

# GAME THEORY AND EMERGENCE OF INSTITUTIONS

*Bernard WALLISER (ENPC, EHESS)*

January 2006

‘The institutions are the rules of the game of a society’

D. North (1990)

‘A institution is a self-sustaining system of shared beliefs about how the game is played’

M. Aoki (2001)

## **Summary :**

The paper recalls the usual features of institutions as concerns their concrete support, their economic and social functions and their influence on basic agents. In relation with methodological individualism, it considers the benefits of studying the genesis of an institution in a game theoretic framework. First, an institution may be obtained as an equilibrium by a voluntary procedure, which introduces an auxiliary game with regard to the main game. Second, it may be interpreted as the equilibrium resulting from an eductive process, where the players’ reasoning stays implicit or becomes explicit. Third, it may be analyzed as the equilibrium resulting from an evolutionary process, either belief based learning or reinforcement learning. Some problems stay unsolved, for instance the collective naturalization and the normatization of the institution.

**Key-words :** eductive view, equilibrium, evolutionary view, game theory, institution, norm

## 0. Introduction

Economic theory describes any economic system as resulting from various forms of interaction between three autonomous entities: agents, nature and institutions. In first instance, agents and nature are considered as primitive entities, even if they are obviously evolving under outside influences. On the contrary, institutions are analyzed as artificial devices whose emergence modalities have to be analyzed, at least when adopting a long term perspective. According to methodological individualism which is dominant in economic thought, they have to be explained as the result of some combination of agents' actions, even if they influence the agents by a feedback effect. Hence, an institution appears as a clear illustration of an 'emerging phenomenon' in social sciences, i.e. a macroscopic salient phenomenon related in both ways to microscopic entities.

Such a research program is especially taken up in the encompassing framework of (non cooperative) game theory. Game theory looks generic enough to deal with various modes of action, does not explicitly presuppose any prior institution and agrees apparently with methodological individualism. The main idea consists in interpreting an institution as some equilibrium state consciously or unconsciously reached by several players. Such a view was already present in the founding book of game theory (von Neumann and Morgenstern, 1944), where the 'solution' of a game is interpreted as a 'standard of behavior'. It was further developed by a lot of local models, exploiting all the principles of game theory as soon as they were elaborated. These models are more or less restricted to specific forms of institutions, i.e. regulating entities, behavior norms or social conventions.

A correlative idea consists in considering the genesis of an institution as the genesis of some equilibrium state. According to game theory, as for any emergent phenomenon, three main approaches of the genesis of an institution can be considered. The first approach considers an institution as the result of a collective choice process in which it is voluntarily designed in an auxiliary game with regard to the main game. The second approach considers an institution as a behavioral regularity which is directly obtained by the agents in some game and reinterpreted as an institution. It divides into two sub-approaches for which the institution is obtained either by some eductive or by some evolutionary process (Binmore, 1992). In any case, it must be stated how an institution once appeared becomes recognized and further legitimized by the players.

The first part of the paper recalls the usual properties of institutions as concerns their concrete support (§1.1), their means of influence (§1.2) and their socio-economic functions (§1.3). The second part deals with the ontology of economic systems (§2.1), the role of methodological individualism (§2.2) and the benefits of an appeal to game theory (§2.3). The third part gives the principles of an a voluntary design of an institution (§3.1), its expression as an outside procedure (§ 3.2) or as an inside procedure (§3.3). The fourth part details the eductive emergence of an institution (§4.1), by an implicit reasoning of the players (§4.2) or by a more explicit coordination through beliefs (§4.3). The fifth part details the evolutionary emergence of an institution (§5.1), taking the form of a belief-based learning process (§5.2) or a reinforcement learning process (§5.3). The last part deals with the naturalization of an already emerged institution (§6.1), the normatization of a naturalized institution (§6.2) and with other open problems related to organizational levels and time scales (§6.3).

## **1. Properties of institutions**

### ***1.1. Nature of institutions***

An institution has a concrete support which oscillates between a cognitive and a social one, giving rise to two extreme types (Greif, 1994). A ‘conceptual institution’ is a mental representation which is endorsed by a collection of individuals who agree on it. For instance, some beliefs (trust in money) are commonly shared, some habits (clothing customs) are extensively followed and some conventions (technological standards) are collectively internalized. An ‘organisational institution’ is a material grouping of several agents intended to satisfy some collective aims. For instance, people gather into families, syndicates as well as firms. Of course, a conceptual institution may be defended by an organisational one and conversely, an organisational institution may incarnate a conceptual one. For instance, money is supported by central banks, faith in god by churches, property rights by courts; conversely, a firm is grounded on a ‘hierarchy’, a family on ‘marriage’ and the State on a ‘social contract’.

An institution is more or less explicitly stated and perceived by the concerned agents, leading again to two contrasted types (North, 1990). A ‘formal institution’ is precisely introduced in the concerned group by some external or internal authority. For instance, some rules (traffic rules) are explicitly imposed by the State while some associations (churches) are designed by themselves. An ‘informal institution’ is only loosely defined by a tacit acceptance of interested members. For instance, some conventions (politeness rules) hold by mutual acceptance while some organizations (family friends) have fuzzy frontiers. Of course, a formal institution needs to receive a more informal support and an informal institution may become formally reinforced. For instance, deontological rules imposed by a professional association are generally accepted by its members; conversely, spontaneous language conventions may further be imposed by a public authority.

Institutions are generally not independent entities, but form networks of institutions which sustain one another in a complementary or substitutive way. For instance, a culture is often defined as a bundle of beliefs and norms and the State gathers many organizations and rights. Especially, institutions form hierarchical structures where upper level ones justify lower level ones. For instance, if formal institutions are naturally structured into hierarchies (constitution, laws, decrees), the same is true for informal ones (generic savoir-faire norms, specific behavior rules). In a more taxonomical way, an institution is conceived as a general concept which may give rise to specific occurrences in restricted contexts. For instance, money appears under different modes in different countries, marriage satisfies different rules within different religions, language spreads under adapted forms in different groups.

In economics, an institution is still considered as a hotchpotch concept since it covers all entities which are not reducible to agents or nature. The central institution is the ‘market’ which confronts the offers and demands of some good and renders them compatible through an associated price. It is organisational since it appears as the gathering of merchants in a same place, but is conceptual too since it can be considered as a node of contracts. Some institutions act as substitutes of the market, always with the aim of coordinating the exchanges: auction mechanisms, planning procedures, bilateral negotiations. Other institutions just contribute to facilitate the exchanges: exchange rights fix legal conditions for exchange, money fluidifies the practice of exchange, trust renders possible the two opposite movements of exchange. Further institutions intend to frame the global system of

transactions : technical conventions codify the goods and technologies, moral norms determine the exchangeable goods, property rights fix the owners of the goods.

## ***1.2. Influence of institutions***

An institution shapes the social system in a more or less original and profound way, leading to adapted categories (Searle, 1995). A ‘constitutive institution’ creates ex nihilo new activities which may be of some interest for the agents. For instance, a new language can be created by arbitrary rules or a new parlour game can be conceived with some conventional rules. A ‘regulative institution’ controls some already existing activities which present some failures. For instance, spontaneous road traffic is controlled by some legal rules while queues in an entertainment place are regulated by some conventional rules. Of course, there is no sharp distinction between the two categories of institutions and a same institution may act in both ways. For instance, money helps to realize the ordinary transactions and creates an original opportunity of reserve of value.

More precisely, an institution is considered as acting on either of the three choice determinants of an agent (opportunities, beliefs, preferences), giving rise to two main types. A ‘deontic institution’ acts on an agent’s preferences by incentives or sanctions. For instance, a solidarity norm such as care to person in danger modifies his preferences. A ‘doxastic institution’ acts on an agent’s beliefs by learning or persuasion. For instance, a theory such as ‘mankind results from biological evolution’ influences their beliefs. Besides, an institution may act on opportunities too by constraining or enlarging an agent’s action set. For instance, a law prevents the agents from buying votes. However, since constraints are considered by the agents in their preferences (if violating them is costly) or in their beliefs (if they are non transgressable), an institution acts essentially on these last two determinants.

In a more direct way, an institution can be considered as acting on an agent’s behavior rule or even on a definite action. In the first case, since a behavior rule relates an action to the environment (‘if context C, then do action A’), the institution amounts to a ‘social norm’ which obeys the same scheme. For instance, a norm of savoir-faire makes precise how precisely to behave in a given social setting. In the second case, by forgetting the context, an institution is even considered as an allowed action. For instance, marriage just fixes a legal opportunity and duty for agents sticking to it. However, such a shortcut institution may nevertheless be interpreted as acting more or less directly on choice determinants. The problem is that any given behavior can always be justified by various combinations of beliefs and preferences. For instance (Orléan, 2002), an imitation behavior may be interpreted as stemming from preferences (interest to act in the same way than the other) or beliefs (interest to follow the other assumed to have more information).

In economics, a deontic institution suggests evaluation criteria for assessing goods (exchangeable goods excluding for instance human organs), activities (work conditions excluding for instance child labor) or global states (wealth distribution excluding for instance very low wages). It is now common to distinguish for each agent between ‘personal preferences’ which are purely self-interested, and ‘moral preferences’ which reflect a collective interest, even if it is not clear in what circumstances the ones or the others are used. Likely, a doxastic institution suggests taxonomic principles for classifying goods (nomenclature of goods), agents (nomenclature of professions) or global states (taxonomy of economic regimes). In order to help them to decipher their environment, it is now usual to

assume that some ‘cultural background’ is common to all agents involved in an economic system, even if its content may change from one system to another.

### ***1.3. Functions of institutions***

An institution helps the concerned agents to face material as well as human forms of uncertainty, leading to specific categories. A ‘protective institution’ prevents the agents against the random effects of natural phenomena. For instance, an insurance system alleviates the effects of a disaster or an illness by splitting the risks among all agents. A ‘coordinative institution’ prevents the agents against the unexpected effects of others’ behaviors. For instance, a reservation system channels the demands for some restricted activity. Of course, outside the essential function just described, other functions can be assigned to institutions, especially symbolic ones. For instance, if marriage has the (internal) protective function of insuring the couple and the (internal) coordinative function of managing common resources, it has moreover the symbolic (external) function of exhibiting a social link.

An institution helps to overcome uncertainty either about the past or about the future, giving again rise to two contrasted classes. A ‘knowledge oriented institution’ helps an agent to assess the structural regularities of his material or social environment and the effects of his actions. For instance, ‘statistical institutes’, ‘scientific committees’ or ‘formation institutes’ compensate for agent’s limited information and knowledge. An ‘action oriented institution’ helps an agent to forecast the future opportunities and threats and to adjust his behavior to his environment. For instance, ‘common references’, ‘shared expectations’ or ‘social roles’ help an agent to expect the others’ behaviors. Moreover, some institutions act as meta-institutions which are necessary for the others to be functional. For instance, a common language is necessary for a common understanding of the surrounding world and for a common action on that world.

More precisely, an institution compensates for ‘system failures’ which appear at the collective level with regard to some reference state (no institution, ideal institution). For instance, a misconception happens when agents are not able to agree on the interpretation of a common phenomenon while a misjudgment happens when agents are not able to agree on the assessment of a common action. Formally, many failures are expressed under the form of ‘social costs’ acting as undesirable frictions. For instance, information and computing costs appear when agents are looking for a better knowledge, while negotiation or transaction costs appear when agents are looking for a better coordination. Of course, an institution which solves a first order failure may often give rise to a second order failure. For instance, an insurance system involves high organizational costs or induces inequalities between insured people.

In economics, three ‘market failures’ are usually considered, by contrast with an ideal perfect competitive market (which is Pareto-optimal). First, ‘technological failures’ arise from externalities or non convexities which prevent an equilibrium state to happen. For instance, a production technology may show increasing returns while preferences of two consumers may be related by positive or negative externalities. Second, ‘informational failures’ arise from the lack of relevant information and make it more difficult for the agents to reach an equilibrium state. For instance, agents have imperfect or incomplete information on states of nature (adverse selection) or on other’s actions (moral hazard). Third, ‘organizational failures’ arise from biased interactions between agents which distort the competitive equilibrium conditions.

For instance, imperfect competition results from the existence of a few producers acting as an oligopoly, while sticky prices prevent that offers and demands adjust rapidly enough.

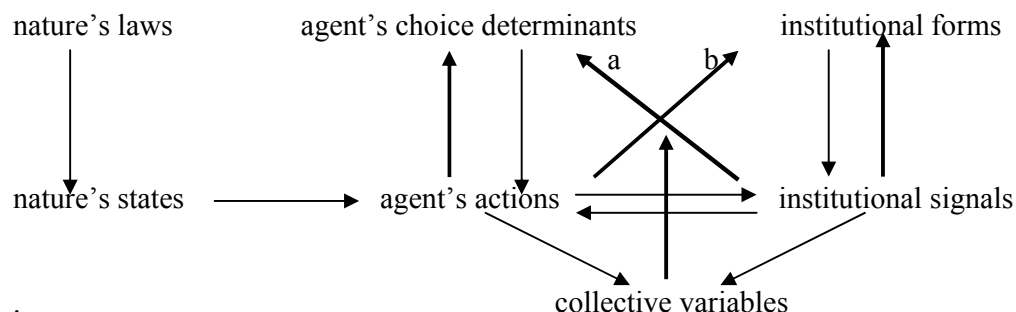
## 2. Institutions and games

### 2.1. *Ontology of a social system*

The three basic entities which constitute a social system can be described by both structural and factual variables. Agents behave in a rational way, are structurally characterized by their three ‘choice determinants’ and factually by the implementation of ‘actions’. If ‘cognitive rationality’ concerns the adequation of beliefs to available information, ‘instrumental rationality’ refers to the adequation of actions to preferences (for given beliefs). Nature behaves in a mechanistic way, is structurally defined by (probabilistic) ‘causal laws’ and factually by the fixation of ‘states’ of nature. Institutions behave in a mechanistic way too, are structurally characterized by ‘institutional forms’ and factually by the fixation of institutional ‘signals’. Furthermore, collective (factual) variables have to be considered, obtained either by composition or by aggregation of individual ones.

The relations inside or among the basic entities can moreover be differentiated into short term and long term ones. In the short term, for each entity, the structural variables influence the factual ones in a top down way. Moreover, at the factual level, there is a two-way influence between agents and institutions and a one way relation between nature and agents; likely, the actions and signals determine together the collective variables. In the long term, for each entity (except eventually for nature), the factual variables influence the structural variables in a bottom up way. Moreover, two interlevel relations may be considered between the two main entities. Relation (a) concerns the possibility of a ‘social conditioning’ on agents, stemming from the signals and/or collective variables and acting on their determinants (essentially their preferences). Relation (b) concerns the main influence considered in the paper, i.e. the genesis of the institutional forms by the agents’ actions and/or the collective variables.

The following scheme makes explicit the relevant relations between variables (the structural variables are on the upper line and the long term relations are in bold):



In economics, on a competitive market, nature furnishes resources and technologies and fixes environmental conditions, all items summarized in well defined states of nature. Agents are split into two types, producers who offer (and demand) goods by maximizing their profit under a technological constraint, and consumers who demand goods by maximizing their utility under a budget constraint. The market is symbolized by the ‘Walrasian auctioneer’, a fictitious entity who computes the prices, treated as institutional signals, by equalizing the

demands and offers. Hence, offers-demands as well as prices result from a fixed point (influenced by nature) between the agents and the Walrasian auctioneer at the factual level. However, an agent's determinants may be influenced by prices and other global phenomena, as is the case for fashion or addiction. Conversely, the market features are determined by the agents' actions as is the case for search of cartels or fight for novelty.

## 2.2. Institutions and methodological individualism

According to the fact that relations (a) and (b) are introduced or not, four epistemological positions can be defined :

	b	not b
a	methodological interactionism	methodological holism
not a	methodological individualism	methodological isolationism

In methodological individualism, institutions are designed by agents without feedback from the first towards the last. In methodological holism, agents are conditioned by institutions (and social environment) without feedback from the first towards the last. In methodological isolationism, institutions are autonomous and agents are self-contained. In methodological interactionism, agents and institutions influence each other in both ways.

The emergence of institutions due to individual actions precisely assumes relation (b), hence relies on methodological individualism (or methodological interactionism). But such an emergence still obeys to different schemes progressively detailed by different schools. On one hand, an institution is intentionally designed in order to solve some collective failures and introduces a 'voluntary order'. For instance, an original language is constructed (esperanto), an artificial market is decided (financial market), a new traffic rule is imposed (priority rule). On the other hand, an institution results from an unintentional process naturally followed and assumes a 'spontaneous order'. For instance, a language, a market or a traffic rule appears progressively as a 'behavior regularity mutually accepted' (Schotter, 1981). In fact, a spontaneous order results itself from two types of processes, according to the agents' more or less demanding reasoning capacities.

In an 'eductive view', hyper intelligent agents come to an institution by their sole crossed reasoning. For instance, Lewis (1969) considers that a 'convention' results from such a process as can be illustrated by the standard examples of money or language. More precisely, in some socialgroup G, in a recurrent situation S, a 'convention' is a regularity **R** when the following conditions (ordered differently from Lewis) happen to be common knowledge (each agent knows, knows that the other knows, ...). The first condition states that each player prefers a conformity to **R** to a less general conformity. The second condition states that each player has a decisive reason to conform to **R** if the others do. The third condition states that all players actually conform to **R**. The fourth condition states that **R** is not the only regularity which satisfies the preceding conditions.

In an 'evolutionary view', boundedly rational agents come to an equilibrium by some collective learning or evolution process. For instance, a functionalist approach (Merton, 1957; Stichcombe, 1968; Elster, 1989) considers that an 'institution' **I** is explained by its function F in a group G under the following conditions. The first condition states that the function F is just an effect of I. The second condition expresses that this effect F (as well as the link from I to F) is not intentional for G. The third condition assumes that the effect F is useful for G. The

fourth condition indicates that the effect F maintains the institution I by some feedback mechanism.

### ***2.3. Institutions and game theory***

In order to study institutions, it is convenient to refer to ‘non cooperative’ game theory for three main reasons. First, game theory works at a level of generality which is higher than economic theory, hence allows to explain economic and non economic institutions as the result of diversified actions generated by (strong or weak) rational behaviors. Second, game theory considers bilateral or multilateral relations between players which are not mediated by previous institutions, thus allows (apparently) to ground an institution on an institutional vacuum. Third, game theory is just a specification of the preceding ontology, and respects methodological individualism insofar as it considers as its basic concept an equilibrium state which is (apparently) only generated by players’ actions. But the last two points are somewhat controversial.

As concerns the second point, if no institution is explicitly stated in the definition of a game, some prior institutions - the so called ‘rules of the game’ - are nevertheless introduced, in a more or less implicit way. These prior institutions are conceptual ones, either formal or informal, and may ensure a constitutive or regulative function. They are assumed both to be stable since they do not evolve during the successive plays of a same game and exogenous since the players are not able to feed back on them. Moreover, they are generally fully endorsed by the players, who simultaneously recognize them, respect them and ground their play on them. Hence, the prior institutions are usually directly defined by their effects on players’ choice determinants, especially on their opportunities (tolerated actions) and preferences (incentives and sanctions), which are already considered as stable and exogenous.

As concerns the third point, classical game theory is grounded on some equilibrium notion which defines an expected state of the game. In a game, any state is defined as the set of strategies and eventually of beliefs held by the players. An equilibrium state is defined as a steady state, i.e. a stable state in the absence of external perturbations. More precisely, it looks like a state from which no player has a perceived interest of deviating unilaterally. Hence, it appears as a fixed point of the loop which relates the players’ actions through their reaction functions. It is generally considered as computed and achieved by a fictitious entity, the ‘Nash regulator’, rather than by the players themselves. But the concrete process leading to an equilibrium state is not explicitly described, expressing an ‘implementation failure’ and preventing to satisfy methodological individualism.

Three main ‘game failures’ are moreover considered, by contrast to a game which admits a unique and Pareto-optimal equilibrium state. First, ‘the codetermination problem’ stems from the absence of any equilibrium state. For instance, ‘matching pennies’ admits no equilibrium, at least in pure strategies. Second, ‘the coselection problem’ stems from the multiplicity of equilibrium states, with different cases. For instance, the ‘battle of sexes’ exhibits two asymmetric equilibria; the ‘meeting point game’ admits two equivalent equilibria; the ‘stag hunt game’ admits one equilibrium which Pareto-dominates another one. Third, the ‘cooperation problem’ happens when the equilibrium state is no more Pareto-optimal. For instance, the ‘prisoner’s dilemma’ admits a unique equilibrium which is dominated by another issue. Since an equilibrium state exists in the majority of games, the last two failures are the most relevant and need themselves to be compensated by some institutions..



### **3. Institution as voluntary design**

#### **3.1. Principles**

The fundamental idea followed by game theory consists in identifying an institution with an equilibrium state. An alternative view worth to be considered assimilates an institution to a collectively optimal state, especially a Pareto-optimal state. But game theory does not consider a collectively optimal state as the natural output of a game since it has no reason to be stable and self-enforcing. However, confronted to the cooperation problem, which disconnects an equilibrium and an optimal state, game theory examines how to reach an optimal state by an equilibrium state of a modified game. An extended game frequently multiplies the number of equilibrium states, hence may transform an optimal state into an equilibrium one. For instance, this happens by considering a repeated game or by introducing some uncertainty in the game.

From now, the ‘equilibrium-like’ definition of an institution will be favoured since a general function of coordination is naturally attributed to both entities. As already stressed by Aoki (2001), an institution has some specific features of an equilibrium state and even of a stable equilibrium state. It is endogenous if one accepts that it only results from the combination of agents’ actions (or strategies). It is self-enforcing since each agent has an interest to follow its injunctions if the others do the same. It is relatively robust to exogenous changes in the agents’ characteristics as well as to their natural environment. It admits many substitutes in the attainment of a given social function or in the solution of a given failure. Moreover, it may not be Pareto-optimal and other institutions may be more performant for all agents.

The basic approach which may be adopted is to consider the institution as the result of a voluntary hence conscious choice process. More precisely, two virtual collective choice processes are simultaneously or sequentially involved. In some basic situation, expressed or not in a game framework, some type of failure arises from interactions between agents of a ‘target group’. In an auxiliary game, a set of players, the ‘deciding group’, is assumed to choose an institution able to respond to that failure. The target group of the primitive situation and the deciding group of the auxiliary game may not coincide. Moreover, if the basic situation may be of any type, the auxiliary game is of a specific type since it represents a ‘social choice’ intended to select one and only one institution. In that very case, the assimilation of an institution to an equilibrium state looks one-to-one.

However, two configurations have to be distinguished depending on the precise link assumed between the basic situation and the auxiliary game. In an ‘outside procedure’, the auxiliary game is completely independent of the basic situation. The decision process is decomposed into two steps, a first step where an institution is selected and a second step where the basic situation is played with the chosen institution. The target group and the deciding group are independent. The institution can be of any kind even it has further to be adapted to the basic situation. In an ‘inside procedure’, the auxiliary game coincides with the basic situation. This involves that the basic situation is already expressed as a game and that the two steps by which an institution is designed collapse. The players of the deciding group have to be included in the concerned agents of the basic situation. The institution has moreover to be adapted beforehand to the basic game.

### 3.2. Outside procedure

An institution results first from a voluntary action implemented by one or several players and realizing what can be called an 'institutional equilibrium'. On one hand, the institution is created by a planner who has the prerogative of such a decision and results then from an individual optimal choice relative to the institution. On the other hand, the institution stems from a contract between several players and appears as an equilibrium in the negotiation process between these actors about the institution. The members of the deciding group are allowed with action sets already formed of possible institutions and have preferences defined either directly on the institutions or on their expected consequences. The institution which is adopted is self-enforcing, at least when the deciding group and the target group coincide.

For the allocation of limited resources or goods, especially of public goods, various choice mechanisms can be selected by the State. For instance, concerning the allocation of broadcast rights in the communication sector (Milgrom, 2000), the public authority may select some specific auction mechanism. Different types of auctions (English, Dutch) are then compared along several criteria: high reservation value for the buyer, high profit for the seller, low possibility of manipulation. More generally, the 'mechanism design' literature (Hurwicz, 1972; Maskin, 1985) considers precisely a single decision-maker involved in the construction of an institutional device. The possible institutions are only specified by some features and they have to satisfy some collective efficiency criteria in a given environment. Such a framework can easily be extended to several decision-makers.

In traffic regulation, the 'car lights game' (analogous to the 'prisoner's dilemma') considers two car drivers (players 1 and 2) who may use two types of lights with following payoffs :

1 / 2	dipped	full
dipped	(2,2)	(0,3)
full	(3,0)	(1,1)

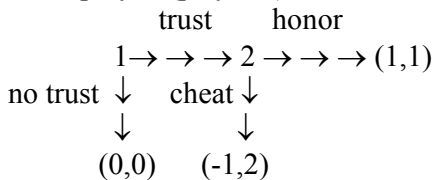
Since the equilibrium state (1,1) is Pareto-dominated by the non equilibrium state (2,2), the players face a cooperation problem. In order to solve this game failure, a planned institution aims at influencing in a favorable way the choice determinants of the agents. It acts on the opportunities when forbidding some actions, namely when preventing the use of full lights. It conditions the preferences when designing incentives and sanctions, namely when fining the use of full lights. It modifies the beliefs when enlightening the structure of the game, namely when stressing the bad consequences of full lights.

As concerns first the implementation problem, it is apparently solved since the chosen institution results from an explicit procedure. However, it is problematic if one considers that the deciding group is itself based on a second order institution. It is generally nominated by some public authority when it does not coincide with the public authority itself. Even when the institution results from a voting procedure, such a procedure can only be decided by a higher level institution. Since the higher level institution has itself to be designed by some even higher authority, one is involved in a infinite regression on institutions. Such a regression can only be stopped by considering a prior institution at some level. As concerns now the coselection problem, the chosen institution is unique when it is selected by a unique decision-maker, but may be multiple if several agents are involved in the choice. As concerns finally the cooperation problem, the chosen institution is optimal when resulting from a unique decision-maker, but may not be Pareto-optimal if resulting from a collective choice.

### 3.3. Inside procedure

An institution results here from a basic game which is played anyway for some exogenous reasons. Players reach an equilibrium state which incorporates some item able to be interpreted as an institution, such an item being of two types. On one hand, a player retains an available action which appears as an institution regulating the players' interactions. When this player has only to choose between adopting or not the action/institution, the equilibrium state is such that the first option is chosen. On the other hand, a player himself is considered as an institution which regulates some prior agents. Since this player can implicitly play the game or not, the equilibrium is such that the player/institution has an interest to play the game. Moreover, the institution concerns specifically the players involved in the game and is then naturally self-enforcing.

In a firm, the 'trust game' (Kreps, 1990) considers an employer (player 1) who can trust or not an employee (player 2), while the employee responds in a more or less opportunistic way :



In the static game, the subgame perfect equilibrium consists for the employee to cheat and for the employer not to trust. However, when the game is repeated, a dynamical Nash equilibrium consists for the employer to trust and for the employee to honor. It is supported by a threat of the employer to stop to trust if the employee cheated once in the past. Hence, trust appears as an institution which corresponds to an equilibrium action.

In a Middle Age town, the 'trade game' (Greif, Milgrom and Weingast, 1994) confronts alien merchants gathered in a 'merchant guild' (player 1) and the 'city ruler' (player 2). The guild accepts generally to trade, but may boycott; the ruler secures persons and properties, but might renege. The static game has the same structure than the trust game, but with some uncertainty involved. Its Bayesian equilibrium consists for the city ruler to cheat and for the merchant guild expecting this to boycott. However, in the repeated game, it exists a Bayesian equilibrium state in which the ruler protects and the merchants trade. It is again grounded on a threat: the guild trades as long as the ruler has not cheated before. The institution is symbolized by the guild which just creates a coalition of merchants able to trade. However, the reason why such a coalition holds is not incorporated in the model; it could be by considering that if a merchant does not boycott when he should, it is punished by the other merchants.

As concerns the implementation problem, it is still solved since the institution is either constituted or chosen by some player, at least implicitly. No meta-institution is needed since the players are spontaneously playing the game and designing the institution as a by-product. However, the institution is considered as already there, since it is assimilated to some action of a player or even to some player. When considered as an action, which concerns more precisely a relation between the players, it just receives a new interpretation. When considered as a player, which constitutes more precisely a grouping of prior players, it just receives a new justification. Hence, it is necessary to make more precise how an original action is conceived or how an original coalition is formed, for instance in an evolutionary framework. As concerns the coselection and the cooperation problems, the equilibrium state may again be multiple, and it is not automatically Pareto-optimal for the players.

## **4. Institution as eductive emergence**

### ***4.1. Principles***

In voluntary design, an institution results from a choice process in which the institution is directly introduced as a possible opportunity or at least as a correlated action or agent. But an institution frequently stays more implicit and appears only as an emerging item in some situation treated as a game. Such an institution is naturally assimilated by the modeler to an equilibrium state of the game. But a concrete process has then to be exhibited in order to state how such a state is constructively achieved. The same process is further interpreted in order to explain the genesis of the institution. An alternative would be to consider that an institution results from some process converging towards a collectively optimal state. But the main processes introduced in game theory lead to an equilibrium state since the resulting state has to be stable in order to stay unchanged.

From now, the institution will be considered as stemming as an equilibrium state from some well defined collective process. Once again, some specific features of the emergence of an institution look similar to features of the emergence of an equilibrium. For instance, the institution appears as some regularity emerging spontaneously and exclusively from the agents' behaviors and relations. The genesis of an institution needs much more reasoning and time than its disappearance and its evolution manifests some hysteresis. The emerging institution is favoured by some specific contextual conditions which may be gathered in such hand such historical period. The emerging institution looks generally stable at middle term, but may change very rapidly at longer term when the game structure evolves too.

A first approach proposed in game theory consists in an 'eductive process' which assumes that fully rational players achieve instantaneously an equilibrium state by their sole sophisticated reasoning. Stronger and stronger equilibrium notions rest on stronger and stronger conditions imposed to the players' beliefs, especially conditions of shared knowledge (everybody knows) or of common knowledge (everybody knows that everybody knows ...). An equilibrium state, looking like an 'equilibrium in beliefs', is interpreted as an institution obtained and applied in the very same game. It follows that such an institution is necessarily of a very specific kind, especially in a symmetric game. It corresponds generally to a 'behavior function' (or strategy) followed simultaneously by all the players at the equilibrium state and is interpreted directly as a social norm. It may also correspond to a belief shared simultaneously by all players and interpreted as a collective representation.

If the institution is now obtained by an indirect process, the last is more or less conscious, giving rise to two configurations. In an 'explicit process', the players' reasoning which leads to an equilibrium state is precisely conceptualized. By simulating each other, the players simultaneously define the equilibrium state and select one of them. Correlatively, the players infer the institution from various assumptions about the beliefs held about it by all players. In an 'implicit process', the players are not really aware of the reasoning which leads to some equilibrium state. However, one of them appears as privileged for strong reasons shared by all agents. Correlatively, the institution is selected by making explicit the underlying equilibrium state.

#### 4.2. *Explicit process*

In an explicit process, the following nested equilibrium notions are sorted out by the players' reasoning. With the minimal presupposition that the players have common knowledge of the game structure (players' choice determinants) and of the players' rationality (maximization of expected utility), one sorts out a 'sophisticated equilibrium' (obtained by sequential deletion of dominated strategies). Under the additional conditions that the players have common knowledge either that they play independently or that they have a common prior, one sorts out respectively a 'rationalizable equilibrium' and a 'correlated equilibrium'. But a Nash equilibrium can only be obtained under more drastic and unrealistic conditions. The players' conjectures (expectations on others' actions) have to be common knowledge or at least shared knowledge among them.

The conditions stated by Lewis for a convention (see § 2.2) can precisely be reformulated in terms of the eductive justifications of a Nash equilibrium. The first and the fourth conditions indicate that the players face a coordination game, the structure of which being common knowledge. The second expresses that the players' rationality is common knowledge and the third that their conjectures are common knowledge. In economics, 'money' (more precisely one instance of money) acts as a convention and can be grounded on the preceding conditions. Each agent prefers that more and more agents accept one money; he accepts one money if the other agents do so; each agent has a definite interest to accept one money; several supports for money are conceivable.

In linguistics, the 'communication game' (van Rooy, 2004) is assumed to justify a communication channel acting as an institution. It is a signaling game between a sender (player 1), a receiver (player 2) and nature (N). There are two (equiprobable) states of nature, harmless and harmful, the actual one being observed only by the sender. The sender can send a message, negative or positive, to the receiver. The receiver can make two actions, nothing or protect. The payoff matrix depends only on the state and on the receiver's action :

N / 2	nothing	protect
harmless	(1,1)	(0,0)
harmful	(0,0)	(1,1)

The sender has four strategies: always negative (NN), always positive (PP), negative iff harmless (NP), positive iff harmless (PN). The receiver has four strategies : always nothing (nn), always protect (pp), nothing iff positive (np), protect iff positive (pn). Among the six equilibrium states, only two are 'separating' (reveal the actual state): (NP, np) and (PN, pn).

As concerns the implementation problem, it is solved since the agents come to an equilibrium by sophisticated reasoning. Money or language are just shared by all agents in their interest and have no 'holistic' existence. As concerns the coselection problem, it stays present since the precise criteria which allow to select one convention rather than another are not precisely stated. But it is always possible to consider that the selection criteria are themselves higher level 'conventions' which are common belief. Some criteria apply to all games and are defined on the game structure: the selected equilibrium state is assumed to be either Pareto-optimal or symmetric or 'fair'. Some other criteria are specific to the environment of each game: the selected equilibrium state is conditioned on these contextual elements. Some final criteria are specific and grounded on some 'salient' features of the situation at hand: the selected equilibrium state appears as a 'focal point' grounded on background knowledge (Schelling, 1960).

### 4.3. Implicit process

In an implicit process, players come more unconsciously to some equilibrium state, but such an equilibrium state looks rather natural to the players. The institution is just an interpretation of the players' combined actions in that state and adopts two forms. On one hand, it appears as a behavior rule which expresses a conditional action coordinating the players: 'in context C, play action A'. On the other hand, it reduces to an indicative action which coordinates the players: 'play action A'. When the game is symmetric, the behavior rule or the action is the same for both players and acts then precisely as a 'social norm'. When the game is moreover a coordination game, the behavior rule or the action becomes a 'social convention'. Of course, in some cases, the equilibrium or even the game behind a social norm can be reconstructed.

For traffic regulation, the 'driving side game' considers two drivers (players 1 and 2) who drive right or left while the 'technology game' considers two drivers (players 1 and 2) who buy an electric or a petrol car (Sugden, 1986, 1989) :

1 / 2	left	right
left	(1,1)	(0,0)
right	(0,0)	(1,1)

1 / 2	electric	petrol
electric	(3,3)	(0,1)
petrol	(1,0)	(1,1)

In the driving side game, there are two equivalent equilibrium states (1,1). A 'coordination convention' just makes precise what side of the road to use : 'drive left' or 'drive right'. In the technology game (analogous to the 'stag-hunt game'), one equilibrium state (3,3) dominates the other equilibrium state (1,1). A coordination convention consists in recommending one type of car : 'choose an electric car' or 'choose a petrol car'. But the two conventions are not equivalent since the first is better for both players, but less cautious.

Always for traffic regulation, the 'crossroads game' considers two drivers (players 1 and 2) arriving at a crossroads and able to stop or to pass :

1 / 2	stop	pass
stop	(2,2)	(2,3)
pass	(3,2)	(0,0)

The game has now two asymmetric equilibrium states (2,3) and (3,2). In order to define a convention, the drivers have to refer to some 'label' relative either to an individual or to a contextual property. A binary label A/B (corresponding to player in line or to player in column) leads to the convention : 'if label A, stop ; if label B, pass' (or conversely). The labels receive several interpretations : coming right or left, having a big or a small car, coming North-South or East-West, driving on a big or a small road. Especially, 'traffic lights' are based on the label green/red (randomly attributed to the drivers) and constitute an institutional device which materializes a 'correlated equilibrium'.

As concerns the implementation problem, it is solved since the institution appears as the direct result of the players' actions. No meta-institution is needed since the players play some game and design an institution in the same game. As concerns the coselection problem, it is still relevant since the equilibrium states may be multiple, hence the possible norms or conventions. But a norm or a convention may be precisely sorted out in order to select one equilibrium state. In the driving side game, the 'focal point' on which a plain convention is established is exogenous to the game. In the crossroads game, the 'label' on which a contextual convention is conditioned is exogenous to the game. In order to deal with the emergence of a precise convention among possible ones, an evolutionary view was proposed

(Sugden, 1995; Binmore-Samuelson, 2004). The associated results stress that it is not always the most efficient focal point or the most efficient label which is selected.

## **5. Institution as evolutionary emergence**

### **5.1. Principles**

In an eductive approach, an institution can be directly identified to an equilibrium state by the modeler. This is possible since the agents are assumed to have a full cognitive rationality and to reason conjointly at the same time. However, concrete agents have only a bounded rationality which limits heavily their crossed knowledge and crossed simulation. In other respects, concrete agents may interact sequentially through time if the basic situation is repeated over several periods. Hence, game theory developed various dynamic processes which lead asymptotically to some equilibrium state. In these processes, the work of crossed beliefs is replaced by the work of time, under the condition that the repeated situation is (sufficiently) stationary. Accordingly, the progressive emergence of an institution is assimilated to the progressive emergence of an equilibrium state.

The Austrian school already considered that some regularity, later interpreted as an institution, may emerge from sequential interactions between agents, looks advantageous for them, hence is maintained by them. Menger (1892) assumes that (a kind of) money simply follows from barter exchanges in which one good is progressively privileged by the agents due to its weak transaction costs. Hayek (1973) asserts that the market (or at least market rules) spontaneously ensues from a cultural evolution process directed by a group selection mechanism. Schumpeter (1934) suggests that the firm (or rather the 'entrepreneur') straightforwardly results from the opportunity given to some individual to make profits by taking advantage of novelty.

A second approach proposed in game theory consists in an 'evolutionary process' which assumes that boundedly rational players realize an asymptotic equilibrium in a sequential way by some type of learning or evolution process. The equilibrium state depends on the structure of the game (players' determinants), on exogenous factors (nature's state) and eventually on initial conditions (prior beliefs). It is said to be structurally stable if it evolves in a continuous way when the social structure and the material environment of the game varies. It is said to be asymptotically stable if it returns to the equilibrium state when it were initially displaced in some neighbourhood of it. When corresponding to an equilibrium state, the asymptotic state of the process is now interpreted as an institution. It is expressed in terms of strategies (relating an action to the history of the game) and it corresponds again to some 'social norm'.

The players are essentially unaware of the institution they are generating together, which appears necessarily as an involuntary device. Two types of evolutionary processes are distinguished according to the information and cognitive capacities of the players. In a 'belief-based process' (or epistemic learning process), each player observes the past actions of his opponents, modifies his expectations about them and chooses an action in consequence. Accordingly, an institution is progressively emerging from a process of mutual adaptation and even of imitation. In a 'reinforcement process' (or behavioral learning process), the players observe only their past performances and just favour for the future the more performant actions. Accordingly, an institution is progressively emerging from a process of trial-and-error type.

## 5.2. *Belief-based learning*

A belief-based learning process is grounded on the following principles. First, a player is no more aware of the others' choice determinants, but observes partially their past actions. Second, a player has a bounded cognitive rationality since he expects the others' future action from past observations, by assuming that their behavior is stationary. Third, a player has a bounded instrumental rationality since, even if choosing his best response to the others' expectations, he acts myopically by considering only the next period. The best example is the 'fictitious play rule' stating that each agent transforms the observed past frequency with which an opponent uses an action into a probability that he uses that action in the future. Even if cycles are possible, the asymptotic state (defined under various stability conditions) is frequently a (mixed strategy) Nash equilibrium.

As an example of a 'distribution norm', Young (1993) considers the share of a crop between a landlord and a farmer. The basic game assumes that each player can propose three shares (in his favour): 30%, 50% or 70%. When the demands are compatible, they apply; otherwise, nobody gets nothing. When the game is repeated between populations of agents matching randomly, an 'adaptive play' process is applied. Each agent observes only the  $m$  preceding bargains (where the demands were compatible or not). Each agent considers only a sample of other's past demands of size  $s$  and computes the past frequency of these demands. Each agent forms a random expectation of other's future demand and chooses a best reply with probability  $1-e$  and a random demand with probability  $e$ . If the amount of information is sufficiently large and sufficiently incomplete ( $s/m$  lies under some threshold) and if the experiment process is sufficiently low ( $e$  lies under some threshold), the 'stochastically stable' state obtained asymptotically is (50%, 50%), hence an equal share of the crop.

In other respects, a belief-based process can be applied to belief diffusion and especially to the diffusion of opinions about the price emerging on some market. The basic game is frequently a mixture of two games. A majority game (like the 'El Faro bar game' or the 'beauty contest game') indicates that an agent wins if he has the same opinion than the others. A minority game (like the 'matching pennies game') indicates that an agent wins if he has a different opinion than the others. The agents are assumed to observe the past opinions, to expect the future average opinion and to react to it. A 'collective belief' may emerge, as was early recognized by Keynes for the future asset prices. In the same vein, a belief-based process can be applied to traffic regulation and some conventions may emerge. For instance, in the driving side game, all drivers converge finally on the same side of the road after a process where each driver always drives on the side with a majority of drivers.

As concerns the implementation problem, it is fully solved since an equilibrium state is obtained without reference to any prior or external entity. For instance, the obtention of 'distribution norms' implies no consideration of fairness grounded on prior moral norms since they just result from well understood interests. Likely, the formation of 'collective opinions' involves no holistic point of view since they are just shared opinions again sustained by players' interests. As concerns the coselection problem, it is likely solved since no idea of exogenous 'focal point' grounded on some cultural features is needed. For instance, a specific share of crops is selected because it is advantageous to both agents and stable in the long run. Likely, a specific belief is selected since it is adopted by all agents facing similar circumstances.



### **5.3. Reinforcement learning**

A reinforcement learning process is grounded on the following principles. First, a player has no more knowledge about the others and observes only the payoffs obtained with his own past actions. Second, a player has a weak cognitive rationality since he just computes an aggregated index of the past performance of each action. Third, a player has a very weak instrumental rationality since he chooses an action with a probability which is increasing with its index. The best example is the 'CPR rule' where a player computes the cumulated performance of each action and chooses an action with a probability which is proportional to that aggregated performance. The learning process converges rather easily towards a pure strategy Nash equilibrium.

For instance, the genesis of money can be explained by a reinforcement learning process (Sethi, 1999). In a toy economy, there are three goods which are indivisible and costly to store. Agents are of three types, each type producing only one good, consuming another good and storing eventually the third as a mean of exchange. At short term, given their strategy, agents meet randomly and exchange goods if profitable; at long term, they adapt their strategy, i.e. the good to be stored as money, according to their performance. The stochastic process converges towards one of two types of equilibria. In the fundamental equilibrium, it is the good with lowest storage cost which is chosen as money. Agents exchange either to obtain their consumption good or to diminish their storing cost. In the speculative equilibrium, one agent at least accepts some good which is neither his consumption good, nor the good with the lowest storage cost.

In other respects, a reinforcement learning process can explain the emergence of an 'altruistic norm' (Eshel, Sansone and Shaked, 1999). Players are essentially situated on a prior network and they observe and meet the others only locally. A player abandons his ongoing strategy if one direct neighbour at least adopts a different strategy. He imitates randomly, in a larger neighborhood, another player according to his past performance. He interacts finally with the others in an intermediate neighborhood. The process converges towards a strategy which looks altruistic, i.e. the strategy he would adopt in a modified game where his modified utility is a combination of his own and of the other's one. In other respects, for traffic regulation, a similar reinforcement process can be applied. For instance, in the driving side game, the drivers converge towards one side of the road if each driver drives more often on the side where he had less accidents.

As concerns the implementation problem, it is solved since an equilibrium state is achieved by the players acting alone in their own interest. In fact, in order to complete the spectrum of dynamic processes, a third process, the 'evolutionary process', was proposed and is based on biological principles. First, each agent observes nothing about his environment, except the payoff he gets in his interaction. Second, he has neither cognitive rationality nor instrumental rationality since he applies a fixed strategy. Third, he reproduces according to his payoff while original strategies ('mutants') can be introduced. The best example is the 'replicator dynamics' where an agent reproduces proportionally to his payoff. Convergence is rather easy towards an 'evolutionary stable equilibrium', where the agents resist to the introduction of a small proportion of mutants. However, such a process is now abandoned since it violates methodological individualism (it provides no concrete justification for the selection and mutation processes) and it is isomorphic to a (more realistic) reinforcement learning process.

## 6. Open problems

### 6.1. *The naturalization problem*

A first remaining problem consists in describing how an institution comes off the underlying equilibrium state in order to become a separate entity (Douglas, 1986). The so-called ‘naturalization problem’ can be decomposed into two successive steps involving specific cognitive capacities for the players. The ‘recognition step’ shows how the institution is no more considered as a regularity stemming from players’ actions, but appears as an autonomous entity with its own rules. Especially, some players may become aware of the regularity and are induced to diffuse its existence. The ‘focalization step’ shows how the players do not react anymore to the many others’ actions, but respond directly to the institutional signals by disconnecting with other players. Especially, the regularity may become a public information and is even considered as stable.

In case of a voluntary emergence, the naturalization problem is automatically solved. The recognition step is ensured since the target group is informed of the choice of an institution, even if the deciding group differs from it. The focalization step is ensured by the fact that the players respond directly to the institution, without prior direct relations. However, once an institution is implemented, a player has the remaining problem of interpreting how it applies in such and such context. He has to decide if a given institution is really relevant in some precise context and how he is supposed to apply it. For instance, a politeness norm is frequently ambiguous as concerns its precise domain and mode of application. Likely, care to person in danger is relevant when it concerns a wounded person in an accident, but not for a wounded opponent in a war.

In case of an educative emergence, the naturalization problem is still partly integrated in the emergence process itself. The recognition step is natural since the actors consider themselves mentally on an equilibrium state, even if the institution needs to be recognized at the same time that the equilibrium is achieved. The focalization step is ensured when the players break the specularity of crossed beliefs about the equilibrium state in order to react to the institution which summarizes its features in a limited set of signals. For instance, if some type of language appears in some group, it is recognized that everybody uses this very language and each responds naturally in the same language. Likely, if driving right is adopted by some players under some convention, other drivers become aware of the convention and use it in similar situations.

In case of an evolutionary emergence, the naturalization problem becomes far more ticklish. The recognition step is solved when some privileged players become aware of a regularity interpreted as an institution and diffuse the preceding observation towards the other players. The focalization step is treated when the players cease to adapt mutually their behaviors in order to respond to the institution which has been formerly recognized in its essential features. For instance, if a specific money is privileged in usual transactions, it is progressively recognized as an accepted mean of payment and is naturally stored or used by more and more other people. Likely, if a fairness norm is adopted in some place, it is considered as a good compromise by other people and spreads by analogy to other places or even to other situations.

## **6.2. *The normatization problem***

A second remaining problem consists in describing how an institution which is positively chosen may turn into a normative entity. The so called ‘normatization problem’ can be decomposed into two successive steps implying specific normative propensities for the players. The ‘legitimation step’ describes how the institution, even if already self-enforcing, turns into a moral obligation. Especially, some crossed expectations may be formed on how to behave, expectations which influence the players’ preferences. The ‘legalization step’ describes how the institution, even if already legitimated, becomes a law edicted by some kind of public authority. Especially, a whole system of incentives and sanctions enforcing it may be designed, which influences again the players’ preferences.

In case of a voluntary emergence, the normatization problem is partially solved by the procedure followed itself. The legitimation step comes from the fact that the deciding group is constituted by a higher level authority which is institutionally recognized by the target group. The legalization step is frequently achieved by the same decision group which is assumed to be able to impose the chosen institution by law. For instance, a pressure group (like the merchant guild) can participate in some game, becomes legitimate since it is performing well in the given situation and turns later into a publicly approved authority. Likely, some social attitude (like trust) is emerging naturally in some game, is considered as a legitimate basis for success, but it can less evidently be turned into an obligation.

In case of an educative emergence, the normatization problem becomes more relevant and can only happen after the naturalization problem is solved. The legitimation step is especially relevant when the assumptions leading to an equilibrium (sufficient cognitive and instrumental rationality) are not perfectly satisfied. The legalization step is exogenously implemented in order to make precise what institution is adopted when many are available. For instance, a language once recognized by its generalized use, is considered as the working language by a given group and may further be imposed as the official language of some country. Likely, a social norm (like priority to the right) may appear as natural in some game, is approved and diffused by an association of beneficiaries, then made compulsory by a regulation authority.

In case of an evolutionary emergence, the normatization problem is crucial and happens again after the naturalization problem is solved. The legitimation step reinforces the recognized institution, which is only a compressed and simplified representation of the equilibrium (Aoki, 2001). The legalization step is just achieved in order to strengthen once more the existence of the institution, reduced to the features which are strictly relevant for players’ actions. For instance, (an instance of) money appears progressively in some society, is first considered as the usual mean of payment, then receives an official status and a legal definition. Likely, when some rationed good is collectively distributed, a queue spontaneously forms, is recognized as an institutional device, becomes a legitimate device which has to be followed under psychical sanctions, and its precise form may even be imposed.

## **6.3. *Nested institutions and time scales***

The emergence models assume that some institution appears in isolation within a set of independent and exogenous institutions. However, the institutions are strongly related to each

other and are even defined hierarchically. The emergence models have then to be adapted to the fact that an institution always appears in a context formed of simultaneous institutions. On one hand, a model may treat together a whole family of related games leading to related equilibrium states, hence to related institutions. On the other hand, a model may treat in a single game a whole set of complementary or substitutive institutions (Milgrom-Roberts, 1990). For instance, marriage is not independent of other social institutions like resource allocation or nationality and they have to be dealt together.

The emergence models studies only the genesis of an institution and are silent about its further evolution and diffusion. However, the reasons for which the institution emerges are not the same than the reasons for which it stays unchanged or evolves afterwards. The emergence models have then to consider that some institution is developing, spreading and acting at different time scales. On one hand, a model has to explain the life cycle of a given institution, in relation with modifications of the underlying game. On the other hand, a model has to explain how an institution is able to switch towards another one, usually due to a change of equilibrium state (Boyer-Orl  an, 1992). For instance, a language develops all its life long for convenience reasons, but may disappear if other languages become dominant.

More generally, structural nestings and time scales are interfering in the dynamics of institutions. A general correlation indicates that high level institutions evolve more slowly than low level ones. More precisely, the emergence of some institution is conditioned by the development of higher level ones and favours the outcome of lower level ones. Symmetrically, the death of some institution may entail, by some ‘domino effect’, a correlative death in related institutions, at lower as well as higher levels. Moreover, some institutions may not be instantaneously compatible, such an incoherence creating constant pressure for their transformation. For instance, money is constantly evolving in its different forms and some kinds of money simply disappear.

Finally, the evolution of a set of institutions is conditioned by some common material and epistemic factors. Especially, some institutional innovations may be proposed and are rapidly diffused. Institutions may change together with an evolution of the agents’ determinants, of their demographic distribution, of their relations or of their common context. Institutions may be conceived by the players by analogy to existing ones in similar situations or selected by analogy to the performance obtained by similar ones (Douglas, 1986). Moreover, institutions are constrained by local political conditions, but the existence of divergent institutions from one country to another creates tensions which have to be attenuated. For instance, markets are created and evolve all the time, with mutual imitation due to competition.

**Acknowledgments :** I thank Pierre Garrouste and Andr   Orl  an for many suggestions concerning this paper

## References

- Akerlof, G. (1980) : A theory of social custom, *Quarterly Journal of Economics*, 94, 749-75.
- Aoki, M. (2001) : *Toward a comparative institutional analysis*, MIT Press.

- Axelrod, R. (1986) : An evolutionary approach to norms, *American Political Review*, 80, 1095-1111.
- Bergstrom, T.C. (1995) : On the evolution of altruistic ethical rules for siblings, *American Economic Review*, 85(1), 58-81.
- Berninghaus, S.K., Schwalbe, U. (1996) : Conventions, local interaction and automata networks, *Journal of Evolutionary Economics*, 6, 267-312.
- Bicchieri, C. (200 ) : *The Grammar of Society : the emergence and dynamics of social norms*, Cambridge University Press.
- Bicchieri, C., Jeffrey, R., Skyrms, B. eds. (1997) : *The dynamics of norms*, Cambridge University Press.
- Binmore, K. (1992) : *Fun and games*, Heath.
- Binmore, K. (1998) : *Game theory and the social contract I : Playing fair*, MIT Press.
- Binmore, K., Samuelson, L. (1994) : An economist perspective on the evolution of norms, *Journal of Institutional and Theoretical Economics*, 150, 45-63.
- Binmore, K., Samuelson, L. (2004) : The evolution of focal points, mimeo
- Boyer, R., Orléan, A. (1992) : How do conventions evolve? , *Journal of Evolutionary Economics*, 2, 165-177.
- Douglas, M. (1986) : *How institutions think*, Syracuse University Press, traduction française : Comment pensent les sociétés, La Découverte, 1999.
- Elster, J. (1977) : Critique des analogies socio-biologiques, *Revue Française de Sociologie*, 18, 369-95.
- Elster, J. (1989a) : *The cement of society*, Cambridge University Press.
- Elster, J. (1989b) : Social norms and economic theory, *Journal of Economic Perspectives*, 3(4), 99-117.
- Eshel, I., Sansone, E., Shaked, A. (1999) : The emergence of kinship behavior in structured populations of unrelated individuals, *International Journal of Game Theory*, 28, 447-463.
- Gilbert, M. (1990) : Rationality, coordination and convention, *Synthese*, 84, 1-21.
- Goyal, S., Janssen, M. (1996) : Non exclusive conventions and social coordination, *Journal of Economic Theory*, 77, 34-57.
- Greif, A. (1994) : Cultural beliefs in the organization of society : a historical and theoretical reflection on collectivist and individualistic societies, *Journal of Political Economy*, 102(5), 912-950.

Greif, A.- Milgrom, P.- Weingast, B. (1994) : Coordination, commitment and enforcement : the case of merchant guilds, *Journal of Political Economy*, 102(4), 745-76.

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

Hodgson, G.M. ed. (1997) : *The economics of institutions*, Edward Elgar.

Hurwicz, (1972) : On informationally decentralized systems, in C.B. Mc Guire, R. Radner, eds, *Decision and Organization*, North Holland.

Kandori, M. (1992) : Social norms and community enforcement, *Review of Economic Studies*, 59, 61-80.

Kreps, D. (1990) : Corporate culture and economic theory, in J.E. Alt - K.A. Shepsle eds. *Perspectives on positive political economics*, Cambridge University Press, 90-143.

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

Maskin, E. (1985) : The theory of implementation in Nash equilibrium, in L. Hurwicz, D. Schmeidler, H. Sonnenschein, eds : *Social goals and Organization*, Cambridge University Press.

Menger, C. (1892) : On the origin of money, *Economic Journal*, 2, 239-55.

Milgrom, P. (2000) : Putting auction theory to work, *Journal of Political Economy*, 108(2), 245-72.

Milgrom, P.- Roberts, J. (1990) : Rationalizability, learning and equilibrium in games with strategic complementarities, *Econometrica*, 58, 1250-78.

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

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

Opp, K.D. (2001) : How do norms emerge ? An outline of a theory, *Mind and Society*, 3(2), 101-128.

Orléan, A. (1997) : Jeux évolutionnistes et normes sociales, *Economie Appliquée*, L3, 177-198.

Orléan, A. (2002) : Les interactions mimétiques, in J. Lesourne, A. Orléan, B. Walliser eds, *Leçons de microéconomie évolutionniste*, Odile Jacob.

van Rooy, R. (2004) : Evolution of conventional meaning and conversational principles, mimeo.

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

- Schotter, A. (1981) : *The Economic Theory of social Institutions*, Cambridge University Press.
- Schumpeter, J. (1934) : *The Theory of Economic Development*, Harvard University Press.
- Searle, J.R. (1995): *The construction of social reality*, Free Press.
- Sethi, R. (1999): Evolutionary stability and media of exchange, *Journal of Economic Behaviour and Organization*, 40(3), 233-54.
- Shubik, M. (1959) : *Strategy and market structure*, John Wiley.
- Sugden, R. (1986) : *The economics of rights, cooperation and welfare*, Basil Blackwell.
- Sugden, R. (1989) : Spontaneous order, *Journal of Economic Perspectives*, 3, 83-97.
- Sugden, R. (1995) : A theory of focal points, *Economic Journal*, 105, 533-550.
- Ullman-Margalit, E. (1977) : *The emergence of norms*, Oxford University Press.
- Walliser, B. (2000) : *L'économie cognitive*, Odile Jacob.
- Walliser, B. (2002) : Justification des équilibres de jeux, *Revue d'Economie Politique*, 112(5), 693-716.
- Young, P.H. (1993) : Evolution of conventions, *Econometrica*, 61(1), 57-84.
- Young, P.H. (1996) : The economics of conventions, *Journal of Economic Perspectives*, 10(2), 105-122.
- Young, P.H. (1998) : *Individual strategy and social structure, an evolutionary theory of institutions*, Princeton University Press.