Thinking the Social dimension of the Artificial World: Ontological status of Collective Beliefs

Denis Phan [1] Jacques Ferber [2]

[1] Doctor in economics, Administrator P&T, CNRS senior researcher,
GEMAS, UMR 8598 CNRS / University of Paris IV Sorbonne
& CREM, UMR 6211 CNRS / University of Rennes I
http://perso.univ-rennes1.fr/denis.phan/
Email: denis.phan@univ-rennes1.fr
[2] Professor of Computer Sciences, University of Montpellier III
LIRMM, UMR 5506 CNRS / University of Montpellier III
http://www.lirmm.fr/~ferber/
Email: ferberlirmm.fr

Abstract

This paper deals with the modeling of the social dimension of artificial societies with cognitive agents. We provide a critical review of the ontological status of collective beliefs in some recent works in game theory, using a heuristic diagrammatic framework, useful for MAS design, called the "4Quadrant". We introduce first a formal presentation of the problem of the "collective" versus "social" dimension of the beliefs, which states the problem more clearly, in particular from the point of view of software design ontology. This raises the problems of the existence of some "social entities" to be able to acquire some relative autonomy with respect to the agents. The "4Quadrant" allows us to discuss, for a given world / system of the ontological commitment and the articulation between both the individual and the collective on the one hand, and between the observable (reifiable) and the subjective on the other hand The last section provides an empirical counterpart, by reviewing and discussing some experiments from behavioral game theory.

Keywords: artificial societies, behavioral game theory, cognitive agents, collective belief, coordination games, complex adaptive systems, intelligent agent's ontology, emergence, experimental economics, learning in games, multi agent systems, population game theory, social ontology.

1. Introduction

1.1 Collective beliefs or social beliefs?

In his recent work Orléan (2004) proposes a new definition of what a "collective belief" is. From our point of view, this initiative is of particular interest for two reasons. Firstly, it proposes a formal presentation (logical) of the "social" dimension of the beliefs, which states the problem more clearly, in particular from the point of view of software design ontology. Secondly, it raises the problem of the existence of some "social entities" (or "social objects": here the "social beliefs") to be able to acquire some relative autonomy with respect to the mind of individual agents of which they are the necessary support. This question is of great importance for the modeling of artificial society, in particular for the so-called "immergence" (Gilbert, 1995; Dessalles, Phan, 2005; Dessalles, Ferber, Phan, 2007)

The existence of such abstract entities in game theory (and more largely in economics and social sciences) is far from being unanimously established. In particular, according to the methodological individualism approach, standard game theory generally does not use such "social" entities. The explicit objective of the quoted work by Orléan is to promote the interest of such social concepts within the framework of "cognitive economics" (Bourgine and Nadal, 2004, Walliser, 2004b). The ontological status, or mode of existence, of such entities is discussed among philosophers (among those who seek ontological bases with the social entities, one will quote: Gilbert 1987, 1989, 1990, 1996 1997, 2000; Pettit, 1993, 2000, 2004; Searle, 1995; Tuomela, 1990, 1992 without forgetting Descombes, 1996, on a related topic) and among sociologists as well. In this latter field, Boudon, (1998, 2003, 2006) contests the possibility of the existence of extra individual entities, but according to (Durkeim, 1895), there is an important tradition in which "social facts" must be considered objectively "like things". See also Douglas, (1986) for a reinterpretation and (Livet, Ogien, 2000) for a synthesis. Our aim is to propose a new view of the modeling of social phenomena using multi-agent systems. This involves the *reification* of several concepts and thus imposes an ontological commitment, i.e. an analysis about the ontological status of the "social facts". In order to examine the ontological status of these collective beliefs, we will

introduce first a general map, called the "4Quadrant", which helps to understand individual and social systems and their interrelations.

Then, we will present the common views about the ontological status of collective beliefs using Orléan's typology and the 4Quadrant model. We will then propose a set of examples that come from game theory and experimental economics

1.2 A multi-perspective approach on societies

In order to analyze complex social systems, we will use what we call the 4Quadrant approach. This diagrammatic framework, inspired from those of (Wilbert, 2000) is designed in order to provide a two-dimensional heuristic description of the complex relationship within social systems This approach resides on a decomposition along two axes: the individual vs. the collective perspectives on the one hand, and the interior (i.e. mental states, representations) vs. the exterior (i.e. behavior, objects, organizations) perspectives on the other hand. These two axes taken together provide a four quadrant map where each quadrant must be seen as a perspective by which individuals, situations and social systems may be understood, as it is shown on figure 1.
The I-I (*Interior-Individual*, upper left) quadrant, is about emotions, beliefs, desires, intentions, of an individual, i.e. about his/her mental states, his/her subjectivity. The E-I (*Exterior-Individual*, upper right) quadrant describes physical bodies, concrete objects, and also behaviors of individuals. The I-C (*Interior – Collective*, lower left) is about shared knowledge and beliefs, collective representations, ontologies, social norms, and represents the inter-subjective part of individuals, what could be called the noosphere. The E-C (*Exterior-Collective*, lower right) is about material or formal social structures such as institutions and organizations, i.e. collective forms and structures of groups and systems, what could be called the sociosphere.
For instance, Internet, from a purely technological point of view, may be considered as an entity of the E-I quadrant, something which may be analyzed as a network of computers which operates using a common set of communication protocols. But Internet is now at the core of our society. It is not only a technological object, but also a medium which brings forth and maintains social interactions. It may thus be considered as a social object which is deeply immersed into our society: Internet is now part of the organization for our daily life. As such it is part of the E-C quadrant.

Figure 1. The 4Quadrant map.

Interior-Individual (I-I) *subjectivity* <Mental states, emotions, beliefs, desires, intentions…> Interiority	**Exterior-Individual (E-I)** *Objectivity* <Agent behavior, objects, process, physical entities> Objectivity
Interior-Collective (I-C) *Inter-subjectivity* <Shared/ collective knowledge invisible social codes and implicit ontologies, informal norms and conventions…> Noosphere	**Exterior-Collective (E-C)** *Inter-objectivity* <reified social facts and structures, organizations, institutions…> Sociosphere

One may like, dislike, have beliefs about Internet, and those mental states (love/hate, beliefs, etc.) will be part of the I-I quadrant. Groups may have also views and ideas on the importance, interest, effectiveness of Internet, i.e. have collective representations about Internet. There are also conventions and rules that govern the use of Internet, which show how Internet should behave and how it should be controlled. Although (formal) parts of these rules are E-C based, content of the rules as well as related ideas and concepts are within the I-C quadrant as any shared / socialized forms of ideas and symbolic features.

We will use this conceptual map throughout this paper, as a heuristic way to analyze the issue of collective beliefs and how different theories attempt to solve it. In particular, the two dimensional opposition between individual and collective, on the one hand, and between objective and subjective on the other hand enable us to discuss numerous controversial issues, such as methodological individualism versus holism (for a recent debate including the supervenience / reduction dimension cf. Kinkaid, 1986; Tuomela, 1990) as for the question of the "Interiority of mind" (Wittgenstein, 1958; Bouveresse 1976).

2. How beliefs can be "social"?

There are a lot of circumstances where the common vernacular language (folk) assigns belief to a group or a collective entity. In a seminal paper on the subject, (Tuomela, 1992) gives some examples such as: "The Catholic Church believes that miracles exist" or "Our family believes that the schools in this country are inefficient". In this paper, Tuomela distinguishes between normative and non-normative (factual) group beliefs. Normative group beliefs are concerned by *cooperative approach*, because there is "a joint acceptance of views by the group members (..) creating group commitments for all the members to accept (and keep accepting) the view in question". By opposition, non-normative group beliefs rely on a *non-cooperative approach*, i.e. the question of the distributions of *personal beliefs* within the group. Our concerns are of the second type. How such "personal" belief can be "social"?
According to the Weberian tradition and the methodological individualist approach in sociology (Boudon 2003) argues that a collective belief must *make sense* for an *individual ideal type*. The key psychological assumption is that belief must be founded on some "good reasons" that could be apprehended and understood (*verstanden*). In other words, motives could be reconstructed rationally (in a large cognitive sense). In that case, collective belief arises when personal beliefs are founded on *similar reasons* from an individual to another, despite personal idiosyncratic characters. From this point of view, the *collective adoption* of such belief alone is enough to grant the "social" nature of the belief. Then, the common "good reasons" are the *mechanism* (Hedström, Swedberg, 1998) by witch this collective adoption arises. Boudon calls "trans-subjective" such shared "good reasons". The underlying process can be viewed as a "top down" (or weakly) emergent phenomenon from the interactions between individuals (Dessalles Phan, 2005). However, a controversial question is the possibility of a purely cognitive phenomenon leading the individuals

separately by a purely internal process without communication to the same "good reasons" (such enigmatical process could be related to the different variants of *solipsism* in philosophy). No-cooperative game theory deals with some of these situations, where agents interact only by actions, not by linguistic communication. In the following, we discuss these questions within this perspective. Ontological basis of artificial intelligent social agents would take benefits by focalizing discussion on this highly formalized approach, completed by the existence of quasi-empirical data about the behavior of real people i.e. in the *behavioral game theory* experiments (Camerer 2003).

2.1 From shared beliefs to common beliefs.

Orléan introduces four definitions of collective/social dimension of beliefs within a group. The "classic" two first definitions correspond to the mainstream view, where the collective is considered only as a *set of individuals*. At the first level, the basic "shared belief" – also called "first degree of shared belief" makes it possible to build the "common belief" in a recursive way. The "common belief" concept can be seen as a reinterpretation of the concept of "common knowledge" within the domain of beliefs. Shared beliefs enable us to build more sophisticated beliefs in the way of recursive construction.

For Orléan "a proposition P is a shared belief for a group if and only if every member of the group believes that P is true". This definition is said to be "first degree of shared belief" because the definition states that each one believes P, but does not necessarily have some beliefs about the beliefs of others. More formally, let us introduce the following notation: C_iP means: "agent i believes that the proposition P is true" that we will simplify in: "agent i believes P", where index i relates the belief C to the agent i. Notation $I \in G_e$ means that an individual i is a part of a group G_e (where G_e is nothing but a *set of agents*). The number of agents within the group is denoted by $N_e \equiv \text{card}|G_e|$. Given these notations, there is a *shared belief* for a set of agents G_e when one has:

$$(1) \qquad C_iP \text{ for all } i \in G_e$$

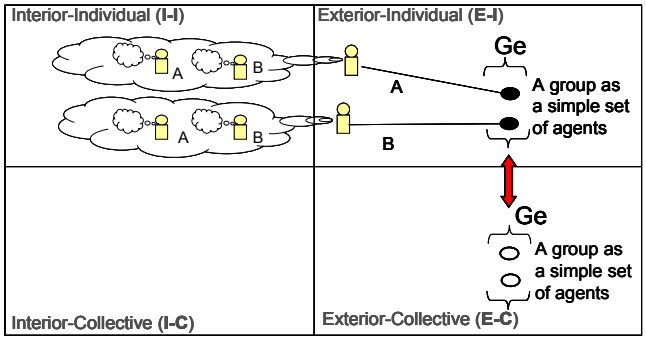
Which states: "every agent within G_e believes P". This definition of the shared belief is homologous to that of the "mutual knowledge" in standard game theory (one just weakened it by replacing "knowledge" by "belief"). By introducing beliefs of the second degree, i.e. beliefs about the beliefs of the others (Stahl and Wilson, 1995), one can define a "second degree of shared belief" by a situation where we have the conditions for the first degree shared beliefs and in addition where every agent in the group G_e believes it to be as such. The "common belief" can be built recursively as the situation where the set of agents G_e is such as:

$$(2) \qquad C_iQ \; ; \; C_jC_iQ \; ; \; C_kC_jC_iQ \; ; \ldots \textit{ ad infinitum} \quad \text{for all } i,j,k\ldots \in G_e$$

Then, "P is a common belief if everyone believes P, everyone believes that everyone else believes P, everyone believes that everyone else believes that everyone else believes P, and so on ad infinitum" This definition of common belief is thus homologous to that of *common knowledge* in standard game theory[1] (Aumann, 1976).

If we decompose the definition of common belief into hierarchical levels, the "common sub-belief" of level (2) is just a shared belief defined on the common sub-belief of level (1) which is simply a standard shared belief. In the same way, the common sub-belief of level (3) is a shared belief defined on the common sub-belief of level (2), etc.
The common characteristic of these two definitions is that they are strictly based on individual beliefs: "They are only collective insofar as all the individuals in one group believe them, in one way or another" Orléan said. From this point of view, they are in conformity with the definition of the strict methodological individualism retained by Boudon (2003). Figure 2 shows the conceptual view of these shared beliefs using the 4-Quadrant approach. A group, which is merely a set of agents, is composed of agents who have beliefs. Common beliefs are simply fixed points in the subjective beliefs (and beliefs of beliefs) about something

Figure 2. Representation of common beliefs using the methodological individualism approach.



2.2 Interpreting "Social" beliefs of first level

Orléan introduces later a *collective entity* called the "group" that we will note G_s to distinguish it from the *set of agents*, G_e. This entity is not a real actor like the other agents, which means that it cannot effectively have beliefs on its own. Although a group is not a human being, agents tend to attribute "beliefs" to such an abstract entity as they would do to individual agents. This can be formalized by using the notation G_sQ which means "the group G_s believes that the proposition Q is true". Orléan qualifies this assertion a *social belief*, because "beliefs are attributed to an *abstract entity*, namely the group itself" (Orléan 2004, p200, underlined by us).

[1] Many Scholars have discussed the role of knowledge in coordination, from (Hume, (1740) to (Schelling), (1960). But Philosopher David Lewis was the first to provide an explicit formal analysis. A common knowledge isf such as: (1) everyone knows it, (2) everyone knows that everyone knows it, (3) everyone knows that everyone knows that everyone knows it, and so on *ad infinitum* (Lewis, 1969). Other Scholars (including Aumann) independently provide alternate definitions of common knowledge. More recently, Margaret Gilbert proposed an alternative approach of the question. (Gilbert (1989) and (Cubitt and Sugden, (2003) have argued that Aumann's and Lewis' visions of common knowledge are radically different and irreconcilable.

The first justification given by Orléan for introducing this "enigmatic" abstract entity is empirical and relies on the analysis of ordinary speech. There are many situations where individuals attribute anthropomorphic characteristics to abstract collective entities. One of them can be found in Orléan's example of propositions of the type: "the market believes that shares are under-valued". This can be considered as a *cognitive shortcut*: nobody believes that strictly speaking "the market" has any beliefs. But using such cognitive shortcut could introduce cognitive biases. It would be interesting to explore the collective consequences of such individual cognitive biases, as we show in the next subsection.

With the possibility of attributing beliefs to a group seen as a social entity, Orléan proposes two new categories of collective beliefs, that we call "social beliefs". According to the following definition[2],

Definition (d1) : "an individual i believes that the group G_s believes a certain proposition P insofar as he believes that a *great part* of individuals in the group G_s believe P"

Orléan does not propose a formal definition, but here is one of the possible interpretations:

$$(3) \qquad C_iC_{Gs}P \Leftrightarrow_i C_iC_{Gpi}P \equiv (C_iC_jP \text{ for all } j \in G_{pi} / N_e \geq \text{card}|\, G_{pi}\,| > N_{pi})$$

Which means: one can say that "agent i believes that the group G_s believes P" if, in an equivalent way for him ($\Leftrightarrow_i$) it believes that an amount of members, higher than a *specific threshold,* N_{pi} (or possibly "*the majority*"), also believe it. We will note $C_iC_{Gpi}P$ (b1)(agent i believes that a "*sufficiently large*" part of the group, G_{Pi} believes P). In this formalization, the basic reference relies on the individual beliefs on P (C_iC_jP for any $i \in G_{Pi}$) and not on the group as an identified social entity. The conditions of equivalence ($\Leftrightarrow_i$) with the social belief of the agent i, namely: $C_iC_{Gs}P$ (b2) must still be specified. Our formalization goes beyond the seminal suggestion of the informal paper by Orléan, since our notation involves that the beliefs of agent i *explicitly integrate* a perimeter of agents G_{Pi} which is based upon a subjective threshold specific to the agent i. This makes possible to rely on both social and individual beliefs models. This subgroup G_{Pi} "exists" only in the eyes of an agent, and therefore is an *ontologically subjective entity* specific to each agent (Figure 5). In this model, individual meaning of the "group" is distinct from the common reference to the group as a whole (noted G_s). André Orléan points out that this first "social" concept of collective belief is close to the first degree of shared belief.

If one considers only the right part of the $C_iC_{Gpi}P$ equivalence, one could even estimate that this form of collective belief is only a weakened form of shared beliefs. It is true that, G_{Pi} can be reduced to a list of agents, but the right part $C_iC_{Gs}P$ is of *another nature* which would have *to be specified*, just like the mechanism which makes it possible for each agent to see the left and right parts of (3) as equivalent.

This operation substitutes a single abstract entity for a set of agents. The notation used here allows us to highlight the "fuzzy" logic subjacent with the operation. This formalism translates a first manifestation of the "cognitive shortcut" which was evoked previously. More specifically, the fuzzy appears in particular in the passage regarding the individualized sub-group C_{Gpi} which is a *subjective construction specific to the agents*, which has the statute of a personal belief upon concrete entities towards the abstract notion of the "*group as a whole*"; thus, G_s is a "collective entity" theoretically recognized by all, even if by definition it is not necessary that this shared reference relates to the same evaluation of the equivalent requirements for the agents (i.e. the sets G_{pi} and the N_{pi} thresholds can differ according to agents). In other words, when one speaks about the"group" (or "market") as a collective entity, everyone makes *as if* this word indicated without ambiguity the same thing for everyone, although these terms can relate to different representations according to agents

The vagueness of the threshold concept of "great part" used by Orléan (2002), gets more precise in his later work (Orléan 2004), where the relevant threshold becomes "the majority". The examples studied later in this paper suggest that the concept of "great part" (to be specified, as in our formalism for instance) has two advantages in comparison with the concept of majority. The first one is the possibility to introduce a heterogeneous threshold among agents, the second one is the possibility of being able to capture the idea that there are other critical thresholds in the groups than the majority. Both have the disadvantage of making the thresholds very dependant on the context.

2.3. The consequences of the introduction of the collective dimension into individual's schemes of learning and decision.

It is well known the dynamics of a set of agents in interaction, seen as a *complex adaptive system*, has specific properties which generally excludes the use of an agent representing the set of agents (Anderson *et al*, 1988; Weisbuch, 1990; Arthur *et al*., 1997, Schuster 2001, Nataname, 2005). Let us illustrate that feature with a very simple model of opinion formation. Since the seminal work of (Weidlich and Haag 1983), many models have been devoted to the role of social influences on the formation of individual opinions. The model presented hereafter is the simplest form from a class of models of social influence drawn by (Orléan, 1998a, 1998b, 1998c). In this class of models, the access of agents to public information and the taking into account of such information may enhance the convergence of agents towards "right" beliefs (i.e. adequate with the "real" state of some variable) but, if the agents give too much importance to the opinion of the others (the public information), they can also be collectively locked-in within a situation where the majority of agents make a choice which does not correspond to their personal conviction about the state of this variable (i.e. inadequate with the "real" state of this variable)

[2] This is a definition introduced in (Orléan 2002). In (Orléan 2004) "a great part" became "the majority". Although both concepts are unclear, it is possible to claim that the first subsume the second as particular case.

Analytical results are not detailed here, focusing only on the main intuitions of the model, illustrated by an example of simulation that has been conducted on the multi-agent platform "Moduleco-Madkit" (Gutknecht and Ferber, 2000; Phan, 2004; Michel et al., 2005). In this model, there is a population of N agents ($i \in A_N \equiv \{1,\ldots N\}$). Each agent has beliefs about the state of its world. Let us suppose that such a state may be represented by binary values, say $M \in \{0,1\}$. Each agent receives *private information* about the state of the world and observes *public information*: namely the average of the choices of its "neighbors" (on a neighborhood that has to be defined). Each agent has a probability p (which is the same for all agents) to receive the "good" private information (i.e. the real state of the world) and a probability $1-p$ to receive the "false" private information:

$$(4)\quad \begin{aligned} &proba(\sigma_i(t)=1|M=1) = proba(\sigma_i(t)=0|M=0) = p \\ &proba(\sigma_i(t)=0|M=1) = proba(\sigma_i(t)=1|M=0) = 1-p \end{aligned}$$

For each iteration, an agent is activated at random. The "active" agent must choose a state $\omega_i \in \{0,1\}$. When its private and public information are the same, there is no difficulty of choice. If its private and public information are contradictory, it is in a situation of cognitive dissonance (Festinger 1957). Then, it must arbitrate between these two contradictory values. In this case, it is assumed that an agent adopts the majority of the choices of its neighborhood with a probability μ. and maintains a coherent choice with its private information with a probability $1-\mu$. Accordingly, μ. is the probability of taking solely into account the private information. Then, $1-\mu$ is the probability for following the majority rule, and to imitate the members of this majority. When $\mu = 0$ only private information matters. When $\mu = 1$, agents are fully mimetic. Let us denote by $q_u(\sigma_i, \eta_i)$ the probability that an agent with a coefficient of imitation μ chooses $\omega_i = 1$ when it observes private information σ_i if the average choice of its "neighbors" is $\eta_i(t)$:

$$(5)\quad \begin{aligned} &si\ \eta_i(t) < 0{,}5;\ q_u(1|\eta_i(t)) = (1-\mu);\ q_u(0|\eta_i(t)) = 0 \\ &si\ \eta_i(t) > 0{,}5;\ q_u(1|\eta_i(t)) = 1 \qquad ;\ q_u(0|\eta_i(t)) = \mu \end{aligned}$$

Let us first consider the case of global influence: equations [4] and [5] allow us to compute the Markovian probability of transition towards 1 (When the state of the world is $M = 1$) :

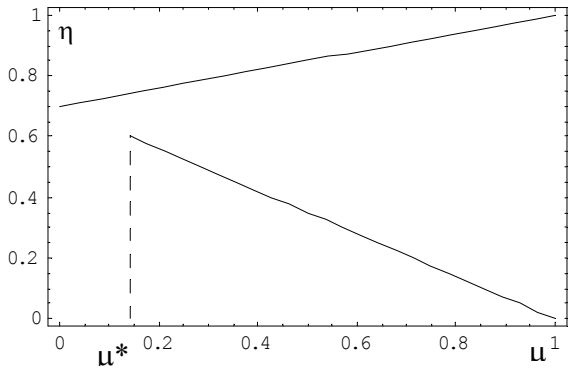
$$(6)\quad p^1_{0\to 1}(\eta(t)) = P^1(\eta(t)) = p.q_\mu(1,\eta(t)) + (1-p).q_\mu(0,\eta(t))$$

This model has an exact (e.g. analytical) solution when $\eta_i(t)$ is the same for every agent, i.e., when there is a complete connectivity between the members of the population, or when public information comes from an average value $\eta(t)$ resulting from the actions of all the agents. This model is *ergodic* and admits an *invariant distribution of probability* (which gives the asymptotic probability for the system to be in a given state). Standard method of resolution of the stationary solutions by the way of the master equation (Weidlich and Haag 1983) gives us two modes (a bimodal distribution with two extrema of the invariant distribution) for important values of the mimetic coefficient, beyond a *critical value* μ^* (equations 7 and figure 4):

$$(7)\quad \begin{aligned} &\mu^* = 1 - 1/(2.p) \\ &\eta_{s+} = p + (1-p)\mu \qquad \eta_{s-} = (1-\mu)p \end{aligned}$$

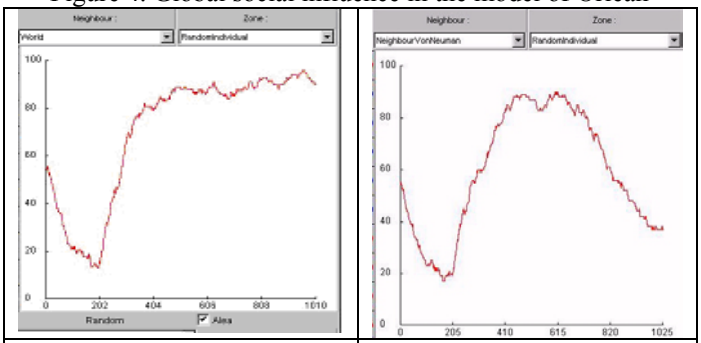
Then, when external effects (social influences) are strong enough, a *broken symmetry phenomenon* can arise. There is a bifurcation at μ^* from a dynamical regime with one attractor to a regime with three attractors (one unstable, two stable).

Figure 3. Extrema of the invariant distribution in the simplest Orlean's model of opinion dynamics



This result is asymptotic, hence not very informative for a short term dynamics. Simulation exercises allow us to make a short term dynamics exploratory, given the long term results. In the short term, simulation provides evidence that agents could be locked-in for a significant period of time (*broken ergodicity*) in a different state than the one resulting from individual preferences (or information) taken in isolation[3]

Figure 4. Global social influence in the model of Orléan



global social influence | localized social influence

In the simulations exhibited on figure 4, at time $t_0 = 0$, 50 % the agents believe that this state is, say 0 (the "white" ones); for all $t>0$ an agent taken at random receives private information on the state of the world and can change his opinion. The "theta" parameter gives the proportion of coherent signals with a world state at "1". In this simulation, theta = 0.1 for both $0 < t < 200$ and for $600 < t$. Between $200 < t < 600$, theta turns out to be 0.9. For the mimetic parameter, $\mu = 0.1$ for all agents until t = 600; after which, $\mu = 0.9$ for all agents, so they became mimetic. As a result, the social opinion is trapped" in the dominant opinion of state {1}, despite 90% of the private information, which suggested that agents are in state {0}. The following simulation results are given for 3 neighborhood structures: World (the original

[3] See (Phan and Gordon and Nadal, 2004) for a statistical mechanics interpretation of this issue.

model) Moore (8 neighbors) and VonNeuman (4 neighbors). In these cases, the initial distribution is a pseudo-random distribution generated with the same seed, but the dynamical random process of learning is different. Results are quite similar for World and Moore. In the Von Neuman case, each agent has only four neighbors, so it is easier to get out of the trap.

3. From individuals to Social: first account.

3.1. Auto-referential social beliefs of the second level in a population game framework.

The last definition (Orléan, 2004, p.200) is right away auto-referential:

Definition (d2) : "individual i's believes that the group G_s believes the proposition Q if, in the majority the members (for us: a great part) of the group believe that the group G_s believes the proposition Q".

The auto-referential nature of this definition arises from the indeterminacy associated with the reference to beliefs of the second degree: since the members of the group do not limit themselves any more "to believe Q", but "believe that the group believes Q"

$$(4) \qquad C_iC_{Gs}C_{Gs}Q \Leftrightarrow_i C_iC_{Gpi}C_{Gs}Q \equiv$$
$$(C_iC_jC_{Gs}Q \text{ for all } j \in G_{pi} / N_e \geq \text{card}|\, G_{pi}\,| > N_{pi})$$

It should be noted that this notation induces beliefs of the type (3) of the second degree upon the relation of equivalence of agent i's himself, which expresses a subsumption of the first degree. It means that we already have a first degree replacement of the beliefs on the set of individuals G_{pi} by a "belief" attributed to the social entity G_s. The main advantage of a direct introduction of beliefs attributed to a collective entity C_{Gs} can be found at the cognitive level: it allows to save computational capacities, in particular to avoid infinite depth recursive calculus, as in the case of common beliefs. On the other hand, the price to pay is the indetermination related to the auto-reference of the second degree: "the group believes that the group believes". But this indetermination is also one of the sources of the relative autonomy of this "social" beliefs. There is autonomy because the action capacity of individuals on a collective object (which is subjective by nature) decreases as the group grows. This can be shown with formal models of social interdependence where a set of interactions may be regarded as a complex adaptive system (Lesourne et al, 2002, Phan et al. 2004). But this kind of autonomy is only relative. This collective entity is only an emergent property of the interactions among agents' subjectivity, and such subjectivity remains the necessary support of this social entity. A group would not exist without its members and their shared beliefs. The dynamics of a group depend on both the "external" interactions between its agents (upper right quadrant) and on their "internal" reasoning and especially the *subsumption process* (upper left quadrant). We will show this by considering population games in the light of experimental economy.

Let us now consider the case of a "population game" (Blume, 1997) where each player is paired with another player randomly chosen (without redundancy) in the population. The players must play a "pure strategy", i.e. they must choose between A or B. What does occur if the players have a priori beliefs on the distribution of the strategies in the population? Let us suppose that player i believes that there is a proportion α of players who play A in the population (of complementarily: $(1-\alpha_i)$ who play B). The expected payoff of his strategy in a random pair-wise framework is thus equal to $\pi_i(S_i, \alpha_i) = \alpha_i.\pi_i(S_i, A) + (1-\alpha_i).\pi_i(S_i, B)$

The reader could check that if player i chooses his best response with payoff defined by the equation (8), then he will choose A if $\alpha_i > 0{,}5$ and B if $\alpha_i < 0{,}5$. It is said that he chooses his best response against the "mixed strategy" α_i. It is important to remark that each agent i has a specific belief: α_i and there is no reason for which this belief should be the same for all agents. Players must choose a pure strategy in a population game with random pair-wise and it looks as if they dealt with only one player playing A (respectively: B) with a probability α_i (respectively: $1-\alpha_i$). In the case of an iterated population game, players can "learn" from their experience. When agents revise their beliefs at each step of the iteration process, these beliefs will change according to the game history. If the environment is stable, there will be a convergence towards a Nash equilibrium (A, A) or (B, B), depending on both the initial distributions of beliefs and the game history.

It is possible to view the initial beliefs as coming from the "culture" of players and the revision of beliefs as resulting from the "history". Then, the equilibrium selection results from historical and cultural factors, in a proportion which depends on the respective part of such factors in the behavior. For example, English people drive on the left part of the road, whereas the continental people drive on the right.

Let us remark that the beliefs described above can be formalized with the notation used in (3) for social beliefs. The reader can verify that if player i believes that a "great part of the group" believe that the best response is to play A ($\text{card}|G_{pi}|/N_e \equiv \alpha_i > 0{,}5$ if one adopts the only majority one: $N_{pi} = 0{,}5$), then the best response will actually be to play A. While noting $Q \equiv (\tilde{S} = A)$ and $C_j(\tilde{S} = A) \equiv$ (the player j believes that his best response is to play A) one have:

$\tilde{S}_i = A$ si : $C_iC_{Gpi}(\tilde{S} = A) \equiv (C_iC_j(\tilde{S} = A)$ for all $j \in G_{pi} / 1 \geq \text{card}|G_{pi}|/N_e > 0{,}5)$

But also, in the case of subsumption :

$$(9 \text{ bis}) \ \tilde{S}_i = A \text{ if: } C_iC_{Gs}(\tilde{S} = A) \overset{i}{\Leftrightarrow} C_iC_{Gpi}(\tilde{S} = A)$$

Why do these beliefs come "from the first level"? Given our previous assumptions, let us assume that the beliefs of agent i on the beliefs of j are exact. If the player j believes that it is necessary to play A : $C_j(\tilde{S} = A)$, it is because he believes that there is a majority G_{pj} in the group who plays A ($\alpha_j > 0{,}5$) then: $C_jC_{Gj}(\tilde{S} = A)$ and under the conditions of subsumption: $C_jC_G(\tilde{S} = A)$. If we do the same reasoning with the beliefs of agent i, we can reformulate our assumptions on the initial beliefs as social beliefs of the second degree according to (4):

$$\tilde{S} = A \text{ if: } C_iC_{Gpi}C_{Gs}(\tilde{S} = A) \equiv (C_iC_jC_{Gs}(\tilde{S} = A) \text{ for all } j \in G_{pi} / 1 \geq \text{card}|G_{pi}|/N_e > 0{,}5)$$

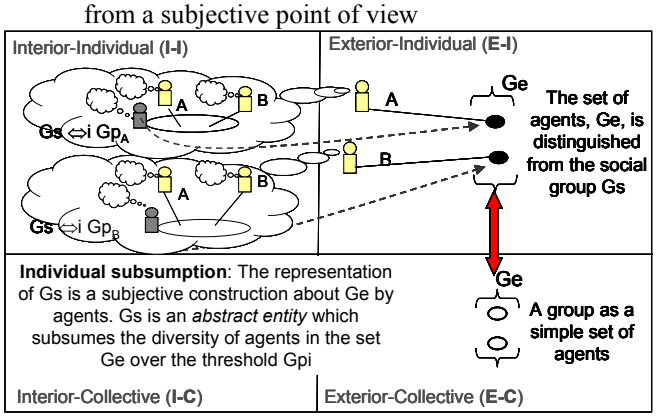
But also in the case of subsumption (auto-referential belief):

$$\text{(10 bis)}\quad S_i = A \text{ if: } C_iC_{Gs}C_{Gs}(\tilde{S} = A) \overset{i}{\Leftrightarrow} C_iC_{Gpi}C_{Gs}(\tilde{S} = A)$$

3.2. From individual subsumption towards social reification of collective entities

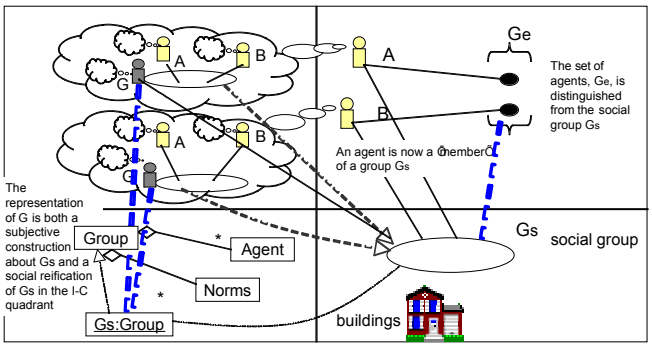
We have distinguished two levels of beliefs involving the meaning of "group". On the one hand, (b1) the individual beliefs $C_iCG_{pi}P$ concern the subgroup G_{pi}, which is an ontologically subjective construction (the subsumption) made by each agent separately (I-I quadrant). Thus, this construction appears in the I-I quadrant (figure 5). It can be seen that while there are representations about the group Gs, these representations are still subjective.

Figure 5. The construction of a collective entity, seen from a subjective point of view



On the other hand, with (b2) $C_iC_{Gs}P$ the beliefs concern an abstract group G_s which is also in part *subjective* – the group does not exist if no agent thinks (subjectively) to that group - like in (b1). But this group is furthermore the result of a social process of *reification:* the social meaning of the group G_s, is the join product of both subjective individuals and collective interactions, here the social communication on that socialized subjective object.

Figure 6. The social reification process, where the representation of a collective entity is both a subjective and a social entity (IMAGE A REFAIRE)



One can view the social construction of such collective entities as an emergent property of complex adaptative systems (Dessalles, Phan 2005). The underlying process is a subsumption operation which works by substituting a set of (ontologically subjective) individual entities for a social entity, the latter being both ontologically subjective and socially constructed (Figure 6).

In particular, it will be noticed that the individual equivalent requirement (3) is a *necessary* but *not sufficient* condition of this process: it is still necessary to specify the conditions under which this reification can be carried out and be maintained. In other words, what are the conditions by which individuals transform beliefs about the behavior of other individuals into beliefs about the "behavior" of a collective entity – the so-called "group" – where this group is a social construction, i.e. a subjective entity supposed to be shared by everyone.

However, this process, by which a group seen as a set of agents (G_e) and the group as a social entity (G_s) are seen as equivalent, is an important but problematical issue. According to the point of view of orthodox methodological individualism, the group exists only as a convenient representation to organize our thoughts, but it is in fact a mere set of individuals. Like the money as a "veil" in the classical economic theory, the group (G_e) is a *neutral abstraction* which can be replaced without change by the list of its members. Conversely, from a more holistic point of view, (Descombes, 1996) opposes this logical vision of an "abstract whole" with his conception of a *concrete whole*, where the whole is now a totality irreducible to the set of its elements. He argues that predicates which may be applied to the whole are not distributable over its elementary components. According to this point of view, the *relational properties* which structured the system are neither at the level of the whole nor at the level of the parts, while being constitutive of both (Descombes speaks about *structural holism*).

The *diachronic dimension* then makes it possible to reveal the particular function of the relation in the shape of existence of this collective object: "*from the historical and social point of view, the communities are precisely not collections of individuals, since their identity is not fixed by a list of individuals*" (Descombes, 1996, p.134). Commenting on Wittgenstein, Descombes shows that a "library" is more than a set of its books (logical totality), because it is *a concretely organized whole*, decomposable according to other levels of individuation that the book itself, and whose meaning is dynamically dependent on this organizational structure.

In a similar way, financial markets described by Orléan are also long lasting entities organized to coordinate exchanges. They cannot be reduced to the set of agents who make exchanges at a given moment independently of the relation they maintain with and within this structure (Orléan 2004). This *social complex structure* is thus significant by itself. The related – although evolving – meaning (as I-C collective representation) maintains itself through time independently of the agents who use this structure and exceed their individual dimension. It's perhaps precisely for this reason that agents attribute spontaneously, in everyday language, anthropomorphic qualities to collective entities in order to

make such structures more interpretable and more meaningful.

An interesting conjecture about a possible interpretation of this anthropomorphic subsumption (the "cognitive shortcut") could be derived from the *cognitivist* explanation of the collective beliefs proposed by Boudon (1995, 1997). For the latter, "Subjects try to solve complex questions by proposing reasonable conjectures. It is because these conjectures are reasonable that a great majority adopt them" (1995, translated from French p.28). Adopting a Simonian-like limited rationality position, Boudon calls "*trans-subjective reasons*" the reasons perceived as sufficient for this "large majority", to adopt these conjectures instead of adopting a more complete (and somehow more adequate but more complex) representation of the phenomenon. We will not discuss here the causes (possibly infra individual) for which this *principle of sufficient reason* is relevant for a large majority of individuals. Instead, we will concentrate on the cognitive shortcut, the object of the cognitive process itself. It is well known that Boudon considers that "ordinary" and "scientific" knowledge use similar cognitive mechanisms (correlatively, he said that it is not necessary to establish "a clear line of demarcation between the scientific and the ordinary thought" - op.cit. p. 48). According to its paradigm of methodological individualism, Boudon considers that a "collective belief is explained only insofar as it makes sense for an *idealtypic subject*, for an *ordinary* individual" (op.cit. p.38). Interestingly, we can apply these principles in a reflexive way beyond the usual usage made by this author.

We could say that a "social belief" which consists in substituting a collective of interacting agent, for an abstract entity (i.e. a group) is a similar cognitive operation as the one which consists in substituting a set of individuals for an "ideal type" (in the Weberian sense).

In other words, the subsumption of a set of agents by a representative abstract social entity can be seen as founded on *trans-subjective reason*. This results in the (false) belief that it is easier to understand a person-like behavior than the complex dynamics of a group. Thus, people tend to reduce a set of interacting agents by a simple abstract entity to which they may attribute anthropomorphic characteristics.

4. One more step: empirical evidences and formal considerations on the social dimension of cognitive reasoning in society.

We can see that there are three different ways to formulate the beliefs problem. Thus, is it possible to determine which one is the most relevant?

4.1. To distinguish between the shared beliefs and the social beliefs: salience and "focal point".

In order to determine which way is the most relevant / adequate interpretation, one can consider a more sophisticated coordination game, where it is not any more question of choosing between two undifferentiated strategies, but where it is necessary to choose a single number between 0 and 99. As in the previous game, the players' payoff is one if they coordinate on the same number and zero elsewhere. In this case there are potentially 100 Nash equilibria since all the coordinated choices are possible equilibrium candidates (the strategic matrix plays would be an identity matrix with payoff (1,1) on the first diagonal and payoff (0,0) elsewhere). However, when experimentalists test this game with "real" players, those are able to coordinate each other by "projecting" some "salient" strategies, such as choosing: 0, 1, 7, 50 or 99. The question of the existence of some salience or "focal point" in coordination problems was early exposed by Schelling (1960) and largely discussed by several authors, with a special mention for the work of Sugden and co-authors (Sugden, 1995; Metha, Starmer and Sugden, 1994a-b; Cubitt and Sugden, 2003, 2005; Bardsley, et aL. 2006). In order to understand well the cognitive mechanism which makes possible the existence of social beliefs, it is convenient, according to Orléan, to use the results of (Metha, Starmer and Sugden, 1994a-b). They achieve this experiment with two groups of reference. In the first group, they just ask to choose a natural number without reference to any situation of coordination. In the second group they propose the game of coordination (payoff 1 if there is coordination, 0 elsewhere). Their results are reproduced in table 1.

One can interpret the results of the test group "P" as a translation of the "idiosyncratic preferences" of the tested agents, which can reveal subjacent personal cultural references. That is for example the case with the "magic" number 7 which is the more frequent choice. This corresponds to the case of the "primary salience" for the authors, where each agent "gives any response that happens to come in mind at the time (Metha *et al* , p 660). In that case, the achievement of coordination in random meeting would result of accidental proximity in cultural references of the players, without any consideration of cognitive rationality as Orléan said "if the two players have shared past experiences and the same cultural background, the choice of primary salience can lead to successful coordination" (Orléan, 2004, p 203).

Table 1: the game of coordination
by (Metha and Starmer and Sugden, 1994; p. 667).

test group « P » (N = 90)		coordination game group « G » (N = 90)	
answer	proportion	answer	proportion
7	11,4 %	1	40,0 %
2	10 ,2 %	7	14 ,4 %
10	5,7 %	10	13,3 %
1	4,5 %	2	11,1 %
r = 28	c = 5,2 %	r = 17	c = 20,6 %

R = number of different answers ; C = probability that two individuals give the same answer

When coordination becomes explicitly the goal, experimented players change their behavior. Then, 40% of the players choose the value 1, which is a useful "focal point" in the sense of Shelling (1960). Let us remark that the dispersion of answers is reduced, and the frequency of the most preferred answer increases. The concentration of the players on the first two most frequent results (44,4% of the population) lead to a significant increase in the probability that two randomly drawn players give the same answer (from 5,2% to 20,6%). Finally, the spectacular inversion in the frequency of number 1 is not done to the detriment of the "magic" number 7, which is increasingly chosen by 3 % of

additional subjects within the population (frequency shift from 10,2% towards 14,4%).

If the individuals of the control group "P" choose their preferred value without any context of coordination, this is no longer the case in the situation of coordination (group "G"), To explain this behavioral change, Metha *et al.* introduce two new distinct levels of salience called "secondary salience" and "Schelling Salience". According to Orléan, it is possible to identify the fist one with the social belief of type (d1) and the second one with the social belief of type (d2). In "secondary salience" the player hypothesizes that others will choose in the same way as himself, that is, he "projects" his own individual beliefs upon the others. In this context, coordination may arise in the same way as in the case of primary salience, for instance if the players share the same cultural background (i.e. stereotypes). In such a situation, the ranking of the choice should be the same as in the case of primary salience (non coordination context). In the case of the experiment; that would have led to the choice of the value 7, which is intrinsically preferred by the greatest number of individuals. From experimental evidences, that is generally not the case. Instead, in the situation of coordination, agents try to identify the dominant belief of the group itself, independently of their own belief. That cognitive attitude corresponds to the case of social beliefs (d2) for Orlean and to the case of "Schelling Sallience" for Metha *et al*. As Orlean said : "Shelling central idea was that in order to coordinate, the players try to find a choice or principle such that, from the point of view of the *group*, it gives a unique reply and results in successful coordination. To do that, agents place themselves on a more general level of abstraction, and try to determine a principle capable of bringing out, in everyone's opinion, a unique equilibrium" (Orléan, 2004, p.204, underlined by us). Thus, the cognitive activity is turned towards the group as a separate significant entity. For Orléan, the adequate modeling of the behavior of the players corresponds to the autoreferential form (d2) of the social beliefs, in which the cognitive activity focuses on the group as an entity, not as a set of individuals: "the strength of the (d2) definition lies in the fact that it totally disregards the variability in individuals" intrinsic preferences in order to define a belief that belongs to the group as a group" (id p.205). If this *cognitive step* is problematic, it is because the dynamics of the group, as an adaptive complex system, can deviate durably from what would be a representative agent. We knew from section 3.1. and related works (Weidlich and Haag 1983; Galam, 1986; Orléan 1998a-c) that in presence of social influence, the state of the group *as a whole* could thus be durably disconnected from the state desired by the majority of its members

To summarize, in a coordination game, the *standard approach* leads to a multiplicity of Nash equilibria (from two towards infinity, depending on the considered game). From the *behavioral* point of view, the agents may to coordinate efficiently and the experimentalists reveal the selection of "Schelling's salience" (or "focal point") within the set of admissible solutions, which help the learning process and the achievement of the selection of a particular equilibrium in iterated games. The problem of the focal point is then to mobilize knowledge that is external to the "standard" framework of the game theory in order to be used as reference in the process of equilibrium selection (as the "stereotypes"). The "culture" and/or the "history" of the players can thus contribute for a significant part to the resulting social choice.

But to look at it more closely, there are the autoreferential "social beliefs" which are the cognitive supports. They are to play a decisive role. In particular, in the absence of a preexistent "cultural" stereotype, the autoreferential form (4) of the social beliefs is adapted to facilitate emergence by a "historical" process such as "focal points". We now will try to propose a cognitive explanation to this focusing on a collective object: "beliefs of the group", rather than the construction of a recurring chain of references crossed on the whole of the agents, according to the approach of the "common belief".

4.2. The « beauty contest »: expectations about the others and *cognitive hierarchy* in a competitive context.

A famous example of recursive reasoning involving hierarchical levels in cognition is the "beauty contest" described by (Keynes, 1936). The formalization of this story in a non cooperative game framework taken off in the Eighties (Mill, 1986, Simonsen, 1988). The first experimentation published based on this game is due to Nagel (1995). The author quotes his sources of inspiration (in particular Roger Gesnerie) in introduction of the article, which is clearly written within a post rational-expectations framework (cf Orléan 2002, for the link with the rational expectations). This game is a very stylized form of the topic evoked by Keynes. It evacuates in particular the "cultural" dimension underlined for the coordination on some "focal point" in the previous section. These simplifying assumptions allow us to focus on the auto-referential mechanisms of such a process. This one received a first interpretation in terms of "cognitive hierarchy" (Stahl and Wilson 1995), and (Camerer 2003, Camerer and Ho and Chong 2004), in the field of the experimental games (cf. also Phan, 2004a, for an application to the MAS).

In the "beauty contest" of Rosemarie Nagel, a group of players, who cannot communicate with each other, must choose a real number between 0 and 100. The winner of the game is the one who will choose the value nearest to the average of the choices, multiplied by a real number p such as $p < 0 < 1$. If several players choose this value, they share the payoff. Let us take for simplicity $p = 1/2$ and payoff = 900.

According to Camerer (2003) one can classify the players according to a *cognitive hierarchy* related to the "strategic depth" of their reasoning, ie the degree to which they can reason about other players. At the "level zero" players play without particular reasoning their idiosyncratically "preferred" value. This corresponds to what Orlean denotes as "primary beliefs"(in the Sugden's sense of "primary salience"). At the "level one" players assume implicitly or explicitly that other players are of level zero and try to expect their default choice. For instance they can try to play some representative "level zero" value multiplied by p, that is lower than $100*p = 50$. If the player is aware of the fact that the others can make like him, then are of the level zero or one, he should play a mix of level zero and level one representative value multiplied by p, (ie. lower than $50*p = 25$ if all co-players are assumed to be of level one). But we are already at "level two" in the hierarchy of reasoning.

Let us come back to the standard game theory and consider now the point of view of an automat which would have been programmed to apply iteratively the principle of the dominated strategies elimination. The latter will apply this recurrent calculus until the variation of the solution is smaller than some threshold (fixed a priori to avoid a fall without end in an infinite recurrence). This series converges towards a unique Nash equilibrium of this game, equal to zero. Let us note that zero is also the unique solution of the linear recurrent equation of the order one: $X(t) = p.X.(t-1)$.

Thus, according to the standard approach, there is a unique Nash equilibrium equal to zero. Experimental results in behavioral economics show that this equilibrium is never reached within a one shot game. However, the players can approach this value by learning in an iterated game. The simplest explanation is that the players do not apply the iterated dominance method, and do not have deep reflexive calculus. A second level explanation suggests that players can anticipate that other players will not go down very far in strategic deepness. To the third level, the problem is now reduced to form beliefs on the average strategic depth of the other players. But one could then say at a fourth level (which it is enough to everyone) that the others believe that the level of strategic depth is so high that this forecast is carried out (it is the phenomenon of Self-Fulfilling Prophecies, well known by the sociologists at least since Merton (1948), but principally developed by the economists within the theory of rational expectations (cf. i.e. Azariadis, 1981; cf. also: Snyder, 1984, for social psychology). What explanation is proposed by the economists?

Table 2 - cognitive hierarchy and strategic deep in the beauty contest

	Level 0	Level 1	Level 2	Level 3
Nagel (1995)	24%	30%	41%	6%
Camerer (2003)	16%	38%	47%	0%

Source : Camerer (2003) p.211 – Because Camerer's rounding, the sum of the two lines is 101%

From the point of view of cognitive hierarchy, interpretations from the controlled experiments by the behavioral economists argue that initially a little more than 50% of the subjects do not go below the level 1 of the previously definite cognitive hierarchy. The choices of the great majority of the population exhibit an apparent cognitive depth of 1 or 2. Returning to our previous example of $p = 0,5$, the distribution of the results obtained by Nagel and Camerer is reproduced on table 3. In the experiments of Camerer, the modal value is established to 25. At Nagel, the mode corresponds to the choice of those who are at level two, that is to say the class [8-10]. In the cognitive hierarchy approach of Camerer, the "group" makes sense only as a *set of agents*, in which the "deepest" cognitive agents try to expect the collective outcome. The relevant 4Quadrant framework is the one of Figure 3, and the cognitive hierarchy can be viewed as a sophisticated mechanism of cognitive rationality, taking into account an heterogeneous distribution of bounded capacity, within the methodological individualist paradigm.

Here the situation, although very near (a parameter $p < 1$ instead of $p = 1$) presents immediately a collective dimension insofar as coordination should not be done with another individual choice, but with a collective object included in the rules of the game. On the other hand, even if a significant share of the players have a cognitive hierarchy limited, the most "refined" players who will make reasoning of level 3 and will form beliefs on the average strategic depth of the other players, cannot make the dead end on the formation of a belief on the share of players in the population which will also go down up to this level of strategic depth, which introduces reflexivity into the process of formation of the beliefs. It is thus necessary to be "a level below" the mean level, but to keep outward journey below. The results mentioned by Camerer (2003) in connection with the pupils of the University of Caltech and other particular groups confirm this last point: the average of the results obtained by these groups is "a half level" with the lower part of that in general obtained (as in table 3, where the totality of the population is on a hierarchical level lower than 3), which shows that a significant share of these groups has certainly gone to "level 3".

4.3 Cognitive hierarchy versus team reasoning

In the problem of pure coordination of the section (4.1.) co-players must coordinate with each other in bilateral meetings. In such a game, contrary to the beauty contest, both players have *common interest* in coordination. In such a case, cognitive hierarchy is not the only way for learning the coordination.

An alternative explanation of the Schelling salience in terms of "team reasoning" has been recently introduced (Sugden, 1993, 1995, 2003; Bacharach, 1999, 2001, Bardsley *et al.*; 2006). This approach involves non-standard forms of rational reasoning in the way of "collective rationality". According to this approach: "An individual i reasons with respect to a group G if he/she works out which profile of actions for members of G would give the best results for G, and then chooses the component of that profile". Roughly, the individual asks 'what should we do'?, and acts upon the answer in the expectation that members of the group think and behave analogously" (Bardsley *et al.*, 2006, p.9).

(Bardley *et al*., 2006) reports experiment about how players try to identify a focal point to select an equilibria in one-shot coordination game. These authors experiments two classes of coordination games (1) text task (ie. for example one has to choose a string between the following: {Ford, Ferrari, Jaguar, Porsche} - pure coordination case) (2) number task (ie for example one has to choose a number between the following: {10,10}- nondescript pure coordination case - or {10,10, 10, 9} or {10,10, 10, 1} – both later – nondescript Hi-Lo case, etc...).

Two experiments are realized in Amsterdam and Nottingham, which consist in a set of tasks, shared in a common basic design, but supporting different presentation and organization of tasks. The experiment proceeds with three treatments. In the picking treatment, the subject has to choose an object and earns the number of points shown on that object (in text task, each choice carries the same number of points). In the guessing treatment, the subject is randomly

paired with another one and must guess what his partner did during the picking treatment. The goal of those two treatments whitout a group is to identify the so-called primary (for picking) and secondary (for guessing) salience (Metha and Starmer and Sugden, 1994), as introduced in section 4.1. The third treatment deals with group coordination as the two subjects, randomly and anonymously paired must play the game simultaneously what is assumed equivalent to play against all the population. If the two partners choose the same object, both win the associated number, otherwise they lose and the payoff is zero.

ICI EXPLIQUER LES HYPOTHESES TESTEES ET LE RESULTAT DES EXPERIMENTS (tableau)

Cognitive hierarchy gives a primary role to pre-reflexive preferences, which is primary salience. On the contrary, according to the results of the 1994 experiment by Metha *et al*., team reasoning gives no role to primary salience.

The main issue of the two experiments is both cognitive hierarchy and team reasoning failed at least one time. Within each experiment, a large majority of subjects used a common mode of reasoning to identify focal point. But that common mode of reasoning differs in the two experiments. The authors underline the sensitivity in the decision context: "the implication is that our subjects were able to use subtle feature of the experimental environment to solve the problem of coordinating *on a common mode of reasoning*." (Brardsley et al. 2006, p. 34). According to Schelling and to various sociological theories (may be except those of Coleman) there is a diversity of methods according to which people achieve coordination, and we must reject any monist explanation through a single form of rationality, or a single formal theory.

Conclusion: challenges for the ontology of cognitive agents in artificial society

The introduction of beliefs assigned to the agents in situation of strategic interactions can lead to recursive reasoning very far away from the effective practices of coordination, as the confrontation of the results of the standard game theory with those of the experimental economy shows it. We showed that the use of some "*cognitive short cuts*" such as the use in the reasoning of social entities could help to simplify certain problems of coordination. In the cases under review, it is precisely because the standard game theory does not use this type of entities which one can lead to results seemingly surprising from a pragmatic point of view, like the multiplicity of equilibrium, and/or of the infinite recursive reasoning.

Linked with both strong methodological individual approach and strictly selfish individually-based instrumental rationality, the paradoxical multiplicity of equilibria, in the coordination games can be reduced by relaxing at least one of these methodological principles. The first way is to relax the scope of rationality, but to maintain the strong methodological individual approach. This is the way of population game theory (the evolutionary turn in game theory). By relaxing the common knowledge hypothesis and by introducing random noise in the agent's choices, such an approach achieves the equilibrium selection problem (Blume 1997, Young 1998). According to (Sugden, 2001) this does not change the superstructure of the theory, namely, the ontological commitment on the sole individuals, stated in figure 3. This is the same thing for the cognitive hierarchy framework. This approach introduces no collective entity, just relaxes the cognitive capacity of the agents by introducing heterogeneous capacity in the depth of reflexive reasoning.

The other way is to maintain the rationality hypothesis, but to relax the strictly individualistic goal of the agents by introducing a collective entity such as the "group" within rational calculus. This does not necessarily lead to a dramatic rupture with standard approach. That is clear in the case of the so-called "we intention" (Tuomela 1990), as Tuomela, claims his methodological individualism. But this is a weaker form of it, which may include larger ontological commitment. This is the case with *team rationality* advocated by Sugden and Bacharach: "Team reasoning forms a bridge between individual and group rationality. It happens in individual heads, but it harnesses the group's causal powers. It makes of the N individual agents a single superagency". (Bacharach 2001). In that case as Bacharach said, group members do not wonder "What should I do" but "What should We do". As we underlined it, in a framework of complex adaptive systems, due to social interactions, this may have some consequences.

As an example: the use of *cognitive short cuts* for subsuming the diversity in a set of individuals in interactions under a generic single "social entity" which then makes it possible to limit the strategic depth of the reasoning, the complexity of the problem and/or the quantity of information to be treated by the agents. But this subsumption could also give the illusion of supporting the comprehension of the agent while replacing a whole complex system by a single entity with anthropomorphic attributes. Consequently, the subsumption supports coordination by a simplification of the selection processes, but not without consequences, on the issue of the corresponding processes. Indeed, the dynamic auto-referential process associated with the social entities thus introduced into the judgement can lead to collective drifts or stabilize the system in sub-optimal situations, in particular when the collective dynamics produce complex effects which make the convenient fiction of the representative entity non-operative or even fallacious.

In the two considered classes of games, knowledge that the players have of the "cultural background" of the other players can be seen as a reducer of uncertainty, which helps (generally) coordination, through the support of several levels of "collective beliefs" which facilitate the elimination of the solutions that seem then little (or not) relevant and the possible convergence towards a focal point. The question of the emergence of such beliefs remains however posed. The appearance of similar behaviors within a population of individuals having an indicator (tag) enabling them to be mutually identified may be at the origin of the formation and of "collective beliefs" of this type. The study of the nature of the emergence of a behavior shared by sets of agents equipped with the distinctive external signs and the conditions of formation of belief associated with this emergent phenomenon are treated in complementary work which relates to the cognitive executives of the observation and the detection of the emergent phenomena (Phan 2004b,

Dessalles, Phan, 2005). More specifically, it is a question of identifying the conditions according to which an approach in terms of "shared individual beliefs" could lead by "emergence" (and subsumption) to an approach in terms of "social beliefs". One shows that the emergence of "social beliefs" in the previously defined direction can be modeled thanks to the conceptualization/formalization of another form of "cognitive hierarchy". designed to formalize at the same time phenomena "of strong emergence" (within the meaning of Muller, 2002) and of the processes of observation and conceptualization of these phenomena per "reduction of complexity" (Dessalles, Bonabeau, 1997).

References

[1] Anderson P.W., Arrow K.J., Pines D. eds. (1988) *The economy as an evolving complex system*, , Santa Fe Institute, Studies on the Sciences of Complexity, Addison-Wesley Pub.Co, Reading Ma.

[2] Arthur W.B., Durlauf S.N., Lane D.A. eds. (1997) *The Economy as an Evolving Complex System II*, Santa Fe Institute, Studies on the Sciences of Complexity, Addison-Wesley Pub.Co, Reading Ma .

[3] Aumann R.. (1976) "Agreeing to Disagree", *Annals of Statistics* 4, p.1236-1239.

[4] Azariadis C. (1981), "Self-Fulfilling Prophecies", *Journal of Economic Theory,* 25, p. 380-396.

[5] Bacharach, Michael (1999) "Interactive Team Reasoning: A Contribution to the Theory of Co-operation", *Research in Economics*, 53(2), June, p. 117-147

[6] Bacharach, M. (2001) "Superagency: Beyond an Individualistic Theory of Games", in van Benthem, Johan (ed.) *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge* (TARK-2001) San Francisco, CA: Morgan Kaufmann, p. 333-337.

[7] Bardsley N., Mehta J., Starmer C. Sugden R. (2006) "The Nature of Salience Revisited: Cognitive Hierarchy Theory versus Team Reasoning", September.

[8] Blume L.E. (1997) ; "Population Games", in Arthur, Durlauf, Lane eds., *The Economy as an Evolving Complex System II* ; Santa Fee Institute, Studies on the Sciences of Complexity, Addison-Wesley Pub.Co, Reading Ma , p. 425-460.

[9] Bonabeau E., Dessalles J-L. (1997) "Detection and Emergence", *Intellectica*, 25 p.89-94.

[10] Boudon R., (2006) "Are we doomed to see the homo sociologicus as a rational or as an irrational idiot?", in Elster *et al.*, *Understanding Choice, Explaining Behaviour, Essays in the Honour of O.-J. Skog*, Oslo, Unipub Forlag, Oslo Academic Press, pp. 25-42

[11] Boudon R. (2003) *Raison, bonnes raisons*, Col. « Philosopher en sciences sociales », Presse Universitaires de France, Paris.

[12] Boudon R. (1998) "Social mechanisms without black boxes", in Hedström and Swedberg eds., *Social mechanisms: an analytical approach to social theory*, Cambridge, Cambridge University Press, p.172-203

[13] Bourgine P., Nadal J.P. eds. (2004) *Cognitive Economics,* Springer Verlag

[14] Bouveresse J. *Le Mythe de l'intériorité, expérience, signification et langage privé chez Wittgenstein*, Paris, Les éditions de Minuit, 1976, réédition 1987.

[15] Camerer C.F. (2003) *Behavioral Game Theory: Experiments on Strategic Interactions*, Princeton University Press.

[16] Camerer C.F., Ho T. Chong K. (2004). A cognitive hierarchy model of games." *Quarterly Journal of Economics* 119: 861-898

[17] Castelfranchi, C., (1998). Modelling social action for AI agents. *Artificial Intelligence*, 103 (1-2), 157-182.

[18] Cherkaoui M. (2005) Invisible Codes, GEMAS Studies in Social Analysis, The Bardwell Press, Oxford.

[19] Cubitt R., Sugden R..(2003) "Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory", *Economics and Philosophy* 19; p.175-210.

[20] [13] Cubitt R., Sugden R..(2005) "Common Reasoning in Games: A Resolution of the Paradoxes of 'Common Knowledge of Rationality'", September..

[21] Descombes V. (1996) *Les institutions du sens*, Les éditions de minuit, Paris.

[22] Dessalles J.L., Ferber J., Phan D (2007) "Emergence in multi-agent systems: review of conceptual and methodological issues for the application to the social sciences" *in preparation*

[23] Dessalles J.L., Phan D (2005) "Emergence in multi-agent systems: cognitive hierarchy, detection, and complexity reduction part I: methodological issues" in Mathieu, Beaufils, Brandouy (eds.), Agent-Based Methods in Finance, Game Theory and their Applications, Series: Lecture Notes in Economics and Mathematical Systems, Vol. 564 Springer, p. 147-161

[24] Douglas M. (1986) *How Institutions Think*, Syracuse University Press,

[25] Durkeim E. (1895) *Les règles de la méthode sociologique,* trad. by W.D. Halls. *The rules of the sociological method.* New York: The Free Press, 1982

[26] Ferber J. (1999) *An Introduction to Distributed Artificial Intelligence.*, Addison Wesley, London,

[27] Gilbert, M.. "Modelling Collective Belief", *Synthese 73*, 1987, p. 185-204

[28] Gilbert M. *On Social Facts*, Princeton University Press, 1989

[29] Gilbert M. (1990) "Rationality, coordination and conventions", *Synthese* 84, p.1-21.

[30] Gilbert M. (1996) *Living together : Rationality, Sociality and Obligation*, Lanham MD, Rowman & Littlechild.

[31] Gilbert M. (1997) "Concerning sociality: the plural subject as paradigm" in J.D. Greenwood ed. The Mark of the Social Discovery or Invention ? London, Rowman & Littlefield Pub Inc. p. 17-36.

[32] Gilbert M. (2000) *Sociality and Responsibility : New Essays in Plural Subject Theory*, Lanham MD, Rowman & Littlechild.

[33] Gilbert M. (2002) "Belief and acceptance as features of groups" Protosociology vol 16 p.35-69.

[34] Gutknecht, O. Ferber, J. Madkit: a Generic Multi-Agent Platform, Autonomous Agents (AGENTS 2000), Barcelona, ACM Press, 2000, p. 78-79.

[35] Hedström P, Swedberg R. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press, 1998.

[36] Hume D. (1740) *A Treatise of Human Nature*, ed. L. A. Selby-Bigge. rev. 2nd. ed., P. H. Nidditch. Clarendon Press, Oxford, 1976.

[37] Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*, London: Macmillan.
[38] Kincaid H. "Reduction, Explanation, and Individualism", *Philosophy of Science*, 53(4) December, 1986, p. 492-513
[39] Lesourne J., Orléan A., Walliser B. dir. (2002) *Leçons de microéconomie évolutionniste*, Odile Jacob, Paris.
[40] Lewis D.K. (1969) ; *Convention, a philosophical Study*, Harvard University Press Cambridge Mass
[41] Livet P., Ogien R. eds. (2000) *L'enquête ontologique, du mode d'existence des objets sociaux*, Editions de l'EHESS, Paris.
[42] Mehta J., Starmer C., Sugden R. (1994a) "Focal Point in Pure Coordination Games: An Experimental Investigation", *Theory and Decision*, 36 p.163-185.
[43] Mehta J., Starmer C., Sugden R. (1994b) "The nature of Salience: An Experimental investigation of Pure Coordination Games" *American Economic Review*, 84 , p. 658-673.
[44] Merton R. K. (1948) "The self-fulfilling prophecy", *Antioch Review*, 8, p. 193-210..
[45] Michel F., Daniel G., Phan D., Ferber J. "Integration Moduleco-Madkit: first results, SMAGET 05, Joint Conference on Multi-Agent Modelling for Environmental Management, March, 21- 25, 2005.
[46] Moulin H. (1986) *Game theory for social science*, New York press NY.
[47] Müller JP. (2004) "Emergence of Collective Behaviour and Problem Solving", in Engineering Societies in the Agents World IV, Fourth International Workshop ESAW-2003, Revised Selected and Invited Papers, pages 1-20, LNAI 3071 Springer Verlag 2004
[48] Nagel R. (1995) "Unraveling in Guessing Games: An experimental Study", *American Economic Review*, 85, p. 1313-1326.
[49] Namatame A. (2006) *Adaptation and Evolution in Collective Systems*, Advances in Natural Computation vol 5, Singapore, World Scientific Pub. Co.
[50] Orléan A. (1998a) "Informational Influences and the Ambivalence of Imitation", in Lesourne, Orléan (eds.), *Advances in Self-Organization and Evolutionary Economics*, p. 39-56.
[51] Orléan A. (1998b) "The Ambivalent Role of Imitation in Decentralised Collective Learning", in Lazaric, Lorenz (eds.), *Trust and Economic Learning*, Elgar Publishers, p. 124-140.
[52] Orléan A. (1998c) "The Evolution of Imitation", in Cohendet, Llerena, Stahn, Umbhauer (eds.) *The Economics of Networks. Interaction and Behaviours*, Springer-Verlag, p. 325-339.
[53] Orléan A. (2002) « Le tournant cognitif en économie », *Revue d'Economie Politique* N°5 septembre octobre, p.717-738.
[54] Orléan A. (2004) "What is a Collective Belief" in Bourgine, Nadal, op.cit., p.199-212.
[55] Pettit P. (1993) *The Common Mind, An essay on Psychology, Society, and Politics* Oxford University Press, 2° ed 1996.
[56] Pettit P. (2000) "Defining and defending social holism", *Philosophical Explorations*, 1 (3) p. 169-184.
[57] Pettit P. (2004) *Penser en société, Essais de métaphysique sociale et de méthodologie*, Col. « penser en sciences sociales », Presse Universitaires de France, Paris.
[58] Phan D. (2004) "From Agent-Based Computational Economics towards Cognitive Economics" in Bourgine, Nadal op.cit., p. 371-398.
[59] Phan D. (2004a) "Hierarchy of cognitive interactive agents and statistical mechanics : how Object Oriented Programming highlights the connection", in Coelho, Espinasse. Seidel .eds., *5th Workshop on Agent Based Simulation*, Lisbon, Portugal, May 2004 SCS Pub. House, Erlangen, San Diego, p. 69-76.
[60] Phan D., Gordon M.B, Nadal J.P.. (2004) « Social Interactions in Economic Theory: an Insight from Statistical Mechanics » in Bourgine, Nadal op cit. p. 225-358.
[61] Schelling T.S. (1960) *The strategy of Conflict*, Harvard University Press, Cambridge, Ma.
[62] Schuster H.G. (2001) *Complex Adaptive Systems, An Introduction*, Scator Verlag, Saarbrücken.
[63] Searle J.R. (1995) *The Construction of Social Reality*, Free Press.
[64] Simonsen M.H. (1988) "Rational Expectations, Game Theory and Inflationary Inertia" in Anderson, Arrow , Pines eds. *The economy as an evolving complex system*, Santa Fe series Vol5 Addison-Wesley CA., p. 205-241
[65] Snyder, M. (1984) "When belief becomes reality", in Berkowitz ed., *Advances in experimental social psychology*, 18, Orlando, FL : Academic Press, p. 247-305.
[66] Stahl, D.O. Wilson P. "On players' models of other players", *Games and Economic Behavior* 10, 1995 p. 218–254.
[67] Sugden, R. (1993) "Thinking as a team: toward an explanation of nonselfish behavior", *Social Philosophy and Policy*, vol 10, pp. 69-89.
[68] Sugden, R. (1995) "A theory of focal points", *Economic Journal* 105 p. 1269-302.
[69] Sugden, R. (2003). "The logic of team reasoning". *Philosophical Explorations* 6: 165-181.
[70] Tuomela, R., "Can collectives have beliefs?" in Haaparanta and Kusch and Niiniluoto (eds.) "Language, Knowledge, and Intentionality: Perspectives on the Philosophy of Jaakko Hintikka", *Acta Philosophica Fennica* 48, 1990
[71] Tuomela R. "Methodological Individualism and Explanation, *Philosophy of Science*, 57, 1990, p 133-140.
[72] Tuomela R. "Group Belief", *Synthese*, Volume 91(3) June 1992, p.285-318
[73] Walliser B. (2004a) "Principles of Game Theory", in Bourgine Nadal eds. op. cit. p.55-78.
[74] Walliser B. (2004b) "Topics in Cognitive Economics", in Bourgine Nadal eds. Op. cit. p.183-198.
[75] Weisbuch G. (1990) *Complex Systems Dynamics*, Santa Fe Institute Studies in the sciences of complexity, Redwood City CA ., Addison Wesley.
[76] Wilber K.(2000) *A Theory of Everything: An Integral Vision for Business, Politics, Science, and Spirituality,* Boston, MA Shambhala Publications, Inc.
[77] Wittgenstein L. *Philosophical Investigations*, Oxford: Basil Blackwell

Author Bios

Denis Phan is Post & Telecommunications administrator and Senior Researcher in Economics and Social Sciences at

the CNRS. His past research interests are in the economics of telecommunications, the software industry and regulation. He has published papers and various books in this field. Since 2000 his main research interests have been cognitive and agent based computational economics. He is the co-author with A. Beugnard of the Moduleco simulation Framework, now integrated into the Madkit Platform as a plug-in. He have recently organized two researcher schools on complex systems and MAS, and published an introductory book on MAS modeling and simulation in Human and Social Sciences in French to be traduced in English.

Jacques Ferber Is Professor of Computer Sciences at Montpellier. His domain of research is about multiagent systems (MAS), distributed systems, multiagent simulation, component based and agent based software engineering. He started to work in MAS in 1987 and is considered as a pioneer of this field in France. His introductory book on MAS and Distributed Artificial Intelligence has been published in French, English and German. Member of several editorial boards and program committee (onferences and workshops) he is the co-author with Olivier Gutknecht and Fabien Michel of the MadKit multiagent platform