the decision-maker observes the past states and computes their frequencies. In the initial period, the frequency is conventional and generally considered to be uniformly distributed. Under a stationarity assumption, the decision-maker assumes that the past frequency will coincide with the future probability. He chooses the action which maximizes the expected utility under that expectation. In one variant, the ϵ-greedy rule, he chooses with probability $1-\epsilon$ the optimizing action and with probability ϵ a random action. In another variant, known as ‘smooth fictitious play’, he chooses an action with a probability which is proportional to the expected utility (probabilistic model).

In ‘behavioral learning’, the second category of learning models, the main learning principle is ‘action reinforcement’. The decision-maker no longer observes the past states (he may not even be aware that he is facing a random environment), but only the results of his past actions. He has weaker cognitive rationality, since he no longer predicts anything, simply computing an ‘index’ summarizing the past performance of each action, which he assumes will remain stable (hence postulating the stationarity of the environment law and consequence law). He also has bounded instrumental rationality, since he simply reinforces the actions with good results and inhibits the actions with bad ones. Exploration and exploitation are directly integrated into the choice rule, since reinforcing an action means choosing it more often without abandoning the others. Here again, the process frequently converges towards the expected utility maximizing action.

In particular, the 'basic reinforcement rule' assumes that the decision-maker observes the past utility obtained with each action. He computes an index for each action, which is nothing other than the cumulative utility it has obtained in the past. At the beginning, the indices have conventional values, generally equal. He chooses an action with a probability which is strictly proportional to its index. Here, due to positive feedback acting on best actions, exploration progressively decreases and converges asymptotically to zero. Alternative models introduce an aspiration level on the (synthetic) utility index, which increases when the level is achieved and decreases when it is not. The decision-maker keeps to his previous action when the aspiration level is reached, otherwise he changes.

In an 'evolutionary process', the third category of learning models, the main principle is 'survival of the fittest'. There is now a population of actors, gathered in subpopulations of decision-makers playing the same action (or strategy). The actors no longer have either cognitive rationality, since they observe nothing, or instrumental rationality, since they always play the same action. But a selection principle states that the actors reproduce according to the utility they obtain in their interaction with nature. Hence, utility here plays the role played by fitness in biology. If the selection principle ensures exploitation, a mutation principle ensures exploration. The mutation principle states that mutating actors are randomly introduced into the population. mutations may be infinitesimal or considerable, regular or decreasing over time. Here again, the (non-stochastic) process generally converges towards a homogenous population playing the maximizing action.

One particular rule, the 'replicator rule', considers a population of actors interacting with nature in each period. The population can be split in subpopulations of actors, acting as automata since all of them always play the same action. Each actor obtains a utility and reproduces in proportion to that utility compared to the average utility. Hence, a subpopulation obtaining a high average utility will see its proportion of the population grow. A variant, the stochastic replicator, considers moreover that a given proportion of new actors are introduced in each period or that existing actors change their actions. However, it is always possible to associate a deterministic process with the corresponding stochastic one, even if their convergence results may differ.

In the surgery example just described (4.7), consider what happens after 12 operations. In epistemic learning, the observed frequencies of types a, b and c are 3/12, 5/12 and 4/12 respectively. According to the fictitious play rule, the expected utility of modes A, B and C are 11/12, 14/12 and 0 respectively. Hence, the surgeon provisionally chooses B; but he will play A when the frequencies approach the (equal) probabilities. In behavioral learning, A was employed 6 times with 2 full successes, 2 moderate successes and 2 failures while B was employed 6 times with 2 full successes, 3 moderate successes and 1 failure. According to the CPR rule, the indices of A, B and C are 8, 20 and 0 respectively, and the surgeon chooses A with probability 2/7 and B with probability 5/7; but with more experience, he will choose A. In an evolutionary process, two subpopulations of surgeons practice modes A and B respectively. Since their numbers of disciples (or patients) vary in proportion to their results, all surgeons end up using the same mode after a certain time.

5

Coordination of players through beliefs

Chance is often
others will.
A. Capus

According to a 'minimal sociology', classical game theory assumes that the players, each in an environment comprising the other players and nature, are coordinated on some equilibrium. An equilibrium state, implicitly achieved by a 'Nash regulator', represents a fixed point on the loop relating the players' crossed expectations about their respective actions. In theory, such a state is achieved constructively by an eductive process, through which the players, endowed with strong cognitive abilities, simulate one another's behavior. In practice, constrained by informational and computational limitations, the players' actions are more likely to be coordinated according to concepts of bounded rationality equilibrium.

In game theory, the players' cognitive rationality has to be adapted to a strategic situation; it leads to an equilibrium state (5.1), the Nash equilibrium acting as a reference (5.2). The players face a new form of uncertainty about their opponents' determinants (5.3) and coordinate through an adapted concept of equilibrium, namely Bayesian equilibrium (5.4). They try to achieve and essentially to select an equilibrium by means of reasoning processes based on conventions (5.5), which give rise to specific concepts of equilibrium (5.6). They face even stronger limitations on their cognitive capacities due to more complex calculations (5.7), hence achieving even weaker forms of equilibrium (5.8).

5.1 Strategic rationality and equilibrium

Actors' interactions are based on a '*separability principle*' which asserts that the different players and their common environment behave in an independent way. This principle can be made more precise thanks to two postulates. The 'isolationist postulate' states that the players hold no permanent relations expressed as prior links, but only occasional relations expressed as transient actions. The 'atomistic postulate' states that the players hold personal determinants uninfluenced by other players' or collective variables, and therefore have no common means, no correlated beliefs, no normative preferences. The separability principle reinforces '*methodological individualism*', since all social facts are assumed to result from the conjunction of independent individual actions. Nevertheless, the players act in a 'strategic context', since their actions have common consequences, but they face them independently.

In game theory, a social system is assumed to consist of two rather than three types of entities: the 'players' and their common 'physical environment'. There is no 'institutional environment', since each player chooses actions which influence the other players directly, without intermediation. In fact, some 'rules of the game' may be introduced, as it is the case for 'parlor games'. But these rules are directly incorporated into the players' determinants, as it is assumed that the players necessarily follow them. They modify the players' opportunities by allowing some actions and forbidding other ones (even if the constraints can be violated). They act on players' beliefs by specifying the outcome for the players of a certain combination of moves. They influence the players' preferences by defining bonuses or sanctions associated with the outcomes.

Players' interactions are, moreover, based on a '*coordination principle*', which asserts that their actions are adjusted by some external device. This principle can be made more precise by the introduction of two postulates. The 'strategic rationality postulate' states that the players continue to act in a rational way, but now with reference to the other players' expected actions. The 'compatibility postulate' states that a virtual entity, called the 'Nash regulator', achieves coordination between the players' actions. If such a coordination process is achieved, it leads to an 'equilibrium state', generally defined as a stationary state in the absence of external perturbations. Once more, the players are not assumed to define a collective plan of action, but each defines an individual plan of action, conditional on those of the other players. The plans of action are causally independent, even if they are obtained through a related deliberation process and influenced by an external regulator.

More precisely, an equilibrium state is defined by three main conditions. Firstly, each player is endowed with instrumental rationality, and therefore adapts his means to his objectives in the light of his expectations about others' actions, called 'conjectures'. Secondly, each player is endowed with cognitive rationality, and therefore adapts his beliefs about others' actions and determinants to his observations. Thirdly, the regulator gives some information to each player in order to ensure the coincidence of players' predicted and realized actions. More precisely, the regulator closes the loop which relates each player's actions to the predicted actions of the others. An equilibrium state appears as a fixed point on that loop; it is 'self-fulfilling' because it induces the realization of the expected actions. It appears simultaneously as an 'action equilibrium' (each action responds to the others' actions) and a 'belief equilibrium' (each conjecture is adapted to the others' conjectures).

The strategic rationality of each player is just an extension to an active environment (constituted of other players) of his usual rationality in a passive environment (symbolized by nature). More precisely, each player is assumed to 'naturalize' his opponent in the sense that he considers his actions as mere states of nature. Even if he knows from his own experience that the others' behavior is strategically and endogenously defined, he treats it as an exogenous variable. As concerns instrumental rationality, the player treats his expectation of the others' actions as if these actions were fixed. As concerns cognitive rationality, his expectation of the others' actions is computed from observations as if these actions were objective.

The framework provided by the game model is too general to be applied as such to a strategic problem. It needs more precise specification of the players involved and of the entity which links up their behaviors. For specific 'contexts' relative to the material and social environment, the game model induces some well-defined 'equilibrium concepts'. Each equilibrium concept is supported by analytical solutions in the form of 'equilibrium states'. An equilibrium state may not exist (co-determination problem) or, on the contrary, be multiple (co-selection problem). Any equilibrium concept can, moreover, be justified (or conversely, be criticized) by one of three types of argument. An 'epistemic justification' supports it in terms of thorough reasoning operated by the players alone. An 'evolutionary justification' shows its achievement by some learning or evolutionary process. An 'empirical justification' simply stresses that it leads to actions consistent with given observations.

Consider two classical games which both involve two drivers on a road. Firstly, the 'crossroads game' (analogous to the 'battle of the sexes') is a symmetric game where the drivers, arriving at the same time at a crossroads, can either keep going or stop. The consequences are physical (material damages) or psychological (the vexation of being the only one to stop). The preferences of each driver are naturally ranked: keep going when the other stops (utility 3), stop when the other stops (utility 2), stop when the other keeps going (utility 1), keep going when the other keeps going (utility 0). Secondly, the 'driving side game' (analogous to the 'meeting point game') is a symmetric game where the drivers may drive on the right or the left. The consequences are material: an accident if one drives on the left and the other on the right, no accident otherwise. The preferences are obviously ranked: driving on the same side as the other (utility 2), driving on the opposite side to the other (utility 0).

5.2 Nash equilibrium

Formally, a static game is represented by its 'normal form' (or 'strategic form'). Each player has a set of possible actions, which may be finite or infinite. The combination of actions for each player defines an 'action profile'. Each player is endowed with a utility function depending on the consequences of an action profile. The combination of the utility of each player for an action profile defines a game outcome. A two-player finite game is represented by a 'game bi-matrix' in which the rows and columns list the possible actions of the players. The two values in each cell represent the utility obtained from the outcome by each player, hence depends for each player on all players' actions. To be more explicit, a player chooses between 'strategies'. A 'pure strategy' is just an action while a 'mixed strategy' is a probability distribution on actions. Mixed strategies are never relevant in decision-making under uncertainty, but they do become useful in games.

The basic equilibrium concept, 'Nash equilibrium', is defined by three conditions. First, each player forms conjectures about the others' actions. Second, each player computes his best response to the others' expected actions. Third, the coincidence of expectations and realizations is achieved by the Nash regulator, which then announces their common value. Hence, a Nash equilibrium state appears as a fixed point of the best response functions of all players. This fixed point already closes the loop between crossed expectations on the second level. In pure strategies, a Nash equilibrium state may not always exist or may be multiple. In mixed strategies, a Nash equilibrium state exists for any finite game (the uncertainty of players' moves favors their coordination). The Nash equilibrium concept, based on a 'best response' point of view, is usually weakened under two forms, next described.

A 'rationalizable equilibrium' is a game state obtained by iterated elimination of inferior strategies. An inferior strategy is a strategy which is never a best response to the opponents' strategies. For a two-player game, an inferior strategy coincides with a strongly dominated strategy, i.e. a strategy which is such that another strategy exists which is better for all players. In order to compute a rationalizable equilibrium state, the inferior strategies are successively eliminated by alternating the players and any remaining outcome forms an equilibrium state. A Nash equilibrium is always rationalizable, since it is formed of best responses to others' strategies, themselves considered to be best responses. Conversely, when a rationalizable equilibrium is unique, it is the unique Nash equilibrium.

A 'correlated equilibrium' is obtained by means of a fictitious entity, the 'correlator' (in fact, a clone of the Nash regulator). The correlator chooses an outcome according to a given probability distribution and suggests to the players the play of the corresponding actions. An equilibrium state (constituted of a probability distribution on all outcomes) is obtained when it is in a player's interest to follow the suggestion of the correlator when the other players follow it. A (mixed) Nash equilibrium is a correlated equilibrium for which the probability of an outcome is the product of the probabilities of the corresponding actions.

The 'epistemic justifications' of the preceding equilibrium concepts are founded at least on two strong assumptions: common knowledge of the game structure, and common knowledge of the players' (Bayesian) rationality. Further, when postulating that it is common knowledge that the players play independently (their conjectures are independent probability distributions), one obtains the rationalizable equilibrium. When postulating, alternatively, that the players' beliefs are precoordinated (their conjectures are derived from the same prior probability distribution), one obtains the correlated equilibrium. However, adopting these two additional assumptions together is not enough to get the Nash equilibrium; more heroic assumptions concerning the players' conjectures are needed. For two players, one has to assume that the conjectures are shared knowledge. For more players, one has to assume that the conjectures are common knowledge (since the conjectures of two players about a third player have to be similar).

The 'evolutionary justifications' of the preceding equilibrium concepts are more diverse, since they have to be sustained by learning or evolution processes followed by the players (see 6.5). In a repeated game, each player uses an explicit choice rule in each period, which expresses his bounded rationality and is conditioned by his past observations. The process may or not converge toward some asymptotic state for a given convergence criterion. The asymptotic state is then simply interpreted as an equilibrium state. It can be proven that the process converges rather easily to some Nash equilibrium state in pure strategies, but converges less easily towards a Nash equilibrium state in mixed strategies.

In the crossroads game, there are two pure Nash equilibrium states, where one driver keeps going while the other stops. A (strong) co-selection problem is involved: one equilibrium state favors the first driver while the other favors the second one. In addition, there is a mixed Nash equilibrium, each player keeping going with probability 1/2. Each outcome is rationalizable, since all actions are; for instance, a driver keeps going because he thinks the other will stop, and he thinks that the other stops since he thinks that the other thinks that he will keep going. Correlated equilibrium states are obtained for specific probability distributions on outcomes; for instance, each driver alone keeps going with probability 3/7, both keep going with probability 0 and both stop with probability 1/7, an equilibrium which is materialized by traffic lights. In the driving side game, there are also two pure Nash equilibrium states, where the drivers both drive on the left or both drive on the right. A (weak) co-selection problem occurs, since the two equilibrium states are utility-equivalent for the drivers.

5.3 Informational limitations

Each player now has three sources of uncertainty, each concerning one of the three entities considered by the modeler. 'contextual uncertainty' concerns his 'passive environment', shared with the other players and summarized in nature. He has fuzzy beliefs about the generation law of states and about the relation of states to consequences. 'Actorial uncertainty' concerns his 'active environment', hence the set of his opponents. He is unaware of their determinants and is unable to simulate the way their actions are chosen. 'Personal ' concerns himself. He does not know his own determinants or even the way in which he chooses an action. However, like in individual decision-making (see 3.3), a general trend consists in reducing any form of to uncertainty on nature.

As concerns actorial uncertainty, each player is assumed to summarize another's determinants in a one-dimensional variable called the player's 'type' (Harsanyi, 1967). A 'reduction' operation is at work here, since a whole structure is transformed into a unique variable. Moreover, the other's type is considered as fixed by nature and is therefore assimilated to a state of nature. A 'naturalization process' is again at work, since the player's determinants are considered as exogenous factors. A type may be defined for each of the three determinants, a capacity type, a doxastic type and a deontic type respectively. But the determinants themselves are progressively reduced to the doxastic type. Opportunities are reduced to preferences when an unavailable action is just considered as having an infinite cost. preferences (and opportunities) are reduced to representations, since they act not directly but through the players beliefs about them. The doxastic type is finally formed of a hierarchy of crossed beliefs, as studied earlier (see 1.7).

As concerns personal uncertainty, the player considers himself in the same way he considers other players. His own determinants (and his own rationality) are summarized in his own type, since he may be uncertain about them. As concerns contextual uncertainty, it is frequently considered that uncertainty about the environment law and about the consequence law can be attributed to some parameter of the analytical relation. This parameter is further treated like a state of nature. Coming back to actorial uncertainty, uncertainty may be directly affected to the choice rule followed by another player, in a similar way as for nature. However, it is not easy to define a set of choice rules which may be followed by a player (optimizing, satisficing, drawing lots, etc.).

The players are further assumed to express all forms of subjective uncertainty they face in a set-theoretic or a probabilistic way. In particular, for actorial uncertainty, it is assumed that each player has a (subjective) probability distribution over the players' types. If the players' types are assumed to be independent, such a distribution is broken down into a probability distribution over the type of each opponent. As concerns personal uncertainty, a player may have a probability distribution over his own possible types. In particular, the probabilistic choice model is subject to one interpretation where a player is uncertain about his own deontic type. More precisely, he holds a specific probability distribution over his utility function (see 3.8).

Finally, the players' beliefs on states (or types) are independent or correlated. The 'Harsanyi doctrine' specifies that the players have a common prior belief about an uncertain structure and receive private information about it. This ‘common prior assumption’, which assumes that the difference of beliefs between players stems only from their different information, is difficult to justify. It essentially induces that the players are pre-coordinated in their initial beliefs. More precisely, such a common belief is generally probabilistic (prior probability distribution), while the private information remains set-theoretic (information partitions). Such an assumption is specifically made for actorial uncertainty (Aumann, 1999). There is a prior probability distribution over all types and each player has privileged information on his own type.

A common assumption made about players’ beliefs is the veridicity axiom. If the players may be fuzzy about the game they are playing, they are nevertheless considered not to be wrong about it. However, it is possible to consider that the players have false beliefs of different kinds, and that as a consequence, they are subjectively playing different games. For instance, a player may be wrong about the other’s type and consider that his opponent has less information, different beliefs or preferences or even weaker rationality than he really has. However, each player may become aware of these errors if he makes unexpected observations about the other’s actions and is induced to correct his beliefs accordingly through a learning process (see 6.7).

In the crossroads game, each player may have some uncertainty about the other’s utility function. Players’ types are therefore introduced according to two alternative configurations. In configuration A, a ‘cautious’ driver has the usual utilities while a ‘go-getter’ driver adds a utility of 2 when he keeps going. Moreover, each driver has a probability p *of being cautious and a probability* $1 - p$ *of being a go-getter, and this is known by both drivers. It can be noted that the action of keeping going becomes a dominant action for the go-getter driver. In configuration B, the first driver has (continuous) uncertainty about the other’s utility. More precisely, driver* 1 *considers that the utility obtained by driver* 2 *when the last keeps going and himself stops is in fact uniformly distributed in some interval* $[2 - e, 2 + e]$.

5.4 Equilibrium under uncertainty

Consider first a situation in which the only source of uncertainty is actorial and deals with others’ determinants. Each player is endowed with a probability distribution on all players’ types, but he knows his own type. The preferences of each player depend on the types of all the players, since he has to internalize their deontic types. A player’s strategy is now a 'conditional strategy', i.e. a strategy conditional to the player's own type. The strategies considered are pure rather than mixed. Of course, when given a conditional strategy, each player knows his own action, but the others are uncertain about it. In the standard case, the is formalized by a prior probability distribution on the players' types. Moreover, the player is assumed to be Bayesian rational when facing such an .

The 'Bayesian equilibrium' concept is nothing other than the Nash equilibrium concept applied to such an extended game. Each conditional strategy of one player is the best response to the conditional strategy (at equilibrium) of the others. It satisfies the same existence and unicity conditions as the Nash equilibrium. It admits the same epistemic justifications as the Nash equilibrium. Each player simulates what a player (including himself) would do if he were of each possible type. This constitutes counterfactual reasoning for all types which are not the real, existing ones. Practically, it is possible to consider that each player is divided into as much clones as his possible types in other's eyes.

The Bayesian equilibrium concept allows a nice interpretation of the 'mixed Nash equilibrium' concept (Harsanyi, 1973). A player's mixed strategy can be seen as a conjecture made by the other players about that very player. More precisely, consider a game where a player has uncertainty about the other's utilities, translated into a probability distribution over the other's type. A Bayesian equilibrium state consists in a (pure) conditional strategy for each player. But a pure conditional strategy defines an action for each type, hence a probability distribution on actions with regard to the type. When actorial uncertainty tends to zero (the player tends to his actual type), the conditional strategy survives and continues to induce a probabilistic action.

Even without actorial uncertainty on other's determinants, each player faces a new form of uncertainty about the other's present action which is called 'strategic uncertainty'. Each player may deal with such uncertainty by assessing a probability distribution on the other's action. Such a probability is precisely defined by considering a (mixed strategy) equilibrium state. However, many such Nash equilibrium states are generally available. A specific selection rule is based on the risk faced by the player when deviating from that equilibrium state. More precisely, a 'risk-dominant equilibrium' is defined as the less risky equilibrium state in a precise technical sense (Harsanyi, Selten). Well defined for 2x2 games (two players, two actions per player), the risk-dominance concept is difficult to generalize to any class of games.

A more complex equilibrium is obtained when the players express the strategic uncertainty about their opponent as non-additive probabilities. In particular, if a player plays an action with a certain probability, the other plays a best response to a subjectively distorted probability distribution expressing his risk-aversion (rank dependent utility choice rule, see 3.4). A 'rank-dependent equilibrium' state is obtained in the spirit of a Nash equilibrium, when each probability distribution of a player is a best (distorted) response to the others probability distribution. As a degenerate case, an equilibrium state may even be defined without a reaction loop between the players, each player playing independently. For instance, a 'uniform equilibrium' state is obtained when each player chooses the best response to the uniform probability distribution attributed to the other's action.

Similarly, less popular equilibrium concepts are obtained when the players express strategic uncertainty in a set-theoretic rather than a probabilistic probability framework. For instance, each player considers any opponent's action as possible, computes the minimal payoff for each action, then chooses the action with the highest minimal payoff (maximin choice rule, see 3.4). A 'cautious equilibrium' state is obtained when each player acts in this way, the players in fact playing independently. However, such an equilibrium concept can only be justified when each player considers that his opponent is completely opposed to him. This is actually true only in zero-sum games, where the players have strictly opposing interests. In that case, if the cautious strategies give opposing utilities to the players, a cautious equilibrium coincides with a Nash equilibrium.

In configuration A of the crossroads game, there is a symmetrical Bayesian equilibrium state when p is less than 1/2: the player stops if he is cautious and keeps going if he is a go-getter. An asymmetrical equilibrium is obtained when p is greater than 1/2: one player always keeps going, the other stops if he is cautious and keeps going if he is a go-getter. In configuration B of the crossroads game, in a Bayesian equilibrium, each player adopts the following conditional strategy: he keeps going if the parameter e is below some precise threshold and stops if it is above. When e tends to zero, one obtains the mixed Nash equilibrium for which each driver keeps going with probability 1/2. As concerns the cautious equilibrium, both drivers stop.

5.5 Cognitive effects

Empirically, some of the general assertions underlying the separability principle and the principle introduced earlier (see 5.1) appear to be too strong. The agents are in fact more closely linked than usually stated, since they all operate within a common cultural environment. Their coordination is based on all the means at their disposal, especially their common background. Here again, psychologists have described several phenomena which seem to violate the basic principles. Some of these can be incorporated just by reinterpreting the usual framework, but others call for a more or less profound transformation of the framework.

A first phenomenon concerns the existence of solid links between the players' determinants, due to some form of 'social conditioning'. As concerns their opportunities, the players internalize certain social constraints which act conjointly on them. As concerns their representations, the players share certain common beliefs acquired through education and from the media. As concerns their preferences, the players are influenced by the same social norms, especially norms of reciprocity or fairness. But the usual equilibrium concepts still remain relevant since the players move independently. Taking one further step in the direction leads us to consider that the players may act as a 'team' and perform 'joint' actions. But even in this case, it is possible to consider that they keep to individual actions and simply have an individual incentive to cooperate.

A second phenomenon concerns the fact that the players pursue not only their individual interests, but have interpersonal concerns. This is taken into account by different interpretations of the player's utility function, which depends both on his own action and on his opponent's one. In the usual interpretation, the two actions lead to common outcomes and the utility of each player just depends on these joint consequences. In a second interpretation, each player is sensitive not only to his own improved outcome, but also to the improved consequences for the other player. Plainly, he may judge his own consequences with reference to the other player's ones, and hence reason in terms of relative consequences rather than absolute ones. More profoundly, he may be altruistic and take into account the other player's consequences even when they are not related to his own. In a third interpretation, each player is directly sensitive to the other's utility. He is benevolent or malevolent according to whether he is satisfied or dissatisfied when the other player obtains an increased utility.

A third phenomenon concerns the perception by the players of their social and natural context. Here again, a 'framing effect' influences the categorization of the material environment, of the rules of the game as well as of the other players' characteristics. In a given game, for example, a player may consider some configurations as similar and he therefore applies the same type of strategy to both (Jéhiel, 2005). Likewise, a player may perceive some games as similar, and he transfers therefore the same type of solution from one to the other. Finally, a player may treat some players as similar in various games and configurations, and he develops therefore the same types of reaction against them. As a matter of fact, such 'framing' is highly language-dependent, in the sense that the type of description of a situation provided by the modeler modifies its representation by the player.

A fourth phenomenon concerns the fact that the players may use cultural information which lies outside the definition of the game in order to choose their actions and coordinate on an equilibrium state. The best example is the resolution of the problem of selection of an equilibrium state in the case of multiplicity. Certain 'conventions', considered to be shared among the players, can act as selection rules at two levels. First-order conventions select salient states according to the context, such 'focal states' resulting from the players' background culture (Schelling, 1960). Second-order conventions define conditions imposed on the equilibrium states, such as symmetry, Pareto-optimality or simplicity. But these conventions are arbitrarily stated outside the game structure.

In other respects, the logification process followed by the modeler is extended from decision-making to equilibrium concepts. The physical laws, the player's characteristics and the rules of the game are defined as (non-independent) formulae. An equilibrium state just appears as a theorem of such a formal system. Logification helps to reveal hidden assumptions, as is the case when considering the epistemic justifications of an equilibrium concept. It also helps in defining the complexity of the computation process of an equilibrium, from the modeler's point of view (see 5.7). But it has not yet made it possible to generate new conditions for the existence and multiplicity of equilibrium states, even less to suggest other equilibrium concepts.

In the crossroads game, the driver may select an equilibrium state by relying on certain social norms. Such norms are imposed to a 'labeled' driver, i.e. a driver which is placed in a commonly observable situation. Some norms are rather informal, such as a male driver who gives way to a female one. Some norms are more formal, such as traffic signals or priority to the right. There are also norms at work in the driving side game. Driving on the right seems to have been developed by convenience *for armed people riding horses. However, even if drivers have a common interest in adopting a convention, whatever it is, it may not be easy to get it universally accepted. For instance, if we consider the most populated countries, China drives on the right while India drives on the left.*

5.6 Contextual equilibrium concepts

A first equilibrium concept is the 'cognitive hierarchy equilibrium' introduced by Camerer (Camerer, Ho). Each player adopts crossed expectations about another's action up to a maximal level (specific to each individual). Moreover, he assumes (wrongly) that any opponent forms his expectations to a lower maximal level than he does. Each player reasoning at level n is even endowed with a probability distribution about the maximal levels of the other players. Finally, he simply plays his best response to this hierarchical belief. At an equilibrium state, these players' beliefs are assumed to be fulfilled, in the sense that the hierarchical belief considered by each player is the right one, however truncated at his own maximal level. A cognitive hierarchy equilibrium coincides with a Nash equilibrium when the maximal level is one for all players.

A second equilibrium concept is the 'fairness equilibrium' introduced by Rabin (Rabin, 1993). Each player forms expectations at two levels, i.e. expectations about the other's action and expectations about the other's expectations about his own actions. Moreover, these expectations are arguments of his utility function in a special way. They express the fact that a player has a sense of fairness regarding the other and is sensitive to the sense of fairness the other has regarding himself. More precisely, a fairness function reflects how a player takes into account the material utility of the other player. Finally, each player maximizes his expected utility with the enlarged utility function. An equilibrium state is obtained when all expectations are realized. A fairness equilibrium coincides with a Nash equilibrium when the material outcomes are high enough with reference to the ethical ones.

In other respects, equilibrium concepts may be defined when each player behaves according to some 'mimetic model' (Orléan, 1998). On one hand, imitation behavior can be grounded on observation of actions or moreover of payoffs. In 'action mimetism', a player simply imitates the opponents' past actions. In 'payoff mimetism', a player imitates the past actions of opponents performing better than him. On the other hand, imitation behavior can be sustained by beliefs or preferences.

‘Preferential mimetism’ is simply founded on the fact that the conjunction of similar actions gives better payoffs than dissimilar actions. ‘informational mimetism’ is more subtly founded on the fact that a player thinks that the other knows more than he does. In fact, mimetic behavior appears as a purely reactive behavior and is more relevant for learning models where the players can observe their performances and adapt to them.

All these equilibrium concepts were introduced to improve the satisfaction of certain basic empirical requirements. They are in accordance with many stylized facts observed by psychologists. But they are not really tested or even testable in laboratory experiments. The main difficulty is that they are under-specified, so they lead to equilibrium states which are consistent with a large spectrum of data. The cognitive hierarchy equilibrium introduces a probability distribution on the levels of expectation which can take any form. The fairness equilibrium introduces a fairness function and a utility function with many unspecified parameters. The mimetic equilibrium is even not well specified and can be expressed in different ways.

These equilibrium concepts seldom receive profound eductive or evolutionary justifications. Only the cognitive hierarchy model is sustained by some general principles. In fact, they appear rather *ad hoc*, since they introduce auxiliary functions which are very particular and difficult to compute for the player (and even the modeler). In the cognitive hierarchy equilibrium, it is not clear how the probability distribution on expectation levels can be evaluated, since the levels of reasoning are hard to observe. In the fairness equilibrium, both the fairness function and the utility function look rather arbitrary. In the mimetic equilibrium, the mimetic mechanisms do not even lead to precise proposals.

Concretely, the definition of contextual equilibria is a difficult task, since the modeler has to integrate two (related) types of concerns for any player. The ‘cognitive concern’ is related to the way a player represents his material and social environment. The ‘social concern’ is related to the way a player takes into account the effect of his actions on the other players. These concerns were first introduced as a ‘cognitive bias’ and a ‘social bias’ on the standard model, but they now appear as fundamental concerns. Moreover, the modeler has to deal with these concerns at two (related) levels. He has to state how the players construct and assess their material and social environment, and he has to state how each player makes the same assessment for the other players. Once more, it is assumed that each player reasons like the modeler, even if his cognitive capacities are weaker (see 5.7).

A third game involving driving can be considered. The 'car type game' (analogous to the classical 'stag hunt game') is a symmetric game where two drivers buy either a gas car or an electric car. An electric car is individually more convenient than a gas car, but it is only operational when both drivers buy it. The preferences are ranked accordingly: buying an electric car when the other player does the same (utility 3), buying a gas car whatever the other player does (utility 2) and buying an electric car when the other buys a gas car (utility 0). There are two Nash equilibrium states, where the drivers either both buy gas cars or both buy electric cars. A (weak) co-selection problem is again involved: the second equilibrium state is better for both drivers than the first. But the first is risk-dominant (and cautious) since it is dangerous to be alone to buy an electric car.

5.7 Computational limitations

Consider first that the players remain coordinated on some equilibrium state by a 'Nash regulator', but are limited in their deliberation process by their cognitive abilities. Each player always defines some response to his conjecture about the other's actions, but both his conjecture and his response are computed in accordance with his bounded rationality. The players' expectations and realizations are still equalized by some external entity symbolized by the omniscient and omnipotent Nash regulator. Two approaches can be distinguished in accordance with the two contrasted views of bounded rationality proposed early on by Simon (Simon, 1982).

In the first approach, each player performs a perfect deliberation process, but he reasons on the basis of simplified determinants. As concerns his opportunities, each player chooses between restricted sets of actions. In particular, he may only consider simple and robust actions deliberately treated in isolation. As concerns his representations, each player constructs simplified beliefs on his environment. In particular, he may take into account only specific or nearby consequences of the joint actions. As concerns his preferences, each player is endowed with simplified utility functions. In particular, he may limit the number of partial criteria considered or shorten the horizon of evaluated effects.

In the second approach, each player considers his original determinants, but reasons with limited capacities of deliberation. Bounded rationality may be expressed, in the framework of epistemic logic, by the players' lack of logical omniscience. It may, in particular, be modeled by considering the player as an 'automaton' which computes the intended action with a finite number of internal states. Bounded rationality can also be directly stated through the specific expectation rules and choice rules used by the players. In the first case, weakly rational expectations or limited crossed expectations can be introduced. In the second case, the 'satisficing choice rule' or the 'probabilistic ' can be introduced.

Consider now that the players are able to choose an action with regard to their information, but it is the process coordinating their actions which is submitted to harsh epistemic constraints. Each player is able to undertake a deliberation process which follows either a strong or a bounded rationality procedure. The coordination process, which always consists in computing a fixed point of the players' conjectures and actions, is now submitted to precise computational limitations. Two contrasted approaches are again possible, depending on the entity assumed to be in charge of the coordination process, i.e. the Nash regulator or the players themselves.

In the first approach, the equilibrium state is always computed by the Nash regulator, but this entity (or equivalently the modeler) cannot deal with computations that are too complex. In computer science, several levels of complexity have been defined for problem-solving, essentially according to the time needed to reach the solution: polynomial, exponential, NP-complete. It is possible to prove that the computation of the solutions associated with a given equilibrium concept in a given class of games involves a given level of complexity. However, the complexity index is defined as an average or an upper limit for a whole class of games, and the complexity of computing some equilibrium in a specific game in the class may well be lower.

In the second approach, the equilibrium is computed by the players themselves, in the spirit of the epistemic foundations of equilibrium concepts. Since they rely on common knowledge of several structural features of the game, bounded rationality acts essentially on these beliefs. In particular, it is possible to consider that the players only have n-shared beliefs or ϵ-common beliefs of the game structure (see 1.8). Likewise, the players only have n-shared beliefs of the players' rationality or believe in an 'irrational type' for their opponents with a small probability (Kreps, Milgrom). However, if the players directly compute a fixed point of the best response functions, bounded rationality acts on their reasoning capacities. In particular, they may again be considered as machines endowed with finite internal states.

In the crossroads game, the equilibrium states seem easy to compute, since the game is a symmetric one with few actions for each player. The *pure Nash equilibrium states, in particular, are easily discovered, since the utilities are simply ordinal and their numerical values are not relevant. The mixed Nash equilibrium state is harder to compute because the utilities now have to be considered as cardinal. The correlated equilibrium states are even harder to obtain, since they correspond to a more sophisticated mechanism and form a whole continuum. Above all, the Bayesian equilibrium states are very difficult to compute, not only for the players, but even for the modeler.*

5.8 Bounded rationality equilibria

Two equilibrium concepts are obtained when each player uses one of the two basic bounded rationality choice rules. The coordination on a fixed point is still computed by the Nash regulator (or the modeler). The 'satisficing equilibrium' obtains when each player plays a satisficing action in response to the other's action. As a special case, the 'ϵ-rational equilibrium' corresponds to the case where each player plays an ϵ-rational action in response to the other's action. The 'quantal equilibrium' obtains when each player uses the probabilistic choice model (which is given a bounded rationality interpretation, see 4.7) in response to the others probabilistic action. The last equilibrium is the most commonly used, since it allows the modeler to represent the deviation of the player from strong rationality according to a unique parameter μ (which may differ from one player to another).

Another equilibrium concept, the 'machine equilibrium', is obtained when each player is able to choose a machine which implements a given strategy. The coordination of the machines is once more achieved by the Nash regulator realizing a Nash equilibrium state. The machine is characterized by its 'complexity', generally the number of internal states it can use (possibly associated with the number of transitions between states). The complexity is exogenous when the player can only choose between machines of a given maximum complexity. The utility function of the player is then restricted to the payoff induced by the equilibrium. The complexity is endogenous when the player can choose both the machine and its degree of complexity. The player then has a multicriteria (often lexicographic) utility, defined primarily by the payoff and secondarily by the cost induced by the complexity.

A more sophisticated equilibrium concept, the 'automaton equilibrium', is obtained when each player is himself an automaton computing a best response to the other automata, such an automaton having a finite number of internal states at his disposal or incurring computational costs. The coordination of the competing automata is still achieved by a Nash regulator. Each automaton may be of a more and more sophisticated type, such as a Moore automaton, a perceptron or a Turing machine. The utility function is generally restricted to the payoff induced by the equilibrium, since the trade-off between the material payoff and the computing costs may lead to a paradoxical situation (see 3.7).

These models receive some empirical justifications, since they at least assume bounded rationality for the players. But here again, they are either under- or over-specified. For the quantal equilibrium (as for the satisficing equilibrium), an additional degree of freedom is introduced which allows a greater number of equilibrium states. Hence, it renders the model less refutable, at least when the parameter μ is not specified. Conversely, for the machine equilibrium (and the automaton equilibrium), some usual Nash equilibrium states become unattainable because the corresponding best responses cannot be computed by the automata. Hence it renders the model more probably refuted since none of the rare Nash equilibrium states may appear realistic.

The models receive few eductive or evolutionary justifications. An eductive justification seems simply irrelevant. Since the players have bounded rationality, it is difficult to assume that they would be able to reach an equilibrium state by pure reasoning. This would only be possible for simple games, but strong rationality would then apply to them. An evolutionary justification seems more plausible. If some (favorable) equilibrium states become attainable or if some (unfavorable) equilibrium states become unattainable, the players may have better performances with bounded rationality than with strong rationality. Hence, when players with strong rationality and players with bounded rationality are both present in a population, an evolutionary process may then select the latter. Moreover, the type of bounded rationality which is selected happens to be adapted to a given situation rather than to another.

To sum up, bounded rationality equilibrium concepts are still relatively under-developed. The first reason is again that there are too many directions in which it is possible, for each player, to depart from strong rationality and enter the realms of bounded rationality. The second reason is that the idea of bounded rationality is not in phase with the idea of equilibrium, at least when it is eductively interpreted. Bounded rationality receives a more relevant interpretation when the players face each other over time, like in learning models. The players then use learning rules which are computationally limited and they are directly engaged in an evolutionist process (see 6.7).

A fourth game concerned with road driving can be considered. The 'car lights game' (analogous to the classical 'prisoner's dilemma') is a symmetric game where the players are driving at night on either high beams or low beams. The consequences are essentially material ones. The preferences of each player are ranked in the following way: driving on high beams when the other is on low beams (utility 3), both driving on low beams (utility 2), both driving on high beams (utility 2), driving on low beams when the other is on high beams (utility 0). The only equilibrium state is for both drivers to drive on high beams, since no one has an interest in deviating. More profoundly, it is in each driver's interest to drive on high beams whatever the other does. A 'cooperation problem' is involved: the equilibrium outcome where both drive on high beams is Pareto-dominated by a (non equilibrium) outcome where both are on low beams. But with some models of bounded rationality, the outcome where both drivers drive on low beams can be obtained.

6

Learning processes among players

What we think we already know
often prevents us from learning.
C. Bernard

In a dynamic setting, the players can communicate through operational actions as well as informational actions, attributing a strategic aspect to the exchange of information. The receiver may be worse off after receiving a (true) piece of information from outside sources and is only guaranteed an improvement under specific circumstances. The sender may be better off keeping his private information to himself, and therefore preventing the communication and homogenization of information among the players. Besides, in order to face uncertainty and limited reasoning, the players exploit the time sequentiality of a game by jointly following more or less sophisticated learning processes.

Players are now able to play in a sequential way (6.1), leading to a refinement of the Nash equilibrium concept, namely the subgame perfect equilibrium concept (6.2). In addition, they face uncertainty about past facts, present structures and future events (6.3), leading to a further extension, the Bayesian equilibrium concept (6.4). In such a context, information received by a player provides him with a positive or negative value (6.5), and the exchange of information between players may or may not lead to a common belief (6.6). Finally, the players are again restricted to bounded rationality (6.7), which is compensated for by various learning processes involving founded on belief revision or on strategy reinforcement (6.8).

6.1 Intertemporal strategic rationality

In a dynamic game, as in a static game, the players operate in a material environment, but not an institutional one. The common material environment is again represented by a specific agent called nature. The players and nature play now sequentially in some given order and determine a path of the game. A game has a finite horizon if the number of successive moves in each possible path is bounded. A game is finite if it only considers a finite number of actions for each player for each possible move. The players are assumed to have rational behavior and to play independently. Nature is assumed to have deterministic behavior following given rules and to behave independently from the players.

The game is represented under an ‘extensive form’ by a ‘game tree’. As concerns the opportunities, each node corresponds to the player (or nature) having the move and the vertices issued from it to the actions (or states) he may implement. Hence, the possible actions of a player depend on the preceding actions of all players. Frequently, nature only acts once at the beginning of the game in order to fix the state, no further message being available after that. As concerns the preferences, they are represented by the synthetic payoffs each player receives for each path in the tree, i.e. for each terminal node. In fact, the players may get payoffs throughout the game which are then aggregated into a synthetic one. As concerns the beliefs, they are only partially integrated into the game tree (see 6.3).

In a situation of certainty, nature is simply considered as defining a precise state and no longer appears as a specific actor. The players are assumed to know perfectly the game tree designed by the modeler. During the game, the players know exactly what actions have already been implemented, hence they know which node they are at in the game tree. As concerns more generally the information the players receive during the play, a ‘hidden’ assumption is at work. It states that each piece of information has a unique interpretation which is the same for all players. The only uncertainty that remains is strategic and concerns the other players’ future actions.

In a dynamic game, each player chooses a strategy, a strategy again being defined by the action the player would implement in each possible situation. More precisely, a ‘pure strategy’ is defined by the action the player would play at each node where it is his turn to move. A ‘mixed strategy’ is a probability distribution he holds over all possible pure strategies. A ‘behavioral strategy’ is a probability distribution he holds over the actions he would implement at each relevant node. It can be proven that a mixed strategy is equivalent to a behavioral strategy when the player has 'perfect recall'. Of course, when a player's number of actions increases, the number of strategies increases dramatically.

A strategy appears in fact as a conditional statement of the type 'if I were in context c, I would take action s'. Here, the context is nothing other than the past history of the game, which is completely summarized in the present node. When choosing a strategy, a player defines, before playing, what he will do in any circumstances, then faithfully applies what he has chosen to do throughout the path played (since no 'surprise' can arise). The strategy is profactual for actions the player may later use. It is counterfactual for actions he will not be able to use, since they are forbidden by the other players' past moves or even his own preceding moves. A player's strategy may appear to another player as a threat or a promise, since it may state: 'if you take action s, then I will take action t'.

The concept of a strategy enables the 'extensive form' of a game to be transformed into a 'normal form'. Obviously, each combination of strategies followed by the players defines a unique outcome. Hence, the game matrix is obtained by considering the strategies of the players as actions and by indicating for each combination of strategies the payoffs obtained by the players. A normal form, on the contrary, cannot always be transformed into an extensive form, since it must satisfy certain constraints. Conversely, a normal form satisfying these constraints may be transformed into several extensive forms. In fact, in the normal form, time is crushed and the two forms are not really time-equivalent.

An extensive form game can be proposed for car driving. The 'parking problem' (analogous to the chain-store paradox) considers two drivers fighting over a car park just big enough for two cars. In the first period, the first driver chooses whether or not to enter the car park. The second driver, already in the car park, chooses whether or not to move his car to allow the first to enter. If the first driver chooses not to enter, he gets a utility of 0 while the second gets a utility of 2. If the first enters and the second gives in, they both obtain a utility of 1. If the first enters and the second resists, they both get a utility of 1. Since each player only plays once, the available strategies coincide with the actions. For the second driver, the strategy of resisting if the first enters appears as a threat he can express.

6.2 Subgame perfect equilibrium

In the normal form associated with the extensive form, it is possible to define the Nash equilibrium concept. Each strategy of a player is a best response to the others' (equilibrium) strategies. However, in a Nash equilibrium state, the dynamic dimension is considerably reduced. One player's strategy is a best response, computed at the beginning of the game, to the other's strategy. But when the strategies are implemented progressively, the actions remaining to be implemented may no longer be in equilibrium, since a player's payoff implicitly changes. In particular, a threat is non-credible if the player has no interest in applying it, once he is in the situation where he should do so. Hence, some Nash equilibrium states are not acceptable, since they rely on non-credible threats.

In the extensive form of the game, it is possible to define an adapted equilibrium concept, the 'subgame perfect equilibrium'. For any game, finite or infinite, a subgame perfect equilibrium state is defined as a state which is a Nash equilibrium not only in the whole game, but also in each subgame (i.e. a truncated game beginning at some node). However, for a finite game under certainty, the subgame perfect equilibrium is obtained in a constructive way by the 'backward induction procedure'. A player situated at a penultimate node chooses his best action with regard to the synthetic payoff. A player situated at a precedent node chooses his best action, with regard to the action chosen by his successors. For a finite game with no ties in payoffs, there exists a unique subgame perfect equilibrium state. For an infinite game, however, the number of equilibrium states may be far higher.

If a subgame perfect equilibrium is by definition a Nash equilibrium, the reverse is not true. The subgame perfect equilibrium is a 'refinement' of a Nash equilibrium. In particular, a subgame perfect equilibrium eliminates the non-credible threats. More precisely, for a finite game, the backward induction procedure ensures the 'dynamic consistency' of the equilibrium path. When arriving at some node where he has to move, a player has no incentive to modify the action he chose at the beginning of the game. In fact, the modeler may again consider that each player can be broken down into successive 'selves', each self acting once and only once. These selves have the same truncated preferences, but may act independently.

The subgame perfect equilibrium concept receives some epistemic justifications. It is again obtained by considering that the players have common knowledge of the game structure and common knowledge of the game rationality. If the assumptions are really fulfilled (as assumed by the modeler), the subgame equilibrium undoubtedly follows (Aumann, 1995). However, the backward induction procedure seems paradoxical since it is grounded on hypothetical moves which precisely are never implemented at equilibrium (Binmore, 1997). In fact, what a player is assumed to play out of the equilibrium path are just counterfactual moves. In this counterfactual reasoning, when a player observes during the play that his opponent does not follow the backward induction path, he has to interpret such a deviation. More precisely, facing a surprise, the player has to call into doubt one of the assumptions which justified in his eyes the presumed equilibrium concept.

Depending on the assumptions that are considered as false, the subgame perfect equilibrium may or may not be confirmed. If it is considered that the player has a 'trembling hand', i.e. that he applies the intended actions with random errors, then the equilibrium concept is kept. If it is considered that the rationality of the players is no longer common knowledge, then the subgame perfect equilibrium ceases to be justified and is replaced by other equilibrium concepts. The same is true when the structure of the game is no longer common knowledge. Consequently, the belief revision rule of a player has to be extended in order to specify how he deals with any new information. It becomes part of his determinants and has to be included in the game structure.

The subgame perfect equilibrium concept also receives evolutionary justifications. The (extensive form) game is repeated identically from period to period and each player is involved in an overall learning or evolutionary process (see 6.7). The learning process may be applied to two different entities. Firstly, it may be the player's strategy that is progressively adapted to past observations. Secondly, it may be the player's action at each node of the game tree which is progressively adapted to past observations. Under rather soft assumptions, implying that all nodes are regularly visited by the learning process, convergence towards a subgame perfect equilibrium obtains.

In the parking problem, there are two Nash equilibrium states. In one, the first driver enters and the second gives in. In the other, the first driver does not enter because the other will resist. However, the second Nash equilibrium state is based on a non-credible threat, since if the first driver actually enters, it is in the second's interest to give in. The first Nash equilibrium state is precisely the unique subgame perfect equilibrium state, obtained by the backward induction procedure (it is in the second driver's interest to give in when the first driver enters; knowing that, the first enters). This situation illustrates a general principle, according to which an effective threat is one which never actually has to be implemented.

6.3 Dynamic uncertainty

The players' uncertainty can now be analyzed by crossing two independent dimensions to generate nine classes. As concerns the object of the uncertain event, a player faces natural uncertainty (about nature), actorial uncertainty (about the other players) and personal uncertainty (about himself). As concerns the occurrence of the uncertain event, a player faces factual uncertainty (about past states, the others' past actions and his own past actions), structural uncertainty (about the generation law of states, the other players' types and his own type) and strategic uncertainty (about the future state, the others' future actions and his own intended action).

Some forms of uncertainty are directly integrated into the game tree in two forms. Uncertainty may be formalized in a set-theoretic way, for instance when a player cannot observe anothers past actions. He groups together in an 'information set' all the nodes where it is his turn to move and which he cannot distinguish, due to his imperfect perception of past history. The information sets form a partition on the player's set of nodes. In an information set, the set of available actions is the same for each belonging node, since other wise, the player could infer which node he is at. Uncertainty is more often formalized in a probabilistic way, for instance when a player is unaware of another's type. He represents uncertainty by a probability distribution on a set of possible types. In fact, it is generally assumed that the type of each player is randomly chosen by nature at the beginning of the game.

The extensive form of a game can be extended from a game tree to a more general type of game, the 'stochastic game'. The game tree is generalized in a game graph which may admit cycles, since the play go several times through similar 'configurations'. Each player's node is associated with such a configuration characterizing the state of the system. For each action of a player, stemming from a node, the configuration is modified according to a transition function, which is stochastic to take uncertainty about nature into consideration. The player's payoff is defined each time he acts and it depends jointly on his action and on the present configuration, again including a random term. The strategy of a player is always defined as his action in each node and each configuration.

As time unfolds, a player gets information about nature through observation, about the others through empathy, and about himself through introspection. When getting information about one form of uncertainty, the player may reduce uncertainty about another form by implementing various reasoning modes (see 2.7). Firstly, if a player gets a message about past events, he may reduce his factual uncertainty. This is achieved thanks to nonmonotonic reasoning, as expressed in standard belief revision (about nature). Secondly, if a player gets factual information, he may reduce his structural uncertainty. This is achieved thanks to abductive reasoning, since an opponent's characteristics (fixed during the play) can be revealed from his observed actions. Thirdly, if a player gets structural information, he may reduce his strategic uncertainty. This is achieved thanks to conditional reasoning, since an opponent's future action is obtained by simulating his behavior on the basis of his known characteristics.

The second operation, which concerns the 'interpretation' by a player of the observations he has made, is the most important one. Interpretation is a form of explanation which attributes what the player observes to more profound factors concerning his environment or himself. The same interpretation problem is faced by the modeler when he attributes the observations made about a player to specific entities or characteristics of them. Attribution is clearly more difficult in game theory than in decision theory, since determinants are more intricate. Attribution is always multivocal since many explanations are available for the same facts. Attribution may be false since certain explanations do not agree with other observed phenomena.

A first attribution problem concerns the explanation by some player of commonly observed consequences. Such consequences can be attributed to nature, to the other's action or to his own action. For instance, schizophrenia consists in attributing to his own action the effects of another's action or, conversely, in attributing to the other's action the effects of his own action. Likewise, rationalization consists in attributing good effects to his own actions and bad effects to the other's actions. A second attribution problem concerns the interpretation by some player of the other's action. Assuming that the other is rational, this action can be attributed to any of the other's determinants. For instance, accusing the other on the basis of his supposed intentions attributes his action to specific preferences (rather than to specific beliefs). Likewise, regretting the other's non-conformist action often consists in attributing this action to ill-informed beliefs rather than a desire to act unconventionally.

The static crossroads game is implicitly an imperfect information game, since each driver acts without observing what the other is doing at the same time. It can be expressed in an extensive form where one driver plays first and the second later, but without knowing what the first did. Hence, it brings to the fore the phenomenon of 'fighting for *the first move'. In fact, each player has an interest in moving first and in not stopping. Each player may even commit himself by blocking the accelerator and informing the other player that he has done this. In other games, one can observe a fight for the second move (each player has an interest in letting the other reveal his action first), agreement on move order (both players agree on a given order of moves) or indifference (the players are indifferent about the order of moves, as in the car lights problem).*

6.4 Perfect Bayesian equilibrium

Consider a dynamic game situation, called a Bayesian game, in which only factual uncertainty is explicitly involved. On the one hand, players cannot observe all past actions of their opponents and consider 'information sets'. On the other hand, nature defines some states which are only partially observed by each player. Such a framework is general enough to deal with a repeated one-shot game. More precisely, in each period, each player plays without knowing what the other does in the same period, but knowing what the other played earlier. Such a framework can also deal with certain forms of structural uncertainty. More precisely, nature plays at the beginning of the game in order to define the players' types, each player observing his own type. In Bayesian games, the strategy of a player is defined as the action he selects for each information set.

A subgame perfect equilibrium can always be defined by considering 'proper subgames', i.e. subgames which intersect no information sets. But it eliminates few Nash equilibria and needs to be strengthened. A 'Bayesian perfect equilibrium' is obtained by specifying more precisely what the player's beliefs are for each information set, i.e. by introducing a probability distribution on the nodes of the set. Such a belief reveals what the player thinks (probabilistically) to be the past history of the game. An equilibrium state is then obtained not only as a fixed point of a loop relating the actions of all players, but as a fixed point of loops relating the beliefs and strategies of each player. As instrumental rationality is concerned, at each information set, the strategy of a player is the best response to the others' strategies, for given local beliefs. As cognitive rationality is concerned, at each information set consistent with the players' strategies, the beliefs are inferred from past actions by the Bayes rule.

The Bayesian perfect equilibrium states are generally very numerous since beliefs are only assumed to be consistent with observations.

This is especially true for repeated games, either with infinite horizon or even with finite horizon. In order to be more selective, several refinements of the Bayesian perfect equilibrium have been proposed. A ‘sequential equilibrium’ is obtained by imposing additional constraints on the players’ beliefs, especially outside the equilibrium path. Another equilibrium concept is based on the ‘forward induction principle’. This principle indicates how to reveal the intentions of a player with regard to his past actions. In fact, it is a special form of interpretation of the other’s action when he deviates from an equilibrium path.

In repeated games, the players communicate through two types of action. Operational actions induce some material and psychological consequences, and hence have a direct impact on the players’ utility. informational actions only provide information to the players, and hence only influence utility through their eventual cost. However, information may be used in order to adapt further operational actions, in which case it has an indirect impact on utility. For instance, ‘cheap talk’ considers informal communication between players before the game really starts. They exchange information at low cost, but are unable to define binding agreements. Since information becomes a strategic item, the players have to acquire or produce information by considering its impact on further actions.

From the point of view of the information receiver, except in the classic case of pure experimentation, the player receives information in two ways. In passive experimentation, he receives information as a by-product of his ongoing actions. This information is free and provides some utility when it is later used. In active experimentation, he obtains information by deliberately deviating from his normal course of action. He loses some short-term utility, but gains long-term utility when exploiting his information. The former ‘exploration-exploitation dilemma’ (see 4.6) is then extended from a passive to a strategic context, where it is harder to deal with because it is embedded in an equilibrium concept.

From the point of view of the information sender, the player giving information through an action can handle it in two ways. Firstly, depending on whether the information provided is ultimately beneficial or not for him, he can highlight or conceal this information. An equilibrium state is said to be ‘revealing’ when the players have an incentive to diffuse their information and ‘pooling’ when they have an interest in keeping it to themselves. Secondly, using the impact of his information on another’s action, he may change his action in order to induce the other to adopt a belief which is favorable to him. For instance, in the ‘bluff’ mechanism, one player tries to make the other believe that he has different information to what he really has, by acting in a way appropriate to the false information. Likewise, in the 'reputation' mechanism, a player aims to be considered as another type than his real type, by playing an action characteristic of the usurped type (see the example in 6.6.

The crossroads problem can be developed by considering that one driver, before arriving at the crossroads, can choose whether or not to make a detour. When he does not make the detour, the game happens as usual. When he does make the detour, the drivers get 2 1/2 each. The 'forward induction' reasoning is then the following. If the second driver observes that the first driver has not taken the alternative route, he must infer that the first driver has renounced a payoff of 2 1/2, and this can only be because he expects a payoff of 3. Hence, at the crossroads, the second driver expects the first to keep going and can only respond by stopping. This reasoning leads to the selection of one of the two Nash equilibrium states, the one in favor of the first driver.

6.5 Value of information

Consider a one-shot game where two players face some factual uncertainty about their material environment. More precisely, at the beginning of the game, nature defines stochastically a state of nature which conditions the matrix of the game. The players have a common prior distribution over the states of nature. Moreover, each player has a private (set-theoretic) belief structure about the states of nature (and the crossed beliefs about it). At some Bayesian equilibrium state (assumed to be adequately sorted out), they get some expected utility. At a given time, they receive a personal message and revise their hierarchical belief structure according to a multi-player revision rule (see 2.6). The value of information for a player is nothing other than the difference between the utility obtained before and that obtained after reception of the message.

More and more averaged information values are defined by the modeler for any player. The 'actual information value' is just the difference of utility obtained by a player with regard to the real world. It is known by the modeler before play, but not by the player. The '*ex post* information value' is the actual value measured on average for all worlds considered as accessible with the player's final beliefs. It can be computed by the player, but only after reception of the message. The '*ex ante* information value' is the actual value measured on average for all worlds considered as accessible with the players initial beliefs. It can be computed by the player before reception of the message. Hence, it can be compared to the cost of the message as a means of deciding whether or not to buy it.

The message is assumed to be a specification message which does not contradict the initial belief (the latter being either true or false). It is usually defined by its content and its status (see 2.6). In particular, one can consider public, private or secret messages. The message is usually compared to the null message, since it is more accurate (for any technical sense of the term accuracy). It may also be compared to any less accurate message (in a precise technical sense). It is already known that the accuracy order between two messages is transmitted, for a given initial belief, to the final beliefs (see 2.6). The problem is then to find out if a more accurate message is also associated with a greater information value.

In a strategic context, the *ex ante* value of information provided by a message may well be negative. This means that the receiver of a message may be worse off after receiving the message than before. In fact, for a private message received by one of two players and assumed to be true, all combinations of information value are possible. Depending on the game under consideration, both players may be better off, both players may be worse off, the receiver may be better off and the other worse off, the receiver may be worse off and the other better off. A negative value occurs, for instance, when new information prevents the players from forming mutually beneficial agreements.

Such a result seems paradoxical at first sight, since the receiver of a message, knowing that its value is negative, can act as if he did not receive it. However, the other player knows that he received the message and acts in consequence, the receiver then being even worse off when he ignores the message. The idea of the negative value of information clearly goes against common sense, which asserts that true information about nature, given to all members of a social system, should ameliorate their performances. However, sociologists have always been convinced that some opaqueness is useful to stabilize a social system, since it prevents certain greedy interests from coming into play.

However, the *ex ante* information value is positive for some specific classes of games and types of messages. Firstly, it is positive for the receiver of a secret message in any game. The reason is simply that a secret message does not change the other's beliefs; it therefore acts like a message in an individual decision-making problem. Secondly, it is positive for the receiver of a private message in a zero-sum game. The reason is that making best use of his information corresponds to the worst use of his information for the other player. Thirdly, it is positive for each player receiving a public message in a pure coordination game. The reason is that the players act as if they formed a unified team.

A last road game is the 'car transaction game' (Akerlof, 1970), where two players act as the buyer and the seller of a car. The main uncertainty concerns the quality of the car, reduced to two states: good car or bad car. The drivers know that the prior quality of a car is equiprobable, but the actual quality of a car is only known by the seller. The value of a bad car is considered as 0 for both drivers while the value of a good car is 6 for the seller and 10 for the buyer. The transaction happens if and only if both drivers agree and the transaction price is exogenously fixed to 4. When information is asymmetric, as assumed, the mean payoff for each driver is 1 when the transaction takes place and 0 when there is no transaction, so the transaction takes place. Conversely, when the buyer learns the quality of the car, the transaction never happens since if the car is good, the seller has no incentive to sell it and when the car is bad, the buyer has no incentive to buy it. Hence, the value of a message informing the buyer of the quality of the car is negative for both players.

6.6 Transmission of information

The 'three hats problem' is a first puzzle dealing with the transmission of information between players about some state of nature. Three boys have hats on their heads, which they are told are either blue or red; in fact, the modeler knows they are all red. Each boy can see the color of the others' hats, but not his own. At the start, an observer gives a prior message that there is at least one red hat in the group. He fixes the rules of the game: in successive periods, a boy announces whether or not he knows the color of his hat. Each boy can observe the others' announcements, acting as messages. The payoffs are not made explicit, but a boy wins if he announces the true color of his hat and loses otherwise. It can be shown (by induction on the number of boys) that, for actors with perfect rationality, everybody makes a negative announcement in the first two periods and a positive announcement in the third one.

In this example, the process evolves in three steps from distributed knowledge to common knowledge about the combination of hat colors (a possible combination acts as a possible world). Several factors explain this result, which corresponds to a complete diffusion of information. Firstly, the players have no strategic interaction, since their payoffs do not depend on the others' actions. Secondly, the prior message transforms a shared initial knowledge into a common initial knowledge and players ground their reasoning on that common basis. Thirdly, since the number of possible worlds is finite, shared knowledge gains one level at each step and must converge towards common knowledge.

The 'two generals problem' is a second puzzle which shows a direct transmission of information. Two allied generals fight against a common enemy in a twofold situation. If the situation is unfavorable, a simultaneous attack is bad for both generals, while a lack of attack is better, whatever the other does. If the situation is favorable, a simultaneous attack is good for both generals while a lack of attack is bad, whatever the other does. One general is on a hill top and observes the situation while the other is in a valley and observes nothing. When the situation becomes favorable, the first general sends a message of attack to the other, but the message has a small probability of getting lost. When he gets the message, the second general sends a message to the first confirming that he has received it, but this message has the same probability of getting lost. The process continues until one message is lost. It can be shown that rational generals never attack, whatever the number of exchanged messages.

In this example, the simultaneous attack could only happen if common knowledge of the situation were achieved, but this never happens. Several factors explain such an unsuccessful result, corresponding to an incomplete diffusion of information. First, the players are in a strategic context since their payoffs depend on the actions of both; however, information is transmitted directly and not through their actions. Second, the only prior belief which is common knowledge is a common probability distribution about the situation. Third, since the number of possible worlds is infinite, shared knowledge of the situation never becomes common knowledge. However, the result is not robust to small changes in the assumptions. For instance, if the generals agree by a prior convention that they attack after exactly n (greater than two) exchanges of messages, the attack will be implemented (if n exchanges are effectively exchanged) and succeeds.

The 'two restaurants problem' is a third illustration of the transmission of information, but in an indirect way. There are two restaurants, one on the left side and one on the right side of a street. One is good quality and the other is bad quality. Successive customers arrive to eat a meal. In fact, they aim for the better restaurant, but they don't know which it is. They have common information under the form of a prior probability indicating that the left one is a bit better. They receive private information under the form of a message correlated with the quality of the left (or right) restaurant. They also receive public information, since they observe the behavior of the preceding customers. It can be shown that after a certain number of periods where the customers alternate, they end up all going to the same restaurant, whether or not it is the better one.

In this example, it becomes common knowledge among the followers that one restaurant is probably the better one, even if it is the wrong one. The arbitration between the players' private information and their public information rapidly turns in favor of the second. Many factors explain this paradoxical result. Firstly, the players have independent preferences, since their payoff does not depend on the occupation of the restaurant. Secondly, they are weakly pre-coordinated by the prior probability distribution acting as a common reference. Thirdly, they have no long-term memory, since it is a different customer who arrives in each period. As a more general framework, consider that two players have a common prior probability of some event and exchange respective messages about their posterior probability in each period. It can be shown that their posterior probabilities of the event finally converge: they cannot 'agree to disagree' (Aumann, 1976).

In the crossroads game, the first driver may initially be uncertain about whether the second is a go-getter or a cautious type. The first driver attributes a probability to each type, constituting the 'reputation' of the second. Such a reputation evolves during the game as a function of the observed action of the second driver. As long as he keeps going, his reputation of go-getter driver slowly increases; whenever he stops, his reputation of go-getter driver drops straight to zero. In fact, it may be to the advantage of a cautious driver to mimic the behavior of a go-getter driver. More precisely, at the (Bayesian perfect) equilibrium state, the second driver keeps going in all the first periods, in order to appear as a go-getter, and then keeps going or stops randomly. After stopping once, he stops for ever since he no longer has a reputation to defend.

6.7 Dynamic processes under bounded rationality

Involved in a repeated (static or dynamic game), the players are now endowed with bounded rationality. Due to the complexity of the game, they are unable to coordinate by their sole reasoning to an equilibrium state. At the opposite, relying on past observations, they progressively adapt to their material and social environment (Egidi, 2007). A self-organization process is at work, which presents transitory states and may or not lead asymptotically to an equilibrium state. When the process converges, the implementation of the equilibrium, formerly achieved from outside by the Nash regulator, is now spontaneously achieved by the work of repeated experience. Seemingly, the selection of an equilibrium state is automatically solved, since a precise asymptotic state is obtained, with respect to the random elements. Such a dynamic process follows a path which depends on initial conditions and exogenous factors. It is governed by five principles which characterize so-called 'evolutionary game theory', only the first being shared with 'classical game theory'.

The 'satisfaction principle' states that the players have certain actions at their disposal and receive a utility as a combination of their joint actions. Hence, the basic game can be represented as usual by a game matrix or by a game tree. A player is frequently represented by a population of agents having the same action set and the same utility function. The agents playing the same action are grouped into a subpopulation. When the game is asymmetric, one agent from one population plays against one agent from another population; hence, the game is a multi-population one. When the game is symmetric, two cases are possible. In a multi-population game, an agent from one population (a row-agent) plays again with an agent of the other population (a column-agent). In a mono-population game, any agent may play against any other agent.

The 'interaction principle' states that, in each period, pairs formed by an agent of each population are drawn and play the game together. Agents are frequently disposed in a network supported by some underlying structure (line, circle, surface, torus, etc.). Each agent can only meet the agents situated in a certain neighborhood called the 'interaction neighborhood'. Moreover, in each period, several pairs of agents are effectively drawn, their number being anything from only one pair to all possible pairs. In particular, a random sample of pairs can be drawn. This pooling process favors anonymity, as players meeting in a given period have a weak probability of meeting again later. Anonymity induces short-term behavior since the agents cannot hope to influence their opponents with threats or promises.

The 'information principle' states that each agent gets some information on past plays. Such factual information is gathered in a neighborhood called the 'information neighborhood', which is frequently included in the interaction neighborhood. Moreover, in each period, the information is collected either in the whole neighborhood or in only a sample of it. The information of a player mainly concerns the opponent's past actions or his own improved utility. Past information is grouped together in a memory which may be limited to a fixed number of past periods. Contrary to classical game theory, the player has little information about the game structure. However, he may abduce some structural information from the factual information. In particular, he may explore some parts of the game structure he is not aware of.

The 'evaluation principle' concerns the way in which information is computed to construct synthetic indices. On the one hand, an agent observing the opponent's past actions may compute statistical parameters (frequency, variance) about these actions. Moreover, he transforms these past observed actions into future expectations. On the other hand, an agent observing his own past utilities may compute an aggregating index (mean, cumulative or discounted value) of the performances of each action. Moreover, he adapts 'aspiration levels' for his choice criteria according to his overall past performance. More generally, an agent tries to discover some regularities in the past observations. He may observe that the others actions have some precise patterns, especially if the other player imitates his own action. He may observe that his own utilities display certain invariants, especially when some of his actions give concentrated utilities.

The 'decision principle' defines how the preceding indices, as well as regularities, are exploited in order to make a choice. The rule is always myopic, since the agent optimizes separately in each period without taking intertemporal effects into account. In practice, any choice rule is a heuristic process which achieves a trade-off between two behaviors (see 3.6). Exploitation behavior is achieved by several principles, ranging from an optimizing to an improving principle. It may be partially mimetic, if the agent thinks the other knows more or obtains better performances than he does. Exploration behavior is achieved by several principles ranging from random deviations to voluntary exploration of specific actions. It can be initiated by past payoffs falling below the current aspiration levels.

In the crossroads game, the global population of drivers in a town may be split or not into separate populations. In a multi-population game, the drivers going from area A to area B always meet exclusively the drivers going from area C to area D. In a mono-population game, the drivers go in various directions and meet randomly inside the town. A driver can gather information and observe that his utility is almost always the same when he stops and is more dispersed when he keeps going. He summarizes his information in indices specifying what per-

centage of other drivers stopped in the past or what utility he personally got in the past when keeping going.

6.8 Learning models

A first class of models involves ‘epistemic learning’ (or ‘belief-based learning’). Since the main driving principle is the revision of the players’ beliefs, they are guided by a ‘Bayesian hand’. Each player knows his own payoff function (but not the other’s) and observes the past actions of his opponent. He modifies his expectation about the future action of his opponent accordingly. Finally, he reacts to the expectation by a choice rule mixing best response with some experimentation. The simplest model in this class is ‘fictitious play’. The player observes the past actions of his opponent. Assuming the other’s behavior to be stationary, he transforms the frequency of his past actions into a probability of his future action. Finally, he chooses a best response to this expectation. In a stochastic variant of fictitious play, a player plays a best response (achieving exploitation) with a high probability and some random action (achieving exploration) with the complementary probability.

A second class of models involves ‘behavioral learning’ (or ‘reinforcement learning’). Since the main driving principle is the ‘reinforcement’ of the players’ best performing actions, they are guided by a ‘Skinnerian hand’. The players have no structural information, and each of them observes exclusively the results of his preceding actions. He computes an aggregated index of the past performance of each action, eventually compared to an evolving aspiration level. He selects an action by playing more often those with a high index and less often those with a low index. The simplest model in this class is the ‘basic reinforcement model’. Each player observes the past utility obtained by his action. He computes an index for each action consisting of the cumulative sum of past utilities obtained with that action. He uses the probabilistic choice rule, associating exploitation (the best actions are chosen more often) and exploration (an action is never abandoned, even if it is chosen more and more rarely).

A third class of models involves an ‘evolutionary process’ (Weibull, 1995). Since the main driving principle is ‘selection’ of the fittest players in a population, they are guided by a ‘Lamarckian hand’. A player observes nothing (except the other’s action if his strategy depends on it). He is represented by subpopulations of agents, all of them having the same fixed strategy. He is involved in a ‘selection process’ which favors the reproduction of agents obtaining the highest utility (likened to fitness in biology, which precisely is a rate of replication); he is submitted to a 'mutation process' where some different agents are randomly introduced into the population. The simplest model in this class is the 'replicator model'. In a mono-population or multi-population framework, agents with fixed strategies meet randomly. They reproduce in proportion to the utility they get from their interactions (ensuring exploitation). In a variant, the 'stochastic replicator', mutant strategies are also introduced randomly into a population (ensuring exploration).

The three classes of models have been extended from decision under uncertainty (see 4.8) to game theory, where their relevance increases still further. They attribute less and less structural knowledge and prior information to the players. They endow the players with increasingly bounded cognitive and instrumental rationality. Moreover, the two last classes present a formal isomorphism, the probability of a given player playing an action being replaced by the proportion of players playing that action. Hence, it is possible to confine ourselves to the two first classes, which appear to be the more realistic ones since they entail no outside entity and depart from biology. In other respects, the pure categories of models can be mixed according to various different principles. They may be considered as relevant in different contexts for a same player. They may be used by different players in a same game. They can also be combined in hybrid models.

The asymptotic properties of the dynamic process depend on the classes of games and learning models used, as well as on the type of stability considered. The only general result is the elimination of the strongly dominated strategies. Otherwise, the process may be asymptotically cyclical or chaotic, but frequently converges towards some equilibrium state. The process converges rather easily (in beliefs or in strategies) towards a strict Nash equilibrium, less easily towards a non strict one (especially a mixed one). Refinements of the Nash equilibrium can sometimes be obtained. With stochastic fictitious play, a 'risk-dominant' equilibrium state (see 4.5) is selected, at least over the very long-term. With an adapted basic reinforcement model, the subgame perfect equilibrium is selected. With the stochastic replicator, the process may converge towards an 'evolutionary stable equilibrium'. Such an equilibrium state is defined by the fact that, if some mutant agents are introduced in small numbers into a population in equilibrium, they are rapidly eliminated.

The transitory process may exhibit long stable periods separated by deep ruptures. The asymptotic process may be slow or fast and lead to a lock in some states. In fact, the path of the system depends heavily on the stochastic factors introduced on interaction modes, information samples, expectation rules or choice rules. Of course, the learning processes can be extended to players situated on a network and limited to neighborhood interactions. It can be applied not only to actions, but also to the more structural characteristics of the players. For instance, players can hold different beliefs which are progressively selected not so much according to their truth as according to their performance. They may even hold different expectation schemes or belief revision schemes which are progressively selected.

In the driving side game, different dynamic processes can be considered, all of them finally leading the drivers to drive on the same side of the road. With epistemic learning, each driver drives on the right (left) when he observes that a majority of drivers drove on the right (left) in the past. With behavioral learning, each driver drives on the right (left) when he observes that he had fewer accidents in the past by driving on the right (left). In an evolutionary process, two drivers who meet die if they are on opposite sides of the road and reproduce if they drive on the same side of the road. In the crossroads game, under a mono-population assumption, an evolutionary process leads one half of the drivers to keep going and the other half to stop. Under a two-population assumption, an evolutionary process leads all the drivers in one population to keep going and all the drivers in the other to stop. In the technology game, with an evolutionary process, the gas car (corresponding to the risk-dominant equilibrium state) is selected.

7

Communication and reasoning in an economic system

Knowledge is power
by itself.
F. Bacon

In economic theory, several types of agents perform transactions involving labeled goods within an institutional context formed of markets and other institutions. A match is achieved between a material sphere where goods are physically transformed and exchanged and a cognitive sphere where related data are computed and communicated. In other respects, knowledge and information may themselves be treated as specific forms of goods which can respectively be accumulated and exchanged by the agents. Finally, if different kinds of institutions are needed in order to coordinate the agents, their genesis receives a first explanation on the basis of common knowledge of the situation.

If the modeler's knowledge of an economic system is based on a shared ontology (7.1), the agents only have an imperfect and incomplete view of that system (7.2). While information acts as a privileged device for the external coordination of agents (7.3), knowledge acts as a necessarily distributed item in the internal organization of an agent (7.4). Likewise, while information appears as a specific commodity exchanged on some market (7.5), knowledge appears as an immaterial capital accumulated by an agent (7.6). Finally, various institutions play an essential cognitive role in coordinating the agents (7.7) and may be obtained through a purely eductive process implemented by all agents (7.8).

7.1 Modeler's view

According to classical ontology, the modeler considers three economic entities as basic ones: goods, agents and institutions. Goods are considered as modular items which may be transmitted voluntarily from one agent to another. Agents are considered as independent entities whose actions consist in producing, exchanging and consuming goods. institutions are considered as distinctive artifacts capable of coordinating the preceding economic operations through institutional signals. Two more entities are sometimes considered. Firstly, the agents are immersed into a physical environment which provides them with certain basic resources. Secondly, the agents may be linked by permanent (economic or extra-economic) relations imposing constraints on their actions.

According to weak methodological individualism, any social phenomenon can be explained exclusively by the agents' actions, with the mediation of institutions and under the constraint of permanent relations. According to strong methodological individualism, the institutions and the permanent relations in turn need to be explained by the agents' actions. As for the agents, they are considered to be rational and endowed with specific determinants according to their economic role. Moreover, the agents are considered as independent in the sense that their determinants are not influenced structurally by external factors, even if these latter may act as parameters. As concerns the physical environment, it is treated as a passive agent ('nature') which follows certain mechanical and often random laws.

Goods are categorized into several classes, according to their physical features or functional properties. A first taxonomy makes a distinction between 'material goods' appearing as objects (such as forks) and 'immaterial goods' appearing as statements (such as patents). A second taxonomy distinguishes between 'investment goods' consumed over a long period (such as machinery) and 'consumption goods' consumed over a short period (such as oil). In fact, the goods are more precisely defined in nested classes according to their similarity. We can, for instance, group together all means of transportation, consider all cars of a given brand, distinguish a car of a given type or even consider separately the cars of a given year or color. The similarity between goods is far higher when they are substitutes (like butter or jam) rather than complements (like car and gas) in economic operations.

Agents are categorized into several classes according to their social constitution or economic role. A first taxonomy distinguishes between 'individual agents' (such as workers) and 'collective agents' (such as firms). A second taxonomy makes a distinction between 'producers' who produce goods from other goods and 'consumers' who definitively destroy the goods. Nested taxonomies are again constructed. For instance, one may consider all workers together, distinguish the subclass of teachers, specify at what level they teach or even distinguish between them according to their age and gender. The similarity index is again higher when the agents are substitutive (like psychiatrist and psychoanalyst) rather than complementary (like judge and advocate).

Institutions are categorized into several classes according to their formal support or coordinating function. A first taxonomy makes a distinction between 'formal institutions' (such as laws) and 'informal institutions' (such as conventions). A second taxonomy distinguishes between 'constitutive institutions' which create a type of link between agents (such as new financial markets) and 'regulative institutions' which regulate already given links between agents (such as traffic regulations). Again, nested taxonomies are constructed. For instance, one may consider all institutions operating by means of their normative content, the subclass of social norms indicating what to do in specific circumstances, among them the norms of politeness indicating how to behave in a collective setting or even more precise norms concerning how somebody should be addressed. Some institutions are complementary (such as property rights and markets) while others are substitutive (such as folk or technical languages).

Relations between agents are categorized into several classes according to their physical nature or functional structure. A first taxonomy distinguishes between 'organic links' (such as physical links between a polluting firm and a polluted agent) and 'epistemic links' (such as loyalty links between a seller and a buyer). A second taxonomy makes a distinction between 'spatial links' based on a physical neighborhood (such as congestion on the road) and 'qualitative links' based on a more diffuse partnership (such as branch relations between firms). The permanent relations are designed along nested networks. For instance, consumers of ice cream are situated on a plane (district) or a line (beach), are subdivided into local graphs (families) among which specific entities are singled out (adults and children). Several networks are also complementary (such as electricity and telephone) or substitutive (such as mail and e-mail).

An employer and an employee jointly define an employment contract. The employer is guided by the profit he gets from the work of his employee, more precisely the difference between the employee's productivity and his wage. The employee is sensitive to his utility, which *combines his wage and the hardness of his work. The contract defines not only the initial conditions of the job (task to be done, initial wage), but also the general conditions of future employment (change of task, rise in wage). It is constrained by many institutional rules (such as wage rising with seniority) frequently expressed in formal rights (such as dismissal rights). It is influenced by the labor market surrounding the firm, which conditions, for instance, a reservation wage for the employee.*

7.2 Agent's view

As concerns his knowledge of the economic system, each agent is assumed to have the same structural view as the modeler about the basic entities. He considers that the other agents are rational and act on goods in some institutional context (and even considers that this is common knowledge). However, he may well overlook certain relations, either between the environment and himself or within his environment. For instance, he considers that the institutional environment concretely influences his behavior but assumes (truly or not) that he has a marginal effect on the institutional environment. Likewise, he considers that the material environment effectively acts on him, but he ignores most feedback effects. Finally, he believes that he can transform his permanent relations with the other agents, but only over the long term.

More specifically, each agent is capable of adopting a personal classification of goods, agents and institutions. The more general distinctions, such as those between investment and consumption goods, are generally endorsed by all agents because of their convenience. But the nested taxonomies may differ profoundly from one agent to the next, according to their culture and past experience. In principle, any official classification acts as a specific institution which is designed with the aim of coordinating agents' knowledge. But such a taxonomy is only partially shared by the agents, who adopt their own categories, adapted to the agents they meet most frequently. Moreover, each agent only has a local view and only considers a relevant subset of goods, agents and institutions. The remaining environment is treated as an undifferentiated whole, usually with mechanical behavior.

Now, as concerns his information about the economic system, each agent is assumed to observe the same characteristics as the modeler does. He is unable to observe the others' determinants when they adopt a generic structure, as is the case for the utility function of a consumer.

However, he may know some determinants which are sufficiently specified, as is the case for the (linear) profit function of a producer. He observes more easily the payoffs obtained by the other agents in production and consumption operations. Finally, he observes essentially the informational or material actions on goods of neighboring agents and the signals (such as prices) emitted by certain institutions.

More concretely, each agent observes any variable in a cruder way than the modeler. He is endowed with successive filters acting between elementary data and basic information (perceptual filters) or between basic information and compound indices (conceptual filters). He faces costs when he gathers information either through observation or by purchase from external sources. He only observes the properties of the environment within a certain neighborhood and may even only observe aggregate properties. In particular, he knows his own characteristics better than those of the other agents. imperfect information leads to asymmetric information between the agents. 'Moral hazard' involves differential information about an agent's action. 'Adverse selection' involves differential information about an agent's determinants.

Finally, it is assumed that all the agents treat knowledge and information in the same way as the modeler does. On the one hand, they produce structural knowledge from past knowledge and present factual information. For instance, a producer may be able to discover the consumer's preferences on goods from the observed demand he receives in different contexts. Likewise, a producer may be able to abduce his competitor's behavior function (relating his production to prices) from his observed past actions. On the other hand, they infer new factual information from the structural information previously produced. For instance, a producer may predict future prices by relying on a simplified equilibrium model of a market for a good.

However, every agent faces strong limitations on his computational capacities. For instance, the strong 'rational expectation' assumption stating that an agent has the same cognitive competences as the modeler depends on three weighty conditions. The agent knows the structure of the system perfectly, he observes past variables perfectly and he makes an optimal expectation in accordance with statistical methods. In practice, the agent forms at best a 'weakly rational expectation', less sophisticated than that of the modeler. Not only is he limited in his structural and factual information about the system, but he also uses approximate statistical methods to form his expectations. Frequently, he contents himself with an extrapolative expectation function, relating the expected variable to its past values.

Let us return to the employment contract example. As concerns the employer, he may only be partially aware of the employee's characteristics, especially the precise specification of his utility function. Likewise, he is generally not aware of the potential 'productivity' of the employee, even if he assumes that it is related to observable traits (such as education). He is no longer able to observe the employee's 'effort' in working, considered as an action variable. As concerns the employee, he is aware, to some extent, of the employer's general goal of maximizing his profit. But he is unsure about the link between his work and the results of the firm. He is no longer able to observe the strategy used by the employer to achieve his profit. Finally, only the wages which are transferred from the employer to the employee are observed perfectly by both parties.

7.3 Information as an external coordination signal

Consider the classical representation of a set of agents willing to exchange certain goods for other goods. They enter with their own initial resources and with given preferences about the goods. They have no permanent relations which might favor some transactions over others. They are assumed to share a common taxonomy of goods and to be able to observe a limited number of data. They are considered to be cognitively rational insofar as that they have a simplified, but sufficient view of their environment. They are considered to be instrumentally rational as long as they are able to optimize the quantity of desired exchanges. Moreover, the exchange of goods between the agents is an operation often assumed to involve no cost.

The main theoretical problem is to coordinate these heterogeneous agents through a commonly accepted institutional device. A standard assumption prescribes that the institution should neither impose physical constraints on the agents' transactions nor modify their preferences structurally. Consequently, the institution is assumed exclusively to produce informational signals influencing the parameters of the agents' constraints and preferences. Concretely, an institution may be represented by an external entity, fictitious or not, capable of realizing an equilibrium. An equilibrium state is achieved when the agents can see no further possibilities of fruitful bilateral exchange. It is obtained by a fixed point of the loop relating the institutional signals to the individual exchanges.

The main institution is the 'competitive market', in which all transactions are coordinated by a single type of signal: the price defined for each good. The price system is computed by a 'Walrasian auctioneer' who is in contact with all the agents. In fact, a competitive equilibrium is obtained under three conditions. Each agent fixes optimally his supply or demand for each good as a function of prices (instrumental rationality). The Walrasian auctioneer fixes the price of each good by equalizing the supply and demand (institutional coordination). Each agent observes the current prices and reasons from them (cognitive rationality). In some cases, in a fish market or a financial market, the Walrasian auctioneer is a real entity. More often, it is just a fictitious entity which acts 'as if' comparing supply and demand.

The attainment of an equilibrium state raises three fundamental problems concerning the Walrasian auctioneer. The 'implementation problem' stems from the necessary existence of a concrete auctioneer or a suitable alternative. The 'synchronicity problem' originates in the fact that the auctioneer has to fix the prices at the same time as the agents fix their supply and demand. The 'selection problem' arises when several equilibrium price systems satisfy the equilibrium requirements. An eductive solution to the two first problems (see 7.8) considers that the agents are able to simulate the computation of prices by the auctioneer and to form 'rational expectations' on these prices. However, the third problem remains, since the agents have to coordinate on the same rational expectation in the event of multiplicity.

Some institutions are designed to be complementary to the market. 'money', for instance, helps to fluidify transactions, because it defines a common value standard which enables exchanges to be carried out exclusively between goods and money. Likewise, 'trust' helps to secure transactions, since it ties together exchanges which are respective counterparts but which do not take place at exactly the same moment. In other respects, 'warranties of quality' ensure that a good has commonly-known characteristics while 'antitrust laws' prevent specific agents from acquiring too much market power. More generally, 'exchange rules' fix the types of goods one is allowed to exchange (excluding, for instance, child labor or human organs).

Some other institutions appear as substitutes to the market. For example, a 'planning mechanism' is used for goods which cannot easily be treated by a competitive market (public goods, goods with increasing returns), the planner playing the role of the concrete coordination entity. Likewise, an 'auction mechanism' confronts a seller and a buyer of a specific good in an asymmetric way, the seller being simultaneously an agent and the concrete coordination entity. Finally, a 'negotiation process' may be directly implemented by the seller and buyer of a good in order to determine its exchange price bilaterally somewhere between its cost and its utility.

The employer and the employee may agree on a first type of employment contract called a 'sales contract'. The employer defines a precise task for the employee with a corresponding wage. Any modification in the task or even in its context corresponds to a new contract with a new wage. A sales contract has a market flavor, since it considers egalitarian agents and tries to define the wage of labor considered as a specific 'good'. The contract is generally incomplete because the task and its context cannot be defined in enough detail. In fact, a sales contract can be considered a basic institution which can be iterated and combined in order to form more complex institutions. For instance, a job market is sometimes considered as the superposing of bilateral sales contracts between employers and employees, coordinated by a common wage for a similar task. But such a mode of coordination involves high transaction costs.

7.4 Knowledge as an internal coordination device

Consider now a set of agents who are members of some economic organization, such as a firm or a union, for example. They enter with their own capacities, but also with their own preferences, which may differ profoundly from the aims of the organization. They only have local information on what is done by the other agents with whom they are in permanent contact. They have bounded cognitive rationality, since they hold local mental models of the structure of the organization. They have bounded instrumental rationality, since they are unable to compute too complex choices. Moreover, they are confronted with several types of costs which appear as frictions. They face information costs when they gather information about their context and reasoning costs when they form a general model of their environment. They face negotiation costs when they design internal contracts and transaction costs when they perform economic operations.

In order to coordinate opportunist agents facing many costs and possessing different information, several internal institutions are established. They tend to promote the specialization of agents in terms of skills and to coordinate them in order to achieve the goals of the organization. They may subject the agents to physical constraints and possibly influence their preferences, especially by inducing some social identity (treated as a type of capital or utility). However, no global equilibrium is generally considered, since the agents are no longer completely free in their actions, being subjected to an external 'authority'. For instance, the 'agency relation' is based on a double asymmetry between the 'principal' and the 'agent'. The principal is able to impose certain rules which create a (psychical or monetary) incentive for the agent to act in the right way. Conversely, the agent has certain private information about the environment which the principal does not possess, but which would be relevant to him.

The main internal institution is the 'hierarchical structure', in which the coordination of activities is achieved through a relation of subordination. This organizational structure is made up of different organizational levels and designates the tasks that the agents have to accomplish at each level. It is in fact constituted of two embedded hierarchies. The 'control hierarchy' is top-down and defines which decisions should be imposed by each level onto the level below. The 'information hierarchy' is bottom-up and specifies which information should be communicated from one level to the level above. The first defines a 'division of labor' concerned with the specification and breaking down of tasks while the second defines a 'division of knowledge' concerned with the filtering and aggregation of information.

The establishment of a hierarchical structure raises three fundamental problems. The first problem is to state what overall authority is supposed to design the organizational structure itself and needs an incentive device. The second problem derives from the fact that the structure is not spontaneously self-enforcing. The third problem involves the multiplicity of available structures, which creates the need for a selection procedure. As a solution to these problems, it is sometimes considered that the highest level of the hierarchy has the power to conceive the hierarchical structure, to stabilize it and even to select it. In particular, it must find a balance between centralization and decentralization in terms of both decision-making and information.

Some institutions are designed to be complementary to the hierarchical structure. For instance, 'corporate culture' promotes not only a common language for the organization, but also common beliefs (so as to have similar interpretations of ambiguous situations). Likewise, 'friendship conventions' define the attitudes expected from members in their interpersonal relations inside the organization. In other respects, some 'communication norms' specify what a member of an organization may transmit about the organization to outside correspondents (duty to preserve secrecy). More profoundly, 'labor rights' determine the rules for hiring and laying-off which are to be followed for any member of the .

Other institutions appear as substitutes to the hierarchical structure. For instance, ‘voting rules’ may be used by the members of an organization for some crucial decisions (especially nominations), even if such decisions need to be prepared beforehand by some sort of authoritative body. Likewise, a ‘team’ is obtained when the members of a group already share a common goal (football team), each agent then acting independently with regard to his local information about the others and the environment. Finally, a ‘network’ simply requires that its members respect certain general rules of participation, each agent then acting autonomously with regard to these commonly accepted rules.

The employer and the employee may agree on a second type of employment contract, called a ‘wage contract’. The employer pays the employee an average wage, but is free to impose a variable task on him, depending on the economic context. A wage contract has a hierarchical flavor, since the employee’s task is completely subject to the employer’s choice and is enforced by sanctions. Such a contract is still incomplete and needs to be interpreted by the agents. A wage contract may be considered as a basic institution which can be iterated and combined to produce more sophisticated institutions. In particular, a company hierarchy is a superposing of wage contracts between agents at different levels, these contracts being coordinated by a prior organizational structure. Such a mode of coordination again involves high transaction costs.

7.5 Information as an exchangeable good

In economic theory, the natural tendency is to consider any exchangeable entity as a good. The category of goods has been extended progressively to include ever more immaterial items: first labor or money, then time or information, finally crime or sex. information is likewise treated as some kind of good, even though it exhibits many features that differentiate it from a material good like a car. information is only partially modular, essentially since it closely links a physical support and a psychical meaning. Obviously, it is possible to define items of information with a fairly independent signification and even to measure their syntactic content in informational quanta (bits). Nevertheless, information is fundamentally holistic, in the sense that different pieces of information are only relevant when they are considered together.

The production of information may be obtained at various costs. It may be very cheap when it results from a natural observation, or more expensive when obtained by means of a complicated apparatus. But its main characteristic is its ‘reproducibility’ once it has been produced. The reproduction cost involves no more than the transcription of its content from one material support to another. Concretely, its production mode depends on the (fixed) costs involved and on the potential users. The production of information is usually public when it concerns global and collectively relevant events (such as meteorological or macroeconomic information). It is likely private when it concerns only local and individually relevant events. Intermediate situations are possible, for instance when public information is treated by private offices in order to adapt it to specific needs.

As concerns its exchange, information may be more or less transferable. It is communicated by means of various material supports, which may be auditory, written, iconic, etc. But its main characteristic is the physical irreversibility of exchange. Once an agent has acquired some information, it is impossible for him to give it back, except through (human or artificial) memory failure. Practically, its transmission depends on the availability of some code shared by the members of an audience. The transmission is public when it uses a common code like natural language, as in general education. It is private when it relies on a specific code, i.e. a shared specialized language, shared background knowledge or shared secret code, as in technical matters. Intermediate situations appear when agents share a common vernacular or technical language.

As concerns its consumption, information may be relevant at different levels. It may be used to satisfy purely intellectual curiosity or to prepare strategic decisions. But its main characteristic is the absence of ‘rivalry’, since the fact that one agent has already consumed it does not preclude its consumption by others. It involves profound externalities, even if a certain degree of control can be exerted on it. Consequently, its diffusion depends on natural and artificial barriers erected to protect it. Diffusion is very large when achieved by mass media or the Web. It is more restricted when directed towards a group of initiated and isolated agents. Intermediate situations are possible, for instance when some TV programs are protected by electronic coding devices.

Despite the specific character of information, it can be bought and sold in an information market. In order to ensure a competitive market, its specific features are attenuated by auxiliary devices or institutions. Free reproduction is forbidden in order to obtain an exclusive good and free diffusion is filtered in order to obtain a good that can be privatized. Nonetheless, some agents may retain monopolistic power over the availability of information and heavily distort market conditions, especially with high storage costs. But then some agents may diffuse information on alternative circuits and re-establish competition outside the official market.

Finally, every agent makes a trade-off between first-hand information directly gathered or inferred and second-hand information purchased on a market. This is especially true for information about the quality of an everyday good, information which is partially reflected in its price. Such an arbitrage may, however, lead to a well known paradox (Grossman, Stiglitz) when nobody initially possesses the information. If somebody buys the information, it is subsequently reflected in the price of the good. Hence, it is in every agent's interest to discover (cost-free) the information from the price, instead of purchasing it directly. However, if everybody does the same thing, nobody has an incentive to acquire the information in the first place, and it will not be reflected in the price.

The employer and the employee are both interested in outside information produced by statistical offices, professional institutions or personal contacts. Both wish to learn about the economic features of the labor market in relation to the availability of workers or the level of wages. The employer also considers the economic conditions of his production sector, as concerns its global activity and the degree of competition. Conversely, the employee looks for more precise information about jobs within his field of competences and about the labor laws protecting his position. Both agents may be sensitive to the reliability of the source of information, depending on whether or not it is first-hand information.

7.6 Knowledge as a factor of production

In economic theory, the tendency is to consider any resource accumulated by an agent as some kind of capital. Following the informal style of sociology, we may speak of symbolic capital or relational capital. In the more formal field of economics, the 'credibility' of the government (concerning public policy) or the 'reputation' of a bank (concerning future interest rates) are considered as capital. Knowledge, in particular, is treated as a form of immaterial capital, as opposed to material capital. It is frequently incorporated into a concrete entity, such as material capital (computer program), an individual (skilled worker) or even an institution (organizational rules).

As with material capital, different forms of knowledge are considered. Firstly, knowledge may be declarative or procedural. Declarative knowledge consists in rules followed by the external world ('to know that'). Procedural knowledge consists in rules modifying the external world ('to know how'). Secondly, knowledge can be explicit or implicit (see 1.6). Explicit knowledge is codified in a language and transmitted through that language. Implicit knowledge is incorporated in an informal way and is only transmitted by imitation and analogy. Thirdly, knowledge may be primary or secondary. Primary knowledge concerns the description of the agents global environment. Secondary knowledge deals with the rules necessary for treating the primary knowledge.

In order to study knowledge as capital, the former is distinguished in a more or less ambiguous way from information. In one definition, information is composed of elementary statements, while knowledge is a structured compound of several statements. This implies that information is declarative, whereas knowledge can be declarative or procedural. In a second definition, information is considered as a transient flow of exchange between two agents, while knowledge is a durable stock incorporated into a single agent. This implies that information is explicit, whereas knowledge can be explicit or implicit. These alternative definitions do not combine well together, since even structured knowledge can be transmitted from one agent to another.

In any case, when it is sufficiently codified and modular, knowledge can be treated as an exchangeable good. Explicit knowledge resembles sophisticated information, and is exchanged as such. Implicit knowledge can only be dealt with by considering the entity in which it is incorporated. On the production side, it is obtained at a higher cost than information, since it is the result of hard reasoning or long experience. Moreover, knowledge is generally more effective in producing other goods than simple information is. On the consumption side, knowledge is consumed more efficiently than information, as it generates increasing returns: prior knowledge helps to build further knowledge. Even more, it generates externalities, since one agent can imitate another who possesses particular skills or farsightedness.

Despite certain special features, knowledge may then be bought and sold on a plain capital market. For instance, patents are bought and sold at a public price between inventors and users. The transaction's originality is that the patent is exchanged on a free market, but its content is protected from public use (inducing a monopoly). Likewise, a football player is transferred between two clubs for a certain price. The transaction's originality is that the player can choose whether or not to accept the transaction and receives part of the price. However, an immaterial capital market remains far-removed from a competitive one, due to increasing returns and externalities. For this reason, many other institutions, such as property rights (royalty rules) or access rights (authorized football transfers), are involved in the transactions.

Like any other capital, immaterial capital can be physically measured in two ways, relating either to its past production or future performance. The past-related evaluation is the sum of all investments which have contributed to its constitution. The future-related evaluation is the sum of all future differential revenues it will induce. For an individual, 'human capital' appears either as the past cost of education or as the future wages induced. For a firm, 'organizational capital' appears either as the past investments in training and design or as the future profits generated. Of course, these two values may differ in the absence of a market for capital, but they coincide when such a perfect market exists.

The employee is endowed with human capital obtained through prior training and past experience and gathered into 'capacities'. He is governed by a personal production function which links his capital and effort to the labor he effectuates. Generally, the capital he brings in or takes out is not priced when he enters or quits the firm. The employer possesses organizational capital represented by the internal structure of the firm and gathered into 'competences'. He has relational capital formed of outside trade relations and gathered into 'reputations'. Moreover, material and immaterial forms of capital can be gathered into 'routines', i.e. programs able to achieve given tasks.

7.7 Role of institutions

The chief aim of institutions is to coordinate the agents' actions in the presence of uncertainty of several types. Firstly, they help the agents to interpret and face their common material environment by explaining and predicting its evolution and by compensating for its effects. For instance, some institutions furnish information about the uncertain physical context and create insurance systems to share the risks involved. Secondly, they help the agents to communicate and synchronize their actions by favoring mutual expectations of their respective behavior and ensuring their compatibility. For instance, some institutions induce predictable, stereotyped behaviors and create a favorable context for agents to coordinate. More precisely, institutions frequently respond to spontaneous failures appearing in both the economic and non-economic spheres.

In game theory, three ‘game failures’ have essentially been considered with regard to usual equilibrium properties. A ‘cooperation failure’ occurs when an equilibrium state is not Pareto-optimal and prevents the players from achieving a mutually better state. A ‘co-selection failure’ occurs when there are several equilibrium states and players cannot agree on one of them. A ‘co-determination failure’ occurs when there is no equilibrium state and players never stop reassessing their actions. These ‘primary failures’ are dealt with by several institutions favoring the players’ information and rationality. But ‘secondary failures’ appear when the players have intrinsically limited information or bounded rationality.

In economic theory, three types of ‘market failure’ can be distinguished with regard to various origins. A ‘technological failure’ occurs when increasing returns or externalities prevent an equilibrium state from being optimal. An ‘informational failure’ occurs when imperfect or incomplete information creates multiple equilibrium states. An ‘organizational failure’ occurs when missing markets or imperfect competition prevents an equilibrium state from arising. Such failures may be compensated for by the state, which imposes certain institutional devices such as public production, a public statistical system, quality norms or antitrust laws. But the functioning of the state is itself liable to ‘government failures’ in the form of bureaucratic costs or the opportunism of civil servants.

An institution acts on agents by modifying their determinants through various channels, often poorly defined. Its influence is causal when it imposes material constraints or designs original relations for the agents. Hence, it simply determines certain physical acts which modify the agent’s environment. Its influence is intentional when it suggests certain representations or motivations endorsed by the agents. Hence, it diffuses beliefs and norms which have to be internalized by the agents. In any case, an institution needs to be accepted by the agents in accordance with their prior determinants, or enforced by incentives and sanctions. It may be consciously perceived or act unconsciously, but even in the first case, it needs to be interpreted by the agents. Finally, when several institutions are relevant in a given context, an agent has to trade-off between their different injunctions.

The influence of an institution is ‘parametric’ when it defines signals which are arguments of the determinants, hence when it only acts on variables which are, in any case, of concern to the agent. The influence of an institution is ‘’structural’ when it modifies the form of the determinants, therefore conditioning these determinants which are no longer exogenous. As concerns opportunities, an institution transferring resources and imposing financial barriers physically constrains the agents' capacities and imposes moral interdictions. As concerns representations, an institution supplies information to the agents (inducing a belief revision) or imposes beliefs. As concerns preferences, an institution influences them through incentives and sanctions or restructures them profoundly by persuasion. However, an institution may bypass the determinants and act directly on behavior rules or strategies. For example, a social norm is an imperative which directly relates a given action to each context, especially when an agent is already assigned to some role.

As a special type of institutions, the cognitive institutions provide to the agents some representations of their environment. It may be scientific theories (for instance Newtonian or Darwinian), pseudo-theories (for instance astrology or intelligent design) or mere ideologies such as myths or dogmas. Together, they form the culture of a society which influences the mental states of the agents and reduce their variance. They appear as a kind of 'collective belief', but stay individual in the sense that they are supported only by individual agents. Their collective character comes exclusively from the fact that they are more or less shared by all individuals. However, a holistic 'social belief' can be assumed to exist in agents' minds. In particular, it is possible that no agent believes it, but that each believes that the others believe it (Boyer, Orléan).

The employer-employee model has been generalized into several models involving institutions. In Spence's model, the employee has an exogenous productivity and acquires education at some cost which decreases with productivity. The employer observes the level of education and, assuming that productivity is proportional to education, offers a wage proportional to the assumed productivity. For low (or high) values of the unitary cost of education, all employees choose a high (or low) education level and the employer's belief is refuted. But for intermediate values of education cost, a low productivity employee chooses low education and a high productivity employee chooses high education. Hence, the employer's belief is self-fulfilling, even if he has got the chain of causality the wrong way round. In Akerlof's model, the employer gives the employee a gift in the form of a wage above the market wage. The employee gives a return gift in the form of more effort under certain circumstances.

7.8 Eductive genesis of institutions

In a first approach, an institution is deliberately designed through a decision process establishing an ‘institutional equilibrium’. On the one hand, it may be chosen by a single planner having the authority to make an individual decision. On the other hand, it may result from a process of negotiation between various agents concerned. Several conditions have to be met in order to implement such a formal choice. Firstly, a list of possible institutions has to be available. Secondly, the consequences induced by each institution have to be known. Thirdly, the agents need to form preferences about the long-term consequences of each institution. Moreover, there must be criteria for selecting the institution if several equilibrium solutions are available. Finally, the institution will have to be reinforced if it is not collectively optimal, which may occur when some of the agents concerned do not participate in the decision process or when the equilibrium state itself is not Pareto-optimal.

However, the deliberate creation of an institution violates the principles of methodological individualism for two main reasons. On the one hand, it requires the existence of a prior (individual or collective) entity whose role is to establish the institution. On the other hand, it requires the existence of a higher-level institution capable of ensuring the institution is respected by means of incentives and sanctions. An alternative approach consists in the spontaneous genesis of an institution. Since it is impossible to obtain an entity ex nihilo, the institution is considered only through its effects on players’ actions. More precisely, an institution is identified with an equilibrium state. But this drastically limits the types of institution that can be considered to those which can be expressed in terms of behavior rules. An institution is then a self-enforcing set of behavior rules in the sense that, once established, it is not to anybody’s advantage to depart from it. As a corollary, the genesis of an institution is identified with the genesis of an equilibrium state.

The genesis of an institution can be conveniently studied in a game theory framework. The main reason for this is that there are no prior institutions in a game, except the ‘rules of the game’ already incorporated into the players’ determinants (see 5.1). Eductive justifications of equilibria assume that the players are able to coordinate on an equilibrium state by means of their reasoning alone (see 6.8).By transposition, eductive justifications of institutions assume that the players can agree on an institution solely by means of their reasoning. The existence of the institution is then founded on common belief about different characteristics of the system. But many institutions can be obtained on that eductive basis.

Lewis (Lewis, 1969) gave an early eductive definition of a 'convention', i.e. a specific institution appearing in pure coordination games (where several equilibrium states are equivalent for the players). A convention is a regularity R in some social system, in a recurrent situation, when the following conditions (which we have ordered differently to Lewis) are common knowledge. Firstly, each player prefers conformity to R to less general conformity. Secondly, each player has a decisive reason to conform to R if the others do. Thirdly, all players conform to R. Fourthly, R is not the only regularity which satisfies the preceding conditions. These conditions can be reformulated in terms of the eductive justifications of equilibria and precisely characterize a Nash equilibrium state. The first and fourth conditions indicate that the game in question is a coordination game, the structure of which is common knowledge. The second states that the players' rationality is common knowledge and the third that the conjectures are common knowledge.

Among the eductively sustained institutions are the cognitive institutions when they are 'self-fulfilling collective beliefs'. A collective belief asserts the existence of some (deterministic or random) regularity relating collective and environmental variables. Such a belief is self-fulfilling if, when it is taken into account by the agents in their choices, it produces the regularity it asserts. Hence, a self-fulfilling belief may not be true for all values of variables, but only for equilibrium ones. A necessary condition for a belief to be self-fulfilling is that all agents hold that same belief. However, two groups of agents may hold respective (random) 'collective beliefs' which are both self-fulfilled, even if an accurate scientist is able to state that something is wrong.

Indeed, several levels of complexity can be considered for a self-fulfilling belief. On the first level, it simply relates an endogenous variable to an exogenous one. For instance, if all agents believe that prices depend on sunspots, the choices they make in consequence actually create the postulated relation. On the second level, it relates an endogenous variable to many explaining factors. For instance, if all agents use the Black & Scholes formula, which determines the price of a financial asset, it becomes self-fulfilling. On the third level, it may constitute a whole theory about the economic system. For instance, if all agents believe in Keynesian theory, it may be realized through their actions. In all cases, a self-fulfilling belief selects one equilibrium state from among a whole set of them.

The employer-employee contract is also framed by many other institutions and especially conventions. These conventions, for instance money and language, precisely obey the emerging conditions stated by Lewis. Each agent prefers that more and more agents accept money (or language); he accepts money (or language) if the other agents do so; it is to each agent's definite advantage to accept money (or language); several variants of money (or language) are conceivable. Moreover, money may give rise to a 'social' belief. For instance, all agents may believe that money is under-valued, but each believes that the others believe it to be over-valued.

8

Evolution of the economic system

Knowledge leads to unity,
ignorance to diversity.
Ramakrishna

The economic system evolves as the result of its struggle against various forms of uncertainty generated exogenously by its material environment and endogenously by its institutional one. It is transformed by a process of co-evolution between its physical sphere and its psychical sphere, in which boundedly rational agents follow self-organizing mechanisms. In particular, the transmission ofinformation and the accumulation of knowledge explain the emergence of social phenomena within organizations or in the economy as a whole. institutions, for example, may be generated by dynamical processes of learning and evolution governed by agents endowed with specific heuristics and meeting in neighboring interactions.

Following the modeler, who follows the economic system changes over different time scales (8.1), the agents try to capture a more local evolution in their own knowledge (8.2). markets evolve as new commodities and new relations are created (8.3), while organizations evolve when new technologies and new modes of governance are discovered (8.4). The diffusion ofinformation is essential to the evolution of a financial market (8.5), while technological innovation is preeminent in the evolution of the firm (8.6). Lastly, institutions appear as emergent structures in some evolutionary processes (8.7), and are more generally created and destroyed during an original life cycle (8.8).

8.1 Evolution of the modeler's model

The economic system evolves over time, in its different manifestations. Time is generally considered as an extra-economic and exogenous variable supporting the evolution of the system. It is considered as continuous for many theoretical models, since most phenomena display a great deal of inertia. Economic growth, for example, is quite regular, even if some accelerations and decelerations are observable. But time can be discrete when exogenous phenomena create natural periods which influence economic operations. Agricultural production, for example, follows annual climatic and vegetative cycles, and market prices fluctuate accordingly. Time may also be discrete (and even endogenous) when economic events create discontinuities. For instance, the successive oil crises induce successive regimes in the economy.

The main properties of basic economic entities may change over the short-term, shifting from one class to another in the basic taxonomies. Goods evolve in terms of their quality through technological and social innovation. The technical or esthetical characteristics of cars, for example, are continually being modified. Agents see their determinants modified, due to exogenous factors, past experience or age. This is especially true for preferences, which vary at long term as to the relative weight attached to partial criteria, discount rates or aspiration levels. For instance, a producer may change his preferences following the discovery of new technological opportunities and a consumer may change his after experiencing a new product. institutions change their nature and even their function. money, for example, successively adopts different supports while keeping its role as a means of transaction. Relations are redistributed as regards their configurations and supports. For instance, coalitions between airlines are reconsidered and reshaped in changing circumstances.

The basic entities also evolve through the creation of new kinds and the extinction of old ones, giving rise to new taxonomies. New sorts of goods become available while old ones disappear. For instance, new labor qualifications are defined, traditional craftsmen being replaced by computer specialists. New types of agents enter the market while others exit. For instance, temporary employment agencies are appearing while traditional unions disappear. New kinds of institutions are created or result from the splitting or unification of old ones. For instance, new financial markets and new auction mechanisms are set up while old tax systems are reshaped. Finally, new forms of relations appear while old ones are abandoned. The web, for instance, has created a completely new system of relations on a worldwide scale.

The evolution of the economic system is subject to nested time scales, since some variables adapt faster than others. For instance, institutions are more stable than economic agents, economic agents are more stable than their determinants, and their preferences are more stable than their representations. The time scales are generally associated with specific relations between variables. The slow variables influence the fast variables over the short term, while the fast variables shape the slow variables over the long term. For instance, a consumer chooses the goods he buys with reference to his present preferences, but the experience of consumption modifies his preferences over the long term. Even more, fast variables establish short-term equilibria while slow variables establish long-term equilibria. For example, producers determine their production levels on a short-term basis, but adapt their goods, technologies or prices over the long-term.

For the modeler, the transformation of entities is generally attributed to explaining factors which may be either causal or intentional. For instance, new means of transportation act on economic activity in a causal way, while new technological devices act on a firm's structure in an intentional way. Similarly, the explaining factors may be extra-economic or economic. For instance, new consumers appear for purely demographic reasons while new producers appear when the economy offers opportunities for profit. Globally, all factors act together in a systemic way and contribute to the production of economic effects regulated by positive or negative feedbacks.

A special feature of evolution is the existence of 'emergent phenomena' arising at a social level. An emergent phenomenon is a phenomenon which looks surprising to the modeler in relation to the basic entities, but may nevertheless be explained. emergence is synchronic when it results instantaneously from the basic entities and diachronic when it appears progressively. It is unidirectional when it results solely from the bottom-up influence of the basic entities and bi-directional when there is a top-down feedback on the basic entities. An emergent phenomenon can be of different kinds, such as a statistical distribution of entities, a relational network between entities or even new entities. For instance, consumer income tends to follow a Pareto law, cartels of producers are progressively formed or new labor institutions are created.

In the employer-employee example, agents' preferences are adjusted over the short term according to the easiness of finding a job. If the employer can easily find another employee prepared to work at the given wage, he lowers his reserve wage, and vice versa. If the employee can easily find a new position, he increases his reserve wage, and vice versa.

In doing so, both agents face low adjustment costs. Over the long term, more profound transformations are taking place. New types of agents like unions appear with the aim of mediating the relation between the supply and demand of labor. New institutional rules are expressed, especially as concerns the hiring and firing conditions of workers by firms.

8.2 Evolution of the agent's knowledge

From the agent's point of view, the economic system is evolving over a personal, subjective timescale. Subjective time is less homogenous than physical time, since it is concentrated around the present time and its reference point is mobile. As concerns past events, they are integrated into the agent's memory. It is frequently stated that there exists a discount rate such that events are considered less and less as their distance in the past increases. For instance, a consumer values his most recent experiences of a good more than his older experiences. As concerns future events, they are integrated into his prospective mind. Again, it is said that there exists a discount rate such that expected events are considered less and less as their distance in the future increases. For instance, a firm considers more the short-term than the long-term effects of a certain investment.

An agent receives newinformation through different channels, about different variables and at different times. He experiments passively wheninformation is just a by-product of his actions, for instance when he observes another's past purchases or when he experiences the satisfaction induced by a newly-tested good. He experiments actively when he performs specific operations with the aim of obtaining information. For instance, a consumer visits various different shops to compare the prices of a good he wants to buy. As usual, he may trade off between exploration for newinformation and exploitation of existinginformation, even if the trade-off is not optimal. As another example, a producer may vary the price of his product through successive adjustments in order to learn the demand function he faces.

An agent modifies his structural knowledge at short term by using belief revision rules. He may simply adjust the parameters of a model of his environment in keeping with his observations. For instance, a consumer may discover the quality of a good for food by observing the demand of another consumer who knows the quality. However, he may hold his theories for a long time before observing that they are refuted. For instance, a consumer may believe for years that the price of some high technology good is regularly decreasing before he observes he is wrong. More profoundly, an agent may define a model of his environment by means of abductive reasoning from data. For instance, a firm may discover the behavior function of some other producer in order to adapt or even to imitate him. Of course, the revealing process is still ambiguous and strategic considerations are involved in it. For instance, if a firm learns that its opponent is employing more workers in a depressed economic climate, it then has to interpret such behavior.

Finally, an agent modifies his expectations by changing his expectation rules. An expectation rule may be based on a more or less crude model of his environment. For instance, a firm forecasts the future price of oil by means of a sector-based model simulating an equilibrium between supply and demand. Due to bounded rationality, the expectation rule may directly relate the expected variable to its past values. For instance, a consumer predicts future prices by means of an adaptive rule, stated in order to reduce the forecasting error. In general, several rules are used simultaneously by different agents to forecast the same variable. For instance, on a financial market, if 'fundamentalists' predict the future price of an asset with reference to its future returns, 'chartists' use rules based on regularities observed in the past.

The agents consider that the evolution process obeys the same types of laws or models as the modeler does. However, the agent is induced to distort or simplify certain explanative schemes. Firstly, he is 'egocentric' in that he attributes any change to himself, to nearby agents or to their common context. For instance, a producer considers that a new technology has been obtained by his own research, by neighboring firms or by academic laboratories which are out of his control. Secondly, he is 'myopic' in that he considers that the slow variables are fixed and that only the fast variables are evolving. For instance, a consumer considers that the prices he observes are exogenous even if he knows that he has some (small) influence on them. In particular, an agent generally considers emergent phenomena as natural phenomena that he cannot influence.

Globally, like his material capital, an agent's immaterial capital evolves in different ways. Firstly, knowledge is increased by incorporating successive pieces ofinformation into it. For an individual, immaterial capital develops through education or training, while for a firm, immaterial capital develops through research and development or in-house training. Secondly, knowledge is enriched by the autonomous internal reasoning performed on it. For an individual, knowledge is transformed by deduction or induction processes; for a firm, knowledge is transformed by redesigning its organization scheme. Thirdly, knowledge can shrink through some kind of cognitive obsolescence. For an individual, knowledge disappears through memory failure; for a firm, knowledge disappears through the loss of skilled agents or the inaccessibility of artificial memories.

In the employer-employee example, over the medium term, their information is modified by deliberate search. The employer looks for new workers prepared to work in the existing jobs for lower wages. The employee looks for jobs outside the firm for which he would be better paid. Each agent conducts his search in a neighborhood and may even limit his search to a sample of that neighborhood. In doing so, he faces relatively high search costs. Over the long term, informational or mediation devices may appear. For instance, employment agencies may be created to diffuseinformation about available jobs and so favor the adjustment of supply and demand.

8.3 Evolution of markets

In a one-period competitive market, the concrete process of price formation is not precisely defined. The Walrasian auctioneer is assumed to fix the prices as a unique signal by following a 'Walrasian tatonnement process' in fictitious time. He increases the price of a given good when its total demand exceeds its total supply and vice versa. But real transactions only take place once the equilibrium prices have been established; hence the material sphere and the cognitive sphere are disconnected. Such a process is quite demanding in terms ofinformation, since all supply and demand must be known at each period. Moreover, the process does not converge in all cases towards a competitive equilibrium state. In a multi-period competitive market, the process of price formation is even more complicated, since the Walrasian auctioneer has to define the prices of all goods in all future periods.

The price formation mechanism is no better defined in a one-period market with imperfect competition, whether the adjustment is achieved mainly in quantities or in prices. A 'Nash regulator' is assumed to fix the quantities and prices by following a 'Cournot tatonnement process'. This regulator observes the quantity offered by one duopolist and sequentially computes the best response of the other. The process is again demanding, since the regulator needs either to observe or to compute the best response functions. Moreover, the convergence of the everlasting process is not guaranteed. However, such a process may be followed in real time by the agents since the price is now fixed by them and not given to them by an outside entity. But these agents are then very myopic in that, in each period, neither agent considers that the other will later react to his action.

A preliminary step towards a more realistic view is to consider that the agents may learn about structural characteristics of the system in which they act, even if they are still coordinated by a fictitious entity. Learning is generally epistemic, since the agents have prior beliefs about their environment, which are revised with reference to new observations. For instance, in an imperfect market, a duopolist may have a prior belief about the demand function which he adjusts with reference to past observations, using least squares or other statistical methods (like the modeler). Likewise, in a competitive market, a consumer may revise a prior belief about the relation between the price and certain exogenous factors he observes.

In fact, the learning process unfolds in a non-stationary environment, since all the agents are learning simultaneously. Nevertheless, the agents' beliefs generally converge towards a reduced form of the actual model (that of the modeler). Some relevant variables may be missing because they are not initially considered by the agents. Such an asymptotic model is only locally rational, since it proves to be true at the equilibrium state but not elsewhere. For example, the duopolist converges towards a reduced demand function which depends only on his own price and not on the other's. Likewise, the consumer converges towards a reduced price function relating the price exclusively to the considered environmental factors.

The main step towards more realism is achieved when the agents define their prices and quantities in each period and implement them simultaneously, without interference from any outside entity. Learning becomes frequently behavioral, since the agents adapt their actions directly to their past observed performances without expectations. For instance, in an imperfect market, a duopolist may use an original learning rule, the 'stubborn rule', which applies only when the action space is one-dimensional. He increases the price of his product if, in the preceding period, he increased his price and got an increased profit or if he decreased his price and got a decreased profit. Likewise, in a competitive market, a consumer and a producer may propose their own prices (adapted to past observations) and the transaction takes place at some intermediate price if the announced prices are compatible.

In fact, the learning process acts as if the Walrasian auctioneer or the Nash regulator were distributed among the agents. The process converges, under certain standard conditions, towards the equilibrium prices. However, the prices remain dispersed among the agents if theinformation, negotiation or transaction costs are too high. For instance, the duopolists converge towards the Cournot equilibrium state with various learning rules, but they converge towards the collusion equilibrium with the stubborn learning rule. Likewise, producers and consumers often converge towards the competitive equilibrium price system. More profoundly, the learning process may lead to the design of an endogenous network among agents. For instance, in a fish market, buyers and sellers may progressively establish lasting relations of loyalty (Weisbuch, Kirman).

In the employer-employee example, different adjustment rules are available, inducing various transaction costs. For instance, if an employer finds a worker who is prepared to accept a lower wage, he asks the current employee if he will work at that wage, and if the employee refuses, he replaces him by the other. When considering many pairs of agents, in each period, each one searches for new partners, changes or keeps his partner according to the above rule and adjusts his reserve wage accordingly. Over the long term, wages converge towards the equilibrium wages when there are no costs of any kind. When there are search costs, on the contrary, the process may converge towards a segmentation of prices in several areas. Likewise, when there are transaction costs, the prices may remain within a certain interval without being unique.

8.4 Evolution of hierarchical organizations

In classical microeconomics, the firm is treated as a compact entity with its own determinants. The evolution of the firm then coincides with the evolution of its determinants, under the influence of purely exogenous factors. The preferences are in fact invariant, since the firm is still assumed to maximize its (long-term) profit. Opportunities are transformed by technological change, without the need to specify its origin. beliefs are passively adapted to change in the physical or institutional environment. Moreover, even when the firm is considered as a set of agents regulated by internal institutions, its evolution is studied as a pure exogenous transformation of these institutions.

More recently, the evolution of the firm, still considered as a single entity, has been studied as a fairly deterministic evolutionary process (Nelson, Winter). The genotype is constituted of the ‘routines’ of the firm, which gather together skilled workers and equipment with the aim of implementing specific tasks. The phenotype is constituted by the firm itself acting in an economic environment. The ‘transmission process' is simply the perpetuation of routines. The 'mutation process' consists in innovations acting on routines. The 'selection process' is likened to the competition process between firms. Hence, through time, the best-adapted routines are developed and selected through the pressure of competition. Correlatively, the firms adopting those routines are themselves selected. Such a process is generally considered as a mere biological analogy, rather than an extension of biological evolution into the domain of the social sciences.

A first step towards a more realistic view consists in studying the evolution of the firm, still considered as a single entity, as a standard learning process. Reference to biology is only made at a generic and metaphorical level. The mutation process is replaced by a search process looking for original opportunities. The selection process is replaced by a filtering process eliminating the less efficient opportunities. The support of the learning and mutation processes may be the firm itself, the firm's determinants (beliefs, preferences) or, more directly, the firm's strategies. The learning process is epistemic when beliefs about demand or competitors are progressively corrected. It is behavioral when some strategies (or even beliefs) are reinforced more and more.

The asymptotic states depend on the precise evolutionary process and on the given environment. For instance, firms adopting good technologies are reinforced while firms adopting poor ones are inhibited. Likewise, firms adopting efficient beliefs develop while firms adopting poorly-adapted beliefs are eliminated. A special case concerns the evolution of the firm's decision rules. The 'Alchian thesis' asserts that, in a learning (or evolutionary) process applied to firms, only the maximizing firms are selected. But this assertion is contradicted by many analytical results (Dutta, Radner). In fact, some non-maximizing firms can survive in specific niches even in the presence of maximizing firms. Even more, non-maximizing firms may adapt better than maximizing ones to some rapidly evolving environments.

A further step towards realism consists in studying the decentralized evolution of the firm. The firm is formed of several units organized into interrelated levels and operating in an environment where demands for different goods are expressed. Each unit follows its own learning process under constraints or incentives imposed by the upper units or the environment. The learning process essentially concerns the units' behavior rules, which change through time. In particular, each unit modifies its behavior rule as a function of past performances (learning curves). The learning process may also concern the design of a network linking the units between themselves and with the environment. In particular, some links are created while others fall into disuse.

The process converges, in a given environment, towards types of structures interpreted as emergent phenomena. For instance, certain distributions of tasks or certain distributions of beliefs among units may emerge. More often, a specific organizational scheme may appear. For instance, a centralized structure (where lower units communicate only with upper units) results more often in a stable environment with homogenous agents, a decentralized structure (where lower units communicate directly with the environment and neighboring units) in a turbulent environment with heterogeneous agents and a hybrid structure (with both types of communication) in a regularly fluctuating environment (Marengo, 1992).

In the employer-employee example, it is possible to consider that the decision process works on two structural levels. On the second level, the employer chooses between a sales contract and a wage contract. On the first level, he implements the sort of contract he has chosen. In dynamics, the changes at each level are correlated to corresponding time scales. The first level evolves over the medium term through an epistemic learning process involving the employees type, which becomes more precise with time. The second level evolves over the long term through a behavioral learning process where each sort of contract is judged according to its past performances.

8.5 Financial contagion

In a financial market, where several traders are buying and selling a given asset, each trader has different sources ofinformation about its value. Firstly, he has private information about the objective value of the asset, defined as the future discounted income that it generates. Secondly, he acquires public information in the form of the price of the asset, assumed to summarize all the traders' evaluations of it. Thirdly, he observes the supply and demand expressed by the other traders, reflecting the precise value they privately attribute to the asset.information is therefore multiform, since it combines private and public components, and asymmetric, since the traders receive different private signals.

Relying on factual information, the trader may abduce structural information about the others' evaluations of the asset. In multilateral trade, when observing a high demand, a trader may infer that good news has been communicated to others. But he may also conclude that the traders assume that the others assume that the others have received good news (when in fact they have not). So the abductive process is rather complex and ambiguous and leads the traders to form crossed expectations about their respective evaluations of the asset. Especially, it is possible that the traders reason at a given level of expectation and that they assume the others reason at a lower level. This means that each trader considers that he is cleverer than the others, which is in fact contradictory for the modeler.

Finally, the trader forms expectations about the future price of the asset in order to determine his own supply and demand. However, the diffusion ofinformation leads to an extinction of all transactions, at least if the traders have the same inter-temporal preferences. In a bilateral trade, for example, when observing that the other wants to trade, a trader may infer that his opponent has more objectiveinformation than he has and therefore refuses to trade. More generally, if fullinformation is known by all the traders, no trader has a specific opportunity for gain and no transaction takes place. Hence, traders will only exchange if they have different attitudes towards risk or if they think they have moreinformation than others.

As a result of all the informational factors, the price of an asset can be additively broken down into two parts, at least by the modeler. The ‘fundamental’ is the objective value of the asset and reflects the technical opportunities and economic preferences of the traders involved with the asset. The ‘bubble’ reflects speculative motives due to the asymmetry ofinformation between traders. In fact, we can distinguish between three different kinds of speculation (Bouleau, 2003). Material speculation occurs when some traders have private information about the properties of the assets, especially their future returns. Psychological speculation occurs when some traders have private information about the others’ beliefs or intentions. Computational speculation occurs when some traders use better reasoning capacities in forming their beliefs and expectations.

When the agents hold rational expectations in a competitive market, it can be proved that no bubbles appear. In order to explain bubbles, it is necessary to introduce imperfect information or bounded rationality of the traders. The communication structure may be defined by an incompletely connected network, where some leaders may have more influence than other traders and act as gurus. The communication process may show strategic information transmission where the trader may retain, blur or distort information in order to prevent it being used against him. The traders may expect the future price of an asset by an adaptive expectation rule which tries more or less to find regularities in its past evolution. The traders may consider that the prices are influenced by either of two gurus, the observation of prices reinforcing or inhibiting the role attributed to each.

Asymptotically, the asset price is more or less efficient with regard toinformation introduced into the system and opportunities left to the traders. A market is said to be 'informationally efficient' when it incorporates well the availableinformation. It is 'weakly efficient' when it incorporates only the public information and 'strongly efficient' when it incorporates both the public and private information. A market is said to be 'efficient by arbitrage' when it leaves no possibility of profits for the traders. Moreover, the asset price depends on the horizon of market transactions. With a finite horizon, bubbles generally disappear when approaching the horizon and the asset price becomes equal to the fundamental. With an infinite horizon, a bubble may be maintained throughout, and the asset price converges towards a 'conventional' value, based on exogenous events. This price reflects what the average opinion expects average opinion to be (Keynes, 1936).

In the employer-employee model, some reputation phenomena may be considered. In each period, the employer chooses whether or not to hire the employee, and if he hires him, the employee may choose to make a high or low effort. In fact, the employee can be of two types, a 'normal' type who may or may not work hard and a 'nice' type who always works hard. When the game is not repeated, the employee (of normal type) makes a low effort, and is therefore not hired. When the game is finitely repeated, the employee develops a reputation of nice type by always making a high effort until the game approaches the horizon, where he makes a low effort with some probability and gets fired once he makes effectively a low effort. The situation may be reversed and the employer may develop a reputation of 'loyalty' towards the employee.

8.6 Technological innovation

Two sources of innovation are generally considered with reference to the frontiers of the economic system. Firstly, an innovation may come from outside, from independent research centers, for example. This is especially true for fundamental innovations concerning new commodities or new technologies. The innovation process is then fairly independent from the economic conditions, except for the budget and organization allocated to research. Secondly, an innovation may come from inside, particularly for large firms with independent research departments. This is especially true for applications resulting from 'learning by doing' or 'learning by using' processes. The economic conditions then work more in favor of the innovation process, since they condition the future of the firm.

The production of an innovation by a laboratory or a firm, if it appears at first as 'problem-solving' (Dosi, 1988), displays quite specific features. It involves entangled physical, human and organizational factors acting in a random way. It is relatively irreversible, since an innovation cannot be deliberately suppressed once it has appeared, except by becoming obsolete. It obeys the law of increasing returns, since the discovery of an innovation frequently favors the production of new ones. Hence, the innovation process defines clusters of innovation giving rise to innovation paths. It results in a shapeless and functionless product which needs to be adapted to a specific context in order to become efficient.

The adoption of an innovation by the originating firm or by others follows more classical principles. It is essentially governed by the future differential profit that the innovation induces, even if other factors come into play. It is subject to physical constraints related to the need for a firm to ensure homogeneity in the quality of the goods and compatibility of the technological standards. It involves positive externalities, since the profit induced by an innovation depends crucially on its adoption by other firms. It frequently leads to preferential mimetism, since when one firm adopts it, other firms in the same sector have a strong incentive to do the same.

The diffusion of an innovation can follow one of two patterns, depending on its institutional protection. If the innovation is not protected by a recognized patent, it can freely diffuse from one firm to another. However, it must be stated in a sufficiently codified form to be transferred, and even then it picks up bias along the way. This sort of transmission presents positive retroactions due to increasing returns and positive externalities. If the innovation is protected by a patent, it temporarily provides the inventor with a rent and is only diffused later, at some cost. It is automatically codified because it is precisely described in the patent, and it is therefore transferred faithfully. Positive retroactions are offset by negative ones, since other innovations are sought in an effort to compete with the existing one.

Globally, each innovation follows a life cycle from its initial discovery through to its replacement by another. Its quantitative evolution often follows a logistical law, with a slow start due to high fixed costs, a fast development once its efficiently has been recognized and a slow end when it can no longer be improved. Its spatial evolution depends on the network of firms, and leads either to the homogenization of a single dominant innovation or to the segmentation of different innovations in specific niches. Its qualitative evolution is path-dependent (because the first adaptors are relatively random), exhibits bifurcations (because innovations evolve differently in different contexts) and may lead to the lock-in of a second-best innovation.

From a collective point of view, the assessment of an innovation is now subjected to the 'principle of precaution'. But such a principle, generally expressed in a loose literary way, remains ill-defined in terms of decision theory, as concerns the environment law or the consequence law. Any innovation, especially a new technology, is the source of unexpected contingencies which first have to be conceptualized. Qualitatively, the decision-maker may be affected by unawareness (see 1.4), since he does not know the consequences of the innovation and does not even know that he does not know them. Quantitatively, the decision-maker considers not only random scientific laws governing the impact of the innovation, but also second-level uncertainty about the first-level laws. Only when the hierarchical uncertainty has been well grasped can a relevant choice rule be suggested to the (generally collective) decision-maker.

In the employer-employee problem, new technologies may modify the relations between the agents. The employer has to consider not only new products but also new tasks, so as to produce his good more efficiently. The employee is assumed to be able to perform these tasks if asked to do so. More generally, new technologies have an impact on the surrounding firm, as is clearly shown in its asset price. The fundamental reflects the cognitive capital of the firm, which incorporates the potential of innovations it may create or import. The bubble is influenced by fashion effects related to new goods or technologies and fails to consider their long-term effectiveness.

8.7 Evolutionary genesis of institutions

The genesis of institutions has already been studied in the context of game theory, by identifying the institution with an equilibrium (see 7.7). The eductive justifications of an equilibrium state were transferred into eductive justifications of an institution. In parallel, the evolutionist justifications of an equilibrium state are transferred into evolutionist justifications of an institution. The evolutionist view automatically satisfies the selection problem since, when several institutions appear, one is selected by initial conditions and exogenous historical factors. But the genesis of an institution is nevertheless history-dependent and context-dependent (with hysteresis effects) and it may lead to the lock-in of the process in a sub-optimal institution.

The evolutionist genesis of institutions was already suggested by Hayek (von Hayek, 1973) who considered an institution as the result of human action, but not of human design. Especially, the market is the mechanical output of a (cultural) evolutionary process, a 'group selection' mechanism favoring the market among the possible institutional structures. A similar view was developed by Elster (Elster, 1989) who considered that even if an institution has a favorable impact on the agents, this was not the aim of the agents' actions. An institution is reinforced as the result of a causal feedback from its consequences to the agents' actions. Both authors assert that an institution is the unintended result of the actions of bounded rational agents, hence is itself limited, but neither developed a precise evolutionist process capable of explaining its genesis.

All the standard learning and evolutionary processes are candidates for explaining the genesis of an institution. In belief-based learning, the agents progressively adapt their beliefs about each other and the process converges towards behavior rules or even expectation rules interpreted as institutions. In one illustration (Young, 1998), the crop share between a landlord and a tenant farmer converges, in a modified fictitious play process, towards a 50/50 proportion. In reinforcement learning, the agents progressively adapt their strategies according to their past performances and the process converges towards strategies interpreted as institutions. For instance (Sethi, 1999), agents who exchange the goods they produce can progressively select one good as a favored medium for exchanges (see the example below). More recently, different classes of models have become available and apply preferentially to specific types of institutions.

Models of diffusion are relevant for cognitive institutions such as collective beliefs or expectations (for instance risk perception). They assume that a representation is progressively diffused among the agents through a process of informational contagion. For example, each agent is endowed with a degree of acceptance of the representation. This acceptance increases when the representation appears reasonable in the light of his background knowledge and when it is shared (to some degree) by a sufficient number of his neighbors. Of course, the representation has first to be introduced by an agent with a specific motive for envisaging and accepting it, and this agent must be recognized as an opinion leader.

Models of interaction are relevant for behavioral institutions such as collective habits or social norms (for instance reciprocity rules). They assume that a behavior rule is progressively favored by the agents through a process of reinforcement. For example, each agent is endowed with a degree of adherence to the behavior rule. This adherence increases when the behavior rule provides the agent and his neighbors with a performance which is sufficiently high (above a certain threshold) in various contexts. Of course, the behavior rule first has to be conceived and tested by an agent acting as a more or less disinterested social experimenter and, moreover, recognized as a policy leader.

Models of coalition are relevant for organizational institutions such as interest groups or functional associations (for instance trade unions). They assume that a social group is progressively extended due to the positive externalities between its members. For example, each agent is endowed with a degree of adhesion to the social group. This adhesion increases when enough agents accept to adhere to the social group or when the agents already adhering achieve good enough performances. Of course, the social group first has to be started up by some far-sighted agent acting as a ‘germ’, hence able to mobilize other agents even if the group is not yet a profitable one.

In the employer-employee example, the counterpart of the employee’s work can be a certain amount of the firm’s product. If such a barter exchange between both agents is conceivable, the exchange is generally a monetary one, as money is already present. In this respect, the genesis of money can be modeled in an economy with three (order) agents and three goods. Each agent produces one of the goods and consumes the good of his predecessor. He can only store one good and storage of a good is costly. Agents need to combine their exchanges and follow an evolutionary process. Among the multiple asymptotic states of the exchange process, one equilibrium state emerges where the good with the lowest storage cost is selected as the only means of payment.

8.8 Naturalization of institutions

The game-theoretic justifications of the genesis of an institution assume that it can be identified with an equilibrium state. However, they do not explain how an institution detaches itself from the underlying equilibrium state to become a separate entity. The ‘naturalization problem’ concerns precisely the exteriorization of the institution as a separate entity. Its solution proceeds in three stages. The ‘recognition stage’ details how the institution is no longer considered as a regularity stemming from agents’ actions, but appears as an autonomous entity with its own signification and behavior. The ‘focalization stage’ studies how the agents cease to react to the others’ signals and respond directly to the institutional signals, by disconnecting from their neighborhood. The ‘legitimation stage’ describes how the institution receives social endorsement and is accompanied by various enforcement devices even if it is theoretically self-sustaining.

As concerns epistemic justifications, the naturalization problem is partially integrated into the eductive process itself. Recognition is rather obvious since, reasoning in a static way, the agents have to recognize the institution at the same time as the equilibrium is attained. Focalization happens when the agents break the specularity of crossed beliefs and react to the equilibrium summarized in a limited set of signals. Legitimation is necessary when the drastic assumptions sustaining the equilibrium (transparency of situation, rationality of players) are not perfectly satisfied. In particular, for cognitive institutions, naturalization consists in translating a (shared) private representation into a public representation, the latter subsequently being assimilated by the agents (Sperber, 1996).

As concerns evolutionist justifications, the naturalization problem becomes a more independent problem. Recognition happens when certain privileged agents become aware of a regularity interpreted as an institution and diffuse thatinformation around them. Focalization happens when the agents cease the mutual adaptation adaptation of their behavior in order to respond exclusively to the recognized institution. Legitimation is necessary when the assumptions leading to an asymptotic equilibrium (anonymous meetings between agents, sufficient cognitive and instrumental rationality) are not perfectly satisfied. In particular, for behavioral institutions, naturalization consists in transforming a private, spontaneously-adopted behavior rule into a social norm which is proposed and even imposed on the agents.

Of course, in an evolutionist view, the naturalization problem is relevant not only for the emergence of a new institution, but for any change of institution. When the economic context changes, a new equilibrium may appear, either a state which was formerly not an equilibrium or which was already an equilibrium and is now selected (Boyer, Orléan). Recognition is simply the progressively diffused observation that a new regularity is emerging in place of the old one, either incrementally or in a more discontinuous way. Focalization is a switch in the behavior of agents who tend to respond to the new institution rather than the old one after a transition phase in which both were present.

Each institution has its own life cycle, which displays a high degree of irreversibility: it is created, becomes operational and finally disappears. This evolution is asymmetrical, since an institution emerges rather slowly, has a long stationary state and disappears rather rapidly. The reasons why the institution lasts may be very different from the reasons why it was created, since it may fulfill new functions with the same content. Likewise, the reasons why it disappears are not symmetrical to the reasons why it was created, since it is generally supplanted by another institution rather than just losing its function. In any case, the institution generally displays robustness, since it endures for some time even when its context is modified.

All institutions co-evolve, since they simultaneously compete with and complement each other. Especially, cognitive institutions contribute to shape more normative institutions and the last favor or constraint the diffusion of the first. Especially, when one institution falls, many other institutions are likely to fall at the same time, through a ‘domino effect’. In principle, however, institutions situated on different structural levels evolve over different time scales. So, when a high-level institution dies, low-level institutions may nevertheless survive. In other respects, all interacting institutions tend to share common features since the possible institutional designs are limited. If an institution plays a certain role in a given context, it is frequently imitated and slightly adapted to another context. Even if two institutions fulfill different roles, the features of one generic institution may be analogically transcribed to another.

As concerns money (or language), the three naturalization stages are well-defined. The recognition stage consists in the agents’ awareness that one good (or language) is becoming used as a privileged means of payment (or communication). The focalization phase consists in the agents considering that means of payment (or communication) as a normal one, while forgetting its origin. The legitimation phase consists in establishing an official definition of money (or language) given by a political authority who fixes the rules to be applied in using it. Of course, when some instance of money (or language) is adopted somewhere, it is also adopted in other places.

Conclusion

To stop playing
is to stop living.
M. Felinto (1982)

Cognitive economics is founded on a renewed *ontology* concerning the economic system, which adds two original dimensions to classical ontology. Firstly, cognition is explicitly introduced, since the modeler enters now the black box of the agent's mind and describes his choice process which compares virtual worlds. Secondly, temporality is more precisely introduced, since the modeler details the nexus of agents' interactions and characterizes their mutual adaptation mechanisms associated to differentiated time scales. Globally, the economic system is seen as the mutual and dynamical adaptation of heterogeneous actors endowed with limited capacities and tied by permanent networks, hence neither pure spirits nor pure automata.

The traditional 'rationalist approach' is replaced by a 'heuristic approach'. The former assumes that any agent is animated by well designed and already given determinants (causal constraints, well-behaved preferences, accepted beliefs). These determinants are suitably combined by a uniquely defined, perfect and costless computational process. The latter assumes that an agent has more diversified determinants, acting as mental states, known to him with some uncertainty and progressively constructed during his decision process (contextual preferences, multilevel beliefs). These determinants, where beliefs play a role at least as important as preferences, are variously combined by crude and myopic heuristics according to the complexity of the environment.

The traditional 'equilibration approach' is replaced by an 'adaptation adaptive approach'. The former assumes that the agents are elementary bricks essentially differentiated by their exogenous determinants (diversified preferences, local beliefs). These agents implement actions which are coordinated without friction by some fictitious entity preventing crossed expectations. The latter assumes that the agents are differentiated by their expectation rules and choice rules and that they co-evolve with an endogenous change in their determinants and behavior rules (evolving tastes, exploration-exploitation trade-off). These agents are involved in pre-given or designed endogenous interactions and adapt their mutual actions through various learning processes which can lead to emergent phenomena.

Three organizational levels of a system, defined by corresponding variables, are assumed to be interconnected. The mental level is characterized by mental states, the individual level is symbolized by agent's actions and the collective level is represented by social constructs. The agents reasoning, supported by mental states, defines his intended actions and these actions, once implemented, determine social phenomena. But conversely, social events may have a retroactive influence on individual behaviors, which in turn feed back onto agents' mental states. More recently, a fourth level, the neural level, has been introduced at the bottom of the hierarchy. It is characterized by neural states, which are assumed to determine directly the mental states and only indirectly the individual behavior.

Cognitive economics is supported by a renewed *epistemology* as concerns modeling practice, with two strong trends that mark a departure from classical epistemology. Firstly, more various data are available to the modeler, since both subjective mental states and objective qualitative patterns are considered as directly observable even if hard to measure. Secondly, more various reasoning operations are exerted by the modeler on his models, as concerns their internal structuring as well as their adequacy to the data. Globally, modeling practice stresses that many social phenomena can be explained by relatively simple assumptions, even if the constructed model remains an ideal one and if its proximity to the actual world remains to be precisely assessed.

The traditional 'positivist approach' is replaced by a 'constructivist approach'. The former assumes that the mental states, essentially preferences and beliefs, are unobservable items and have to be revealed (when coherent and stable) from the observable ones, i.e. actions. Besides, the observable facts are summarized in disaggregated and hopefully quantifiable indicators measured by statistical instruments. The latter assumes that the mental states (and reasoning modes), even if somewhat opaque, may be directly apprehended by the agent in an introspective way and further declared to the modeler. In the same spirit, some original patterns can be directly extracted from laboratory and historical experience and appear as ‘stylized facts’ treated by the modeler as sophisticated data.

The traditional ‘deductivist approach’ is replaced by a ‘counterfactual approach’. The former assumes that a model deduces analytically from well defined assumptions incorporating general laws some consequences which are empirically testable. In some delimited application domain, these consequences are confronted in a projective way to the available data and are assumed to be clearly confirmed or refuted. The latter assumes that a model essentially describes a mechanism linking assumptions which are only approximately true or even known to be false to a set of specific consequences. It can be tested in its assumptions as well as its consequences and is progressively revised with regard to original data by a projective as well as inductive process.

Three successive and contrasted steps in the construction of models can be achieved by the modeler, eventually sustained by different evaluation criteria. A possible model is just a coherent suggestion adapted to the field, an acceptable model is an explanation of a recognized phenomenon, and an assessed model is successfully tested against the data. The modeler transforms a possible model into an acceptable one by conceptual exploration in the ‘context of discovery’, then into an assessed model by empirical validation in the ‘context of proof’. However, empirical refutation of a model leads the modeler to revise the original acceptable model and theoretical critics about an acceptable model to suggest new possible models.

Cognitive economics sustains a renewed *praxeology* concerning public policy, which departs in two main directions from classical praxeology. Firstly, more diversified policy measures are considered by the modeler, since persuasion and incentives become natural means of influencing agents’ behavior through agent’s beliefs and reasoning. Secondly, more adapted interventions are studied by the modeler, since only local and brief shocks may be sufficient in order to induce the system’s path to deviate in a desired direction. Globally, public (or private) policy looks more and more similar to a medical act in search of a careful and progressive treatment than to an engineering act in search of a resolute and instantaneous achievement.

The traditional ‘controlling approach’ is replaced by an ‘inciting approach’. The former assumes that the policy measures are essentially physical and monetary ones and act as additional causal factors even if intentionally designed. The main goal is to face the various sources of uncertainty generated by random factors acting besides physical determinism as well as human will. The latter assumes that policy measures may be cognitive and influence the agent's beliefs by providing relevantinformation or even suggesting strong images. Moreover, it accepts that the system's future is not only subject to hardly predictable bifurcations, but is also submitted to radical , expressed by the occurrence of unexpected contingencies.

The traditional 'regulationist approach' is replaced by a 'self-organizing approach'. The former assumes that public policy is capable of implementing actions that steer the economic system nicely along a desired path despite a changing environment. Moreover, the state is considered as a privileged agent acting in the public interest, just considered as an aggregation of private interests. The latter argues that public policy is only able to make the system deviate from one path to another when it acts at the right time and in the right place. Moreover, the state is an actor situated on the same level as the other agents, following its own objectives, since the actors interests appear too soft, look too heterogeneous and evolve too fast.

Three successive stages can be achieved by the modeler when considering the system's future, based on different goals. A plausible scenario is a future path consistent with the representation of the system, a realizable scenario is moreover achievable with the available means, and a desired scenario is, in addition, in accordance with the pursued aims. The modeler transforms a plausible scenario into a realizable one by verifying that well-adapted measures can be activated, a realizable scenario into a desired one by checking that it meets the main objectives. However, when a scenario fails to be accepted as valuable, the modeler is urged to substitute another realizable scenario and critics about the last lead even to suggest new plausible scenarios.

Cognitive economics still appears as a progressive research program, both a theoretical and an empirical one, able to provide more increasing returns. It still may import some devices from cognitive science, mathematics or even sciences of nature, but has to avoid the danger of 'wild economics', where foreign concepts or models are artificially introduced into economics without precisely studying their relevance. A problem must first be stated in economic terms, with a precise question to be answered, before looking to see whether an analogous problem has already been treated in another field. It aims at giving to economics a systematic cognitive turn, by adding a new dimension to its already recognized corpus, in the manner of other social sciences (cognitive sociology, cognitive linguistics).

It needs first to refine its ontology by developing simple and original schemes, which associate reasoning and learning in a manner hopefully not more simplistic than the classical ones. The study of agents' behavior has to incorporate forms of perceptive patterns such as categorization, kinds of reasoning modes such as interpretation and types of decision factors such as emotion. The study of agents' interactions has to incorporate forms of mutual relations such as symbolism, kinds of adjustment processes such as negotiation and types of social influence such as socialization. The study of system's temporality has to consider different levels of learning processes, transitory states as well as asymptotic states, continuous phenomena as well as emergent structures.

It needs further to shift its epistemology towards more empirical work, in association with cognitive psychology, and to treat it in a more inductive way. Special importance must be given to field experiments, which provide the best data about how concrete agents react to various institutional devices in a natural economic environment. A development of laboratory experiments must be achieved by implementing rigorous protocols (control of environment, sampling of agents, replication of experiments), however avoiding too dispersed experiments in random directions. Some empirical work has also to be pursued in neuro-economics in order to examine if it can give more information than just what brain areas are activated by such and such mental activity.

It needs finally to develop its praxeology, by adapting its achievements mainly deployed in game theory to more specific economic problems. On the one hand, it must avoid being diluted into the overly-general theoretical movement of 'social cognition' or the even more general explanatory program called 'complex systems'. On the other hand, it must avoid being reduced to overly-specific economic problems such as the attitude of investors towards radical uncertainty, the innovation process followed by firms or the formation of financial bubbles in a market. If economists really become persuaded that the cognitive dimension is important and fruitful, they will turn this weak heterodoxy into full orthodoxy and progressively abandon its labeling as 'cognitive economics'.

References

Akerlof G (1970): The market for lemons: quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, **84(3)**, 488-500.

Alchian P (2001): Uncertainty, evolution and economic theory, *Journal of Political Economy*, **58**, 599-603.

Alchourron C.E, Grdenfors P, Makinson D (1985): On the logic of theory change: partial meet contraction and revision functions, *Journal of Symbolic Logic*, **50**, 510-530.

Allais M (1953): Le comportement de l'homme rationnel devant le risqué: critique des postulats et axiomes de l'cole américaine, *Econometrica*, **21(4)**, 503-546.

Arena R, Festré A (2006): *Knowledge, beliefs and economics*, Edward Elgar.

Arrow (1974): *The limits of organization*, Norton.

Arthur W.B (1989): (f) Competing technologies, increasing returns and lock-in by historical events, *Economic Journal*, 116-131.

Aumann R.J (1976): Agreeing to disagree, *The Annals of Statistics*, **4(6)**, 1236-39.

Aumann R.J (1995): Backwards induction and common knowledge of rationality, *Games and Economic Behavior*, **17**, 138-146.

Aumann R.J (1999): Interactive epistemology, *International Journal of Game Theory*, **28(3)**, 263-314.

Aumann R.J, Brandenburger A (1995): Epistemic conditions for Nash equilibrium, *Econometrica*, **63(5)**, 1161-80.

Aumann R.J, Hart S (1992, 1994, 2002): *Handbook of game theory with economic applications, 3 volumes*, Elsevier.

Baltag A, Moss L.S (2004): Logics for epistemic programs, *Synthese*, **139(2)**, 165-224.

Banerjee M (1992): A simple model of herd behaviour, *Quarterly Journal of Economics*, **108**, 797-817.

Barwise (1988): *Three views of common knowledge, in M. Vardi ed., Proceedings of the TARK Conference*, Morgan Kaufmann.

van Benthem J, ter Meulen A eds (1997): *Handbook of Logic and Language*, Elsevier/MIT Press.

Benz A, Jaeger G, van Rooij R eds (2005): *Game theory and pragmatics*, Palgrave Macmillan.

Bernoulli D (1738): Specimen theoriae novae de mensura sortis, Saint-Petersbourg, trad. 'Exposition of a new theory on the measurement of risk', *Econometrica*, **22**, 23-36.

Billot A, Vergnaud J.C, Walliser B (2006): *Multiplayer belief revision and accuracy orders*, Proceedings of the LOFT Conference.

Binmore K (1987): Modeling rational players, *Economics and Philosophy*, **3**, 9-55, **4**, 179-214.

Binmore K (1997): Rationality and backward induction, *Journal of Economic Methodology*, **4**, 23-41.

Binmore K, Brandenburger A (1990): *Common knowledge and game theory*, In: Essays in the Foundations of Game Theory, Blackwell, 105-150.

Blackwell D (1951): *Comparison of experiments. Proceedings of the second Berkeley symposium on mathematical statistics and probability*, University of California Press, 93-102.

Boudon R (1981): *The logic of social action*, Routledge.

Bouleau N (2003): *Martingales and financial markets*, Springer.

Bourgine P, Nadal J.P eds (2004): *Cognitive economics: an interdisciplinary approach*, Springer.

Boyer R, Orléan A (1992): How do conventions evolve? *Journal of Evolutionary Economics*, **2**, 165-177.

Bush R, Mosteller F (1955): *Stochastic models for learning*, Wiley.

Camerer C, Ho T.H, Chong J.K (2004): A cognitive hierarchy theory of one-shot games and experimental analysis, *Quarterly Journal of Economics*, **119(3)**, 861-98.

Conlisk, J (1996): Why bounded rationality?, *Journal of Economic Literature*, **34**, 669-700.

Debreu G (1954): Representation of a preference ordering by a numerical function, in R.M. et alii eds, *Decision processes*, 159-165, Wiley.

Dosi G et al. eds (1988): *Technical change and economic theory*, Pinter Publishers.

Dutta P.K, Radner R (1999): Profit maximization and the market selection hypothesis, *Review of Economic Studies*, **66**.

Egidi M (2007): *Decomposition patterns in problem solving, in R. Topol, B. Walliser: Cognitive economics: new trends*, Elsevier.

Egidi M, Rizzello S eds (2004): *Cognitive economics: foundations and historic evolution*, Edward Elgar.

Ellsberg E (1961): Risk, ambiguity and the Savage axioms, *Quarterly Journal of Economics*, **74 (4)**, 643-669.

Elster J (1989): *The cement of society*, Cambridge University Press.

Fagin R, Halpern J, Moses Y, Vardi M (1995): *Reasoning about knowledge*, MIT Press.

Foray D (2004): *The economics of knowledge*, MIT Press.

Friedman M (1953): *Essays in positive economics*, University of Chicago Press.

Fudenberg D, Levine D (1998): *The theory of learning in games*, MIT Press.

Gigerenzer G, Selten R, eds (2001): *Bounded rationality*, MIT Press.

Gilboa I, Schmeidler D (1989): Maxmin expected utility with non unique prior, *Journal of Mathematical Economics*, **18**, 141-153.

Gilboa I, Schmeidler D (2001): *A theory of case based decisions*, Cambridge University Press.

Gittins J (1989): *Multi-armed bandits allocation indices*, Wiley.

Gossner O, Mertens J.F (2001): *The value of information in zero-sum games*, Mimeo.

Grossman S, Hart O (1983): An analysis of the principal-agent problem, *Econometrica*, **51 (1)**, 42-64.

Grossman S, Stiglitz J (1980): On the impossibility of informationnally efficient market, *American Economic Review*, **70(3)**.

Harsanyi J.C (1967): Game with incomplete information played by Bayesian players, *Management Science*, 159-82, 320-34, 486-502.

Harsanyi J.C (1973): Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points, *International Journal of Game Theory*, **2**, 1-23.

Harsanyi J.C, Selten R (1988): *A general theory of equilibrium selection in games*, MIT Press.

von Hayek F (1973): *Law, legislation and liberty*, University of Chicago Press.

von Hayek F (1937): Economics and knowledge, *Economica*, **4**.

Heiner R.A (1983): The origin of predictable behavior, *American Economic Review*, **73**, 560-595.

Holland J.H (1975): *Adaptation in natural and artificial systems*, Michigan University Press.

Jéhiel P (2005): Analogy-based expectation equilibrium, *Journal of Economic Theory,* **123**, 81-104.

Johnson-Laird P.N (1983): *Mental models*, Cambridge University Press.

Kahneman D, Slovic P, Tversky A eds (1983): *Judgment under uncertainty: heuristics and biases*, Cambridge University Press.

Kahneman D, Tversky A (1979): Prospect theory: an analysis of decision under risk, *Econometrica*, **47**, 263-291.

Kamien M, Taumann Y, Zamir S (1990): On the value of information in a strategic conflict, *Games and Economic Behavior*, **2**, 129-53.

Katzuno A, Mendelzon A (1992): Propositional knowledge base revision and nonmonotonicity, *in P. Grdenfors ed.: Belief revision*, Cambridge University Press.

Keynes J.M (1936): *The general theory of employment, interest and money*, Hartcourt-Brace.

Kirman A (1993): Ants, rationality and recruitment, *Quarterly Journal of Economics*, **108**, 137-156.

Knight F (1921): *Risk, uncertainty and profit*, Kelley.

Kokinov B (2005): *Advances in cognitive economics*, NBU Press.

Kraus D, Lehmann D, Magidor M (1990): Non monotonic reasoning, preferential models and cumulative logics, *Artificial Intelligence*, **44**, 167-208.

Kreps D, Milgrom P, Roberts J, Wilson R (1983): Rational cooperation in the finitely repeated prisoner's dilemma, *Journal of Economic Theory*, **27**.

Laing R.D (1970): *Knots*, Random House.

Lesourne J (1992): *The economics of order and disorder*, Clarendon Press.

Lesourne J, Orléan A, Walliser B (2006): *Evolutionary microeconomics*, Springer.

Lewis D (1969): *Conventions: a philosophical study*, Harvard University Press.

Lewis D (1973): *Counterfactuals*, Blackwell.

LOFT (1994, 96, 98, 2000, 02, 04, 06): *Proceedings of the Conferences 'Logic and the foundations of game and decision theory'.*

Luce R.A (1959): *Individual choice behaviour: a theoretical analysis*, John Wiley.

Macho-Stadler I, Perez-Castillo D, Watt R (2001): *An introduction to the economics of information: incentives and contracts*, Oxford University Press.

Marengo L (1992): *Structure, competence and learning in an adaptive model of the firm, in G. Dosi- F. Malerba eds., Organization and strategy in the evolution of the enterprise*, 124-154, Mac Millan.

McCain R.A (1992): *A framework for cognitive economics*, Praeger Publishers.

Merton M (1936): The unanticipated consequences of purposive social action, *American Sociological Review*, **65**.

Milgrom P, Stokey N (1982): Information, trade and common knowledge, *Journal of Economic Theory.*

Mongin P, Walliser B (1988): *Infinite regressions in the optimizing theory of decision, in B. Munier ed., Risk, decision and rationality*, 435-457, Reidel.

Nash J (1950): Equilibrium points in N-person games, *Proceedings of the National Academy of Sciences (USA).*

Nelson R.R, Winter, S. (1992): *An evolutionary theory of economic change*, Harvard University Press.

von Neumann J, Morgenstern O (1944): *Theory of games and economic behaviour*, Princeton University Press.

Neyman A (1991): The positive value of information, *Games and Economic Behavior*, **3**, 350-55.

North D (1990): *Institutions, institutional change and economic performance*, MIT Press.

Orléan A (1998): *The evolution of imitation, in P. Cohendet, P. Llerena, H. Stahn, G. Umbhauer eds., The economics of networks*, Springer.

Osborne M.J, Rubinstein A (1994): *A course in game theory*, MIT Press.

Peirce C.S (1878): *Deduction, induction and hypothesis, in Collected papers of Charles Sanders Peirce*, Harvard University Press.

Proust, M (1954): *Du ct de chez Swann*, Gallimard.

Quiggin J (1982): A theory of anticipated utility, *Journal of Economic Behavior and Organization*, **3**, 321-343.

Rabin M (1993): Incorporating fairness into game theory and economics, *American Review*, **83(5)**, 1281-302.

Radner R (1980): Collusive behavior in non-cooperative epsilon-equilibria of oligopolies with long but finite lives, *Journal of Economic Theory*, **22**, 136-154.

Rasmusen E (1989): *Games and Information*, Blackwell.

Reny P.J (1993): Common belief and the theory of games with imperfect information, *Journal of Economic Theory*, **59**, 257-274.

Rizzello S (2003): *Cognitive developments in economics*, Routledge.

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

Roth A, Erev I (1995): Learning in extensive form game, *Games and Economic Behavior*, **8**, 164-212.

Rubinstein A (1989): The electronic mail game: strategic behaviour under almost common knowledge, *American Economic Review*, **79(3)**, 385-391

Rubinstein A (1994): *Models of bounded rationality*, MIT Press.

Rubinstein A (2000): *Economics and language, five essays*, Cambridge University Press

Sarin R.K, Wakker P (1998) Dynamic choice and non expected utility, *Journal of Risk and Uncertainty*, **17**, 87-119.

Savage L.J (1954): *The foundations of statistics.* John Wiley

Schackle G (1952): *Expectation in Economics*, Cambridge University Press.

Schelling T (1960): *The strategy of conflict*, Harvard University Press.

Schelling T (1978): *Micromotives and macrobehavior*, Norton.

Schmeidler D (1975): Reexamination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, **4(1)**, 25-55.

Sethi R (1999): Evolutionary stability and media of exchange, *Journal of Economic Behaviour and Organization.*

Shafer G (1976): *A mathematical theory of evidence*, Princeton University Press.

Simon H (1955): A behavioural model of rational choice, *Quarterly Journal of Economics*, **69**, 129-138.

Simon H (1957): *Models of man, social and rational*, John Wiley.

Simon H (1959): Theories of decision-making in economics and behavioural science, *American Economic Review.*

Simon H.A (1976): *From substantive to procedural rationality, in J. Latsis ed. Method and Appraisal in Economics*, Cambridge University Press.

Simon H (1982): *Models of bounded rationality*, MIT Press.

Spence A.M (1973): Job market signalling, *Quarterly Journal of Economics*, **87(3)**, 355-74.

Sperber D (1996): *Explaining culture: a naturalist approach*, Blackwell.

Sperber D, Wilson D (1995): *Relevance: communication and cognition*, Blackwell.

Stalnaker R.C (1968): *A theory of conditionals, in N. Rescher ed. Studies in Logical Theory*, Blackwell.

Stalnaker R.C (1996): Knowledge, belief and counterfactual reasoning in games, *Economics and Philosophy*, **12**, 133-163.

Sugden R (1986): *The evolution of rights, cooperation and welfare*, Blackwell.

Sutton R.S, Barto A.G (1998): *Reinforcement learning*, MIT Press-Bradford Books.

TARK (1986, 88, 90, 92, 94, 96, 98, 2001, 03, 05): *Proceedings of the Conferences 'Theoretical aspects of reasoning about knowledge'*, Morgan Kaufmann.

Topol R, Walliser B (2007): *Cognitive economics: new trends*, Elsevier.

Tversky A (1972): Elimination by aspects: a theory of choice, *Psychological Review*, **79**, 281-299.

Viale R (1997): *Cognitive economics*, La Rosa Editrice.

Walliser B (1989): Instrumental rationality and cognitive rationality, *Theory and Decision*, **27**, 7-36.

Walliser B (2006): Justifications of game equilibrium notions, *in R. Arena and A. Festre eds. Knowledge, beliefs and economics*, Edward Elgar.

Walliser B, Zwirn D (2002): Can Bayes rule be justified by cognitive rationality principles?, *Theory and Decision*, **53**, 95-135.

Walliser B, Zwirn D, Zwirn H (2005): Abductive logics in a belief revision framework, *Journal of Logic, Language and Information*, **14**, 87-117.

Watzlawick P (1970): *La réalité de la réalité*, Seuil.

Weibull J (1995): *Evolutionary game theory*, MIT Press.

Weisbuch G, Kirman A, Herreiner D (2000): Market organization and trading relationship, *Economic Journal*, **110**, 411-36.

Young P (1998): *Individual strategy and social structure*, Princeton University Press.

Index