DRAFT MAY 2004

Metareflexive Mimetism: The prisoner free of the dilemma

David Chavalarias[1],
Center for Research in Applied Epistemology (CREA),
Ecole Polytechnique, Paris, France,
www.polytechnique.edu

Abstract :

Much attention has been given in the last several years to imitation processes for the modeling of social systems in economy as well as in anthropology, sociology and political science. But the diversity of mimetic rules employed by modelers proves that the introduction of mimetic processes into formal models cannot avoid the traditional problem of endogenization of all the choices, including the one of the mimetic rules. This article addresses this question starting from the remark that human's reflexive capacities are the ground for a new class of mimetic rules. This leads us to propose a formal framework, metamimetic games, which advantage is to endogenize mimetic processes while being human specific. A computational study of a metamimetic game around the spatial prisonner's dilemma is given as a first insight into metamimetic dynamics.

[1] chavalar@poly.polytechnique.fr; home page : http://chavalarias.free.fr

I. What do human's meta-cognitive capacities bring about in mimetic processes?

a. Mounting the evolutionary hierarchy

Mimetism is considered as a key component in human social behavior (Girard 1961), and the sophistication of human mimetic processes could have been one of the major evolutionary transition in hominization toward human's social organization, as we know it (Donald 1993). For these reasons, scientists more and more incorporate mimetic processes into formal modeling to account for the extremely rich structures observed in human's social systems.

But the diversity of mimetic rules employed by modelers proves that the introduction of mimetic processes into formal models cannot avoid the traditional problem of endogenization of all the choices, including the one of the mimetic rules. In the literature of social systems modeling, two main processes of imitation have been defined. (1): In the traditional conception of *Homo oeconomicus*, some researchers considered *payoffs-biased imitation,* i.e. imitation of the most successful agents in one's neighborhood (Nowak & May 1992). (2): A growing number of contributions are attempts to introduce what is called *conformism,* in the study of social phenomena (Axelrod 1997 ; Bala & Goyal 2001 ; Galam 1998 ; Orléan 1998). Here, *conformism* is the propensity of individuals to adopt some behavior when it has already been adopted by some of their neighbors, the propensity being relative to the frequency of that behavior in the neighborhood. To a lesser extent, other imitation processes have been studied, among which we can mention (3): *non-conformist*, the propensity of an individual to adopt the behavior of the minority (Arthur 1994), or prestige (Henrich & Gil-White 2001). This list of imitation processes is far from exhaustive and we can already notice that even for *conformism* or *payoffs-biased imitation*, several technical definitions have been proposed, either deterministic or probabilistic (Nowak et al. 1994). On the other hand, it is also possible to propose models including several rules for imitation, as some authors already did (Boyd and Richerson 1985; Henrich and Boyd 1998 ; Janssen & Jager 1999, Kaniovski et al. 2000).

This raises an epistemological question for modelers. Which rule(s) for imitation should be considered depending on the social systems under study? Some scholars have addressed this question in an evolutionary perspective, assuming that the mimetic rules used were the result of natural selection processes (Henrich & Boyd 1998). But the slow dynamics of genetic processes seems to be in contradiction with the quick evolutions observed in social systems (Alvard 2003, Feldman & Laland 1996, Gould 1987, Gintis 2003) that is mostly grounded on cultural evolution through cumulative learning. A more realistic view would be that the set of mimetic rules itself, depends on the culture under study, its history, and quick varying environmental conditions. If we can imagine that mimetic dynamics have been designed by genetic evolution, it is harder to believe that genetic evolution itself it directly responsible for changes in rules for imitation. The problem is here to find an appropriate top-level evolutionary process that could select mimetic rules while being compatible with observation in cultural evolution.

b. Cognitive foundations

Another way to address the question of endogenization of mimetic rules will perhaps come from a recent concern in social system modeling. The complexity of human’s social systems have no equivalent in others species. For example, considering group coordination, only insect’s societies, composed of very simple entities, have social structures involving several thousands members. This feature disappears as soon as the repertoire of behavioral possibilities of species get wider, and reappears only when it comes to humans (Bourgine 2003, Wilson 1975). This remark is noticeable because it is precisely modeling of self-organized systems in ethology that has been a precursor for multi-agents modeling in social sciences. It is clear that the goal for social systems modeling is not to consider humans as cloned insects. What is at stake is rather to find differences between humans and others mammals, which enable emergence of highly structured social groups, while keeping inter-individual heterogeneity. This has lead recently some modelers to propose, as an heuristic in social modeling, to consider in priority models that could be human specific (Alvard 2003, Bowles & Gintis 2003, Fehr

& Fischbacher 2003). In social sciences, a similar heuristic that particularly concerns mimetism, has been formulated few decades ago by René Girard[2] (1978):

> *In order to elaborate a science of human, we have to compare human imitation with animal mimetism, precise human's specific modalities of mimetic behaviors if they exist.*

Following this heuristic, we will thus look for differences between animal's and human's cognitive capacities that could have qualitative impacts on imitation processes. From numerous studies in psychology as well as in ethology, we can see that two elements are playing a crucial role in human behavior while being apparently out of reach of non-human cognition.

First, humans are reflexive beings. To give a low level definition of reflexivity, it is the ability to take as object of cognitive treatment the cognitive treatments themselves by creating new levels of cognitive processing. Emergence of reflexive capacities can be traced in ontogeny with the study of the development of infant's cognitive capacities (Zelazo et al. 1996) and the self-triggered loop that should be the elementary component of reflexive processes is closely linked with the constitution of the self (Damasio 1999, Donald 1991). Reflexivity helps us to think others as we think ourselves and ourselves from other's eye view and thus develop our social skill. From the imitation point of view, reflexivity makes all the difference since, as Eric Gans (1995) says, *"prehuman imitation is non-reflexive; the subject has no knowledge of itself as a self imitating another"*.

The second difference between animals and human's cognitive capacities, closely related to reflexivity, is metacognition (Donald 1991, Sperber 2000, Tomasello 2000), defined here as cognition about cognition. Whether animals have metacognitive capacities is still in debate in this scientific community. Some experiments seem to indicate that great apes and dolphins may have some rudimentary metacognitive capacities (Smith et al. 2002, Rendell and Whitehead 2001), but those are very limited. In particular, there is no evidence that animals can consider learning or imitation processes as object of cognition, and to our knowledge, there is no experiment showing that

---

[2] « *Pour élaborer une science de l'homme, il faut comparer l'imitation humaine avec le mimétisme animal, préciser les modalités proprement humaines des comportements mimétiques si elles existent* »

animal's could add voluntarily a metacognitive level to solve a given problem, although some primates seem to be able to deal with chains of hierarchically organized behaviors (Byrne 1998). This means that animal's metacognition, if it exists, is most probably constituted of rigid chains of process monitoring that can as well be hardwired, without requiring reflexivity to monitor their structure.

There is no space here to give more details about these two differences. But we will try to show that taking them into account makes it possible to build a new class of models that may offer an answer to the problem of the multiplicity of mimetic rules.

II Reflexive mimetic rules and endogenization of meta-choices

Introducing metacognition and reflexivity in formal models reveals two phenomena. First, imitation rules can be identified as cognitive objects, modifiable by way of cognitive treatments like imitation processes. Second, an imitation rule can be reflexive in the sense that it can participate to its own modification. To go further in that direction, we have to be more precise on what we will consider to be an imitation rule. We will give here a definition that fits a multi-agents perspective. Before that, we will expose briefly the prisoners dilemma game that will be our standard example when we will need to fix ideas with a concrete case.

a. A brief description of the prisonner's dilemma game

The prisonner's dilemma game is a two-players game where players have to choose simultaneously one of the two options: to defect (*D*), or to cooperate (*C*). The dilemma lays in the fact that option *D* leads always to the highest reward whatever the other does - rewards associated with playing *D* are *T* (temptation to defect) when playing against a cooperator (who wins *S* with $T>S$), and *P* when playing against a defector (who also receives *P*). But when both players cooperate, they receive $R>P$. This means that mutual cooperation is more advantageous than mutual defection (collective rationality), but given the opponent's action, defection is individually more advantageous than cooperation (individual rationality). The situation is usually synthesised by the following table:

player *B*→ / ↓ player *A*	**B plays C**	***B* plays *D***
A* plays *C	*A* & *B* win *R*	*A* wins *S* & *B* wins *T*
A* plays *D	*A* wins *T* & *B* wins *S*	*A* & *B* win *P*

With the following relations on *T, R, P* and *S*:

$T > R > P > S$, and

$T + S < 2R$ (mutual cooperation is the best you can do collectively).

The consequences of a prisonner's dilemma situation is that if a player is rational and greedy, the best he can do is to play *D,* because whatever his opponent, the payoffs associated with action *D* will always be higher than the those associated with action *C* ($T > R$ and $P > S$). Consequently, if both players are rational and greedy, they will both defect and loose the advantage of mutual cooperation ($P<R$). Thus in this game individual rationality is in conflict with collective rationality. It should be noticed however that agents can have various rules for deciding which behavior to adopt that are not necessarily payoffs maximizing. They can for example be altruists, willing to maximize other's payoffs, or conformist, willing to do as the other do. This point will be detailed below.

In the following, we will consider an *N*-players iterated prisonner's dilemma, which means that players are playing several rounds in a row, with several players each round. We will refer to this game as *game G.* At a given round, the behavior of a player will be the same with all its opponents[3] (*C* or *D*), which means that the current state of an agent can be described by a set of parameters *(b,r,g)* where *b* is its current behavior (or last action *C or D*), *r* the rule it used to choose this action (payoffs maximizing, altruism, etc.) and *g* the payoffs it obtained playing *b* last round.

b. Artificial agents

We will know give some definition that will be useful in the following. First of all, we will consider agents that are defined by a collection of traits. These traits will be categorized in *modifiable traits* and *other traits*.

- *Modifiable traits* are those that an agent can change voluntarily. This is the case for example, for a cooperative vs. defective behavior, the colors of clothes, the

[3] This is a classical situation in Public Good Games or Common Pool Resources.

political party the agent decides to vote for, the strategy it chooses, or the learning rules adopted for a given task, etc. Most of the time, these kinds of changes take place on small time scales (within a day). The set of modifiable traits of a given agent *A* can be represented by an ordered set $s_A=(\tau_1,\dots,\tau_m)\in T_{\text{mod}}$ of *A's* choices for the different categories of modifiable traits.

We will note $K(T_{\text{mod}})$ the set of modifiable trait categories. For example, in game *G,* $K(T_{\text{mod}})=\{$'behavior', 'rule for imitation'$\}$ with *behavior={C,D}* and *rule for imitation={copy the best action of previous round, do what the other did last round, etc.}*.

- *Other traits* are those that do not entirely depend on agent's will. They depend on global dynamics and change generally on larger time scales (months, years, life), like social positions, payoffs, reputation, prestige, age, etc. For example, in game *G,* payoffs (*g)* are the only category *other trait.*

A state of the world ω is the list of all the traits that characterize the agents and their environment. The set of all possible worlds will be noted Ω. Agents will represented by ordered sets. The first components will represent their modifiable traits, that taken as a whole will be called their strategies. The last components will represent the other traits (their payoffs, prestige, etc.). As we saw it, in game *G*, an agent can be represented by an ordered set *(b,r,g).* Then in game *G,* a state of the world is the list of all ordered sets defining the agents, and particular strategy is a couple *(b,r).*

Agents become aware of perceived traits through the social network they are embedded in. For a given agent *A*, the set of all agents from which it can learn some traits, by any mean, will be called its neighborhood: V_A. Agents categorize their neighborhood in sub-neighborhoods on the basis of the perceived traits.

For example, Cavalli-Sforza and Feldman (1981), Boyd and Richerson (1985) consider different types of cultural transmission processes within sub-neighborhoods indexed on age and kinship: vertical transmission from parents to offspring, oblique from elders to younger, and horizontal among peers.

V_A can also contain unidirectional links, like those that are defined by the media network that is known to be an important factor of social influence (Bandura 1977)

Given all this, we are now able to define what we will consider as rules for imitation:

<u>Definition</u>: *Rule for imitation*

A rule for imitation is a voluntary cognitive process that given an agent, a state of the world, and a category of modifiable traits, produces as an output, the set of agent's neighbors that will be taken as reference for the modification of the modifiable trait of the given category, as well as the process for this modification.

Consequently, the imitation rule of an agent A can be represented by a function $r{:}\Omega\times K(T_{\mathrm{mod}})\rightarrow T_{\mathrm{mod}}$. We will note R the set of all possible imitation rules.

For example, if the modification process consists in pure copying, the simpler definition of *conformism* is 'copy the behavior of the majority of agents in your neighborhood'. The simpler definition of *payoff-biased imitation* is 'copy the behavior of the most successful agent in your neighborhood'. Etc. Imitation processes, as we defined them, are part of the category of *cultural learning processes* that Tomasello (2000) defines as *"intentional phenomena in which one organism adopts another's behavior or perspective on some third entity"*. They differ from individual learning in that a mediator is needed for the learning. They differ from other *cultural learning processes* (*exposure, priming, emulation, mimicking*) in that they are voluntary and concern surface behaviour as well as goals and processes.

From this definition, we see that only *modifiable traits* are submitted to imitation although any trait can be used in the definition of the imitation rule. There are potentially as much imitation processes as modifiable traits, and an agent can use the same imitation rule for several modifiable traits. Here, the fact that a rule for imitation takes as input the set of categories of modifiable traits rather than a single modifiable trait is essential. This means that the same rule for imitation can be use to update several kinds of traits which could be qualitatively of very different natures. For example, if you are looking for prestige you might be willing to copy most prestigious person within your neighborhood, having the same education, the same job, the same car, etc. In that case, applying for a particular job or buying a particular car will be part of a single global goal.

c. Metareflexive mimetic systems

When we look at mimetic behaviors in systems of N agents doted with reflexivity and metacognition, we get what we call *metareflexive mimetic systems*. In those systems, rule for imitation are identified as cognitive processes and thus become modifiable traits $(R \subset K(T_{\text{mod}}))$. They can consequently be controlled by others imitation rules (fig. 1.b.2) in a hierarchical way (metacognition). This means that with metacognition, *imitation rules are modifiable traits by way of meta-rules application (fig. 1-b-2).* We can then imagine mimetic systems with several levels organized hierarchically. To a given modifiable trait, we can associate a metamimetic chain that controls its expression. For example, in game G you can imagine that an agent has the following 2-leveled hierarchy: top-level rule: imitate the winner (*payoff biased imitation*), first level rule : do as the others did (because this rule happened to be the most successful), behavior: play C (*cf. fig. 2*). This kind hierarchical organization is very frequent in modeling literature, the originality of our approach lays elsewhere.

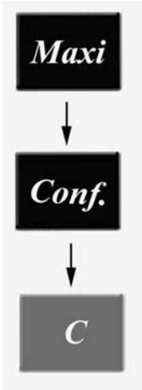

Fig. 1

If our aim is built a model of human behavior, we have to respect what is currently accepted as modeling constraints, and in particular, our agents must have a bounded rationality. The consequence is that *metamimetic chains have to be finite*. For the top-level rule, we face three alternatives. Either we postulate a fixed exogenous rule or a genetic regulation mechanism. These are the two options taken respectively by game theory and evolutionary game theory. This would put the top-level rule in the category *other traits.* But there is a third alternative. Since agents are reflexive, they can use a cognitive process reflexively. According to the definition we gave of a rule for imitation, this means that among the categories an imitation rule can take as inputs, we find the category 'top-level imitation rule'. Thus, *a top-level rule for imitation can update reflexively by acting on itself as a modifiable trait (fig 1-c). We will say that agents are metareflexive.*

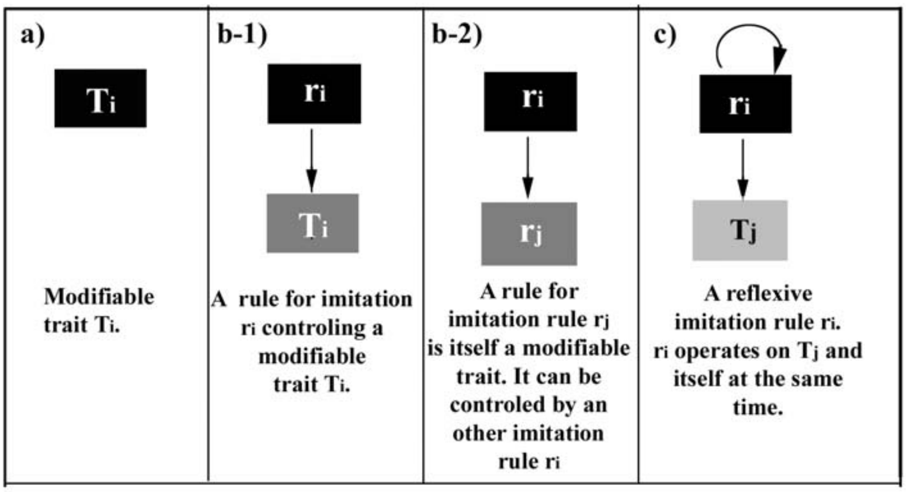

Figure 2 : **The building blocks in metamimetic games**

In metareflexive systems agents can thus have chains of behavior monitoring potentially including several meta-levels. The question is now how agents change the structure of these chains, their length as well as our composition. We will now propose some transition rules for these chains that will define a game we will call the *metamimetic game*. We choose these rules to be as simple as possible and to coincide with standard rules of revision under imitation where this makes sense.

Consider a population of agents that can deal with a maximum of *n* meta-levels (bounded rationality). Call *n* the *cognitive bound* of agents. Consider for clarity that agents have to choose a behavior r_0 for a single type of action (like playing *C or D*). Agents will then be defined by a set of modifiable traits $(r_0, r_1, .. r_k)$ with $k \leq n$, where r_0 is a behavior and r_j are some metamimetic rules. Let for example consider an agent *A* with a cognitive bound of two playing at the game *G* and suppose that *A* is defined by the metamimetic chain (r_0, r_1) where $r_0 = D$ and r_1 is a *payoff-biased mimetic rule* (copy the best agents). We have to define the way this agent updates its different modifiable traits at the different levels.

We have already said that the update of a modifiable traits at and intermediate level (r_0 for example) has no reason to differ from what has been already considered in literature. This kind of updates can be defined in the following way:

Revision by imitation of a modifiable trait at intermediate level (Figure 3)

The revision of a modifiable trait of level $Nc-1$ by the meta-rule of level k can be decomposed in the following steps:

- r_k is used to update the modifiable trait of level *k-1*.
- If during this process, *A* actually chooses to change its modifiable trait of level *k-1,* then two possibilities arise. Either the trait of level *k-1* belongs to the category 'behavior', and the process stops here, or this trait is itself a metamimetic rule, and the agent engages in the revision of the trait of level *k-2*.

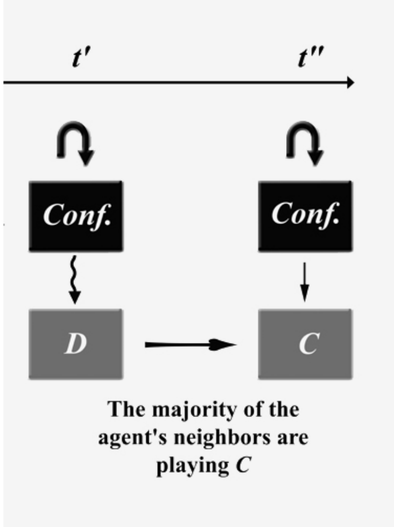

Figure 3 : Update of an intermediary modifiable trait. A conformist Agent *A* observes that the majority has changed for *C,* and decides to play *C*.

Reflexive Update

Reflexive updates take place at the top level of a metamimetic chain. Top-level rules are the most important rules in metamimetic chains since they ultimately determine their dynamics. The way a top level rule evaluates neighbors in a mimetic process is somehow the most important goal for the agent. Suppose that after evaluation, an agent *A* comes to the conclusion that from its point of view, agent *B* is the most successful in its neighbourhood. We will have to distinguish two cases in function of the size $|c_B|$ of the metamimetic chain of agent *B*.

***First case**: Partial metareflexive update ($|c_B|<n$)*

A is able to monitor a *B*-like top-level rule under its own top-level rule r_k. Then, *B* top-level rule is viewed as a mean, eventually temporary, to achieve the goal defined by r_k. For example, if in game *G, A* is a pure maxi-agent (using a payoffs biased rule) that has a cognitive bound of two or more, and if it happens that a pure conformist agent *B* is indeed the most successful, then *A* will perhaps adopt the conformist rule at its first meta-level, keeping in mind that it is only a strategy to maximize payoffs (cf.Figure 4). The payoffs-biased rule will then jump to the second meta-level and *A* will be conformist as long as it will be an efficient strategy.

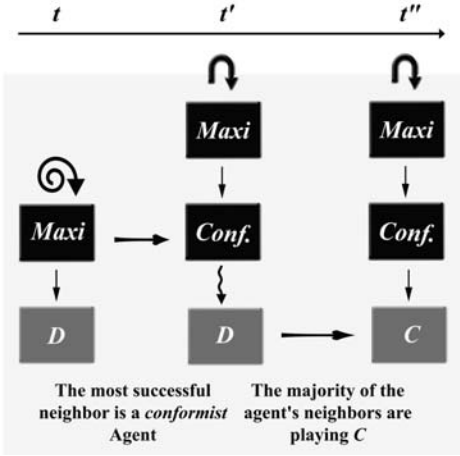

Figure 4 : A case of partial metareflexive update. At time *t*, a *Maxi* agent *A* has a *conformist* neighbour that is strictly more successful than all of its other neighbors. *A* will then adopt the conformist rule at its first meta-level, keeping in mind that it is only a strategy to maximize payoffs (second meta-level). Thereafter, it might be that according to this rule, the current behavior is not the best one, and has to be changed. *A* will be conformist as long as it will be an efficient strategy. .

***Second case** : full metareflexive update ($|c_B|=n$)*

Imagine that the most successful agent in *A's* neighbourhood is using a metamimetic chain of length *n.* Then, if *A* really think that such a structure is needed to have the same performance than *B*, *A* will have to revise entirely its metamimetic chain, i.e. *A* will have to change its top-level rule. For example, in the case of the game *G* mentioned above, it is possible that a maxi-agent *A* with a cognitive bound of only *1,* finds out that one of its conformist neighbour have higher payoffs than any other agents, itself included. To be like *B, A* will then have no other solution than to *become conformist* (Figure 5).

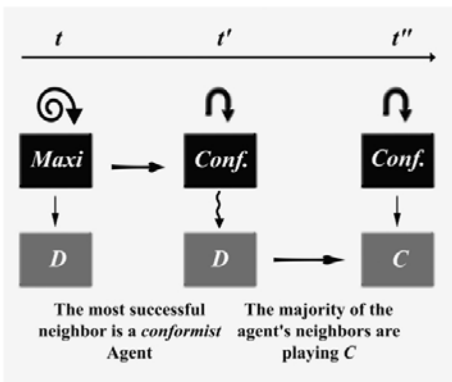

Figure 5 : **A case of full metareflexive update**. At time *t,* a *Maxi* agent *A* has a *conformist* neighbour that is strictly more successful than all other neighbors. Consequently, *A* adopts the *conformist* rule. Thereafter, it might be that according to this rule, the current behavior is not the best one, and has to be changed**.**

Metareflexive updates happen when, after observation of the environment, the metamimetic chain of the agent is not self-coherent. In a way, we can say that the change is motivated by a cognitive dissonance. Contrary to intermediary updates, the *all structure* (length and/or composition) of the metamimetic chain might change in such updates. In particular, it may happen that the changes needed are so deep that the agent has to revise entirely its metamimetic chain, top-level rule included. This happens when the agent reaches the bound of its cognitive capacities and cannot manage both new goals and old ones. Metareflexive update is the core of metareflexive mimetism since has we will see, it is what enables endogenous setting of mimetic rules.

III The metamimetic game

a. Axioms

If we specify the maximum number of meta-levels and adopt the transition rules described above, we get the definition of metamimetic games:

Definition: ***Metamimetic game***

A metamimetic game is a N-players game where agents are imitators and metareflexive. More over, the three following conditions should be satisfied:

*C-I - **Bounded rationality**: in metamimetic chains, the number of meta-levels is finite and bounded for each agent by a fixed integer, its cognitive bound (Bc).*

*C-II - **Metacognition**: at all levels in a metamimetic chain, rules for imitation are modifiable traits by way of meta-rules application. (If R_n is the set of all possible rules for imitation at level n then $\exists i \in \aleph, R_n \subset P_i(T_{\text{mod}})$).*

*C-III – **Reflexivity**: the last level of a metamimetic chain updates reflexively, changing the length of the metamimetic chain in the limit of the cognitive bound of agents. When the cognitive bound is reached, top-level rules update themselves (reflexivity of imitation rules).*

In *metamimetic games* agents will be characterized by a set of metamimetic chains, which, for an analogy with game theory, can be viewed as representing embedded metagames. The study of metamimetic games will thus consist in the study of the evolution in length and composition of these chains, leading to the emergence of structures, at the intra-individual level as well as at the inter-individual level.

b. The unsatisfiability

We will now briefly develop a formalization for metamimetic that will be useful in understanding the global dynamics of meta-rules. In metamimetic games, individuals come to change their modifiable traits when discovering that they are poor performing compared with neighbors. This leads us to define the notion of unsatisfiability:

Definitions

Unsatisfiability: *Let $c=(r_0,r_1,.. r_k,)$ be a metamimetic chain of an agent A. We will say that A is unsatisfied if there exist a neighbour A' of A, with a metamimetic chain c' $\neq$ c, such that evaluated by A, A' performs strictly better than A. The individual unsatisfiability $f(c)$ of agent A will then be defined as the probability for A to be unsatisfied. It is the sum on the set of metamimetic chains c' present in A's neighbourhood (V_A) of the probabilities that A adopt c', $f_{A,c}(c')=P_A(c \to c' \neq c)$. If we look at this phenomenon at the population level, we get the notion of unsatisfiability of a metamimetic chain c by a metamimetic chain c', $F_c(c')$. If we write c_A the metamimetic mimetic chain of agent A, we have:*

$$F_c(c')=\sum_{A,c_A=c}\sum_{c'\in V_A \wedge c'\neq c} f_{A,c}(c').$$

It is the probability that an agent A with the chain c will adopt the chain c'. Finally, we get the definition of the unsatisfiability of a metamimetic chain c :

$$F_c=\sum_{c\neq c'} F_c\,(c')$$

In the same way, we define the n-unsatisfiability of the top n components of a metamimetic chain $(*,c)=(*,r_{k-n} \ldots,r_k)$ as the probability that one of the n top meta-rules in chain C will be modified in an update. For example, we can speak of the unsatisfiability of a (D,maxi) agent, or of the 1-unsastisfiability of maxi agents.

Relative unsatisfiability *: To understand the dynamics of metamimetic games, we have to be able to compare the unsatisfiability of different types of chains according to their frequency. The idea is that the proportion of a given type of metamimetic chain will be stable if there are, in mean, as many agents that adopt this chain than agents that quit for another chain.*

In the case of discrete time dynamics on discrete populations, we can write the equation for the evolution of the proportion of c metamimetic chains :

$$\Delta p_c^t = -\sum_{A, c_A = c} \sum_{c' \in V_A \wedge c' \neq c} f_{A,c}^t(c') + \sum_{A, c_A \neq c} \sum_{c \in V_A} f_{A,c'}^t(c)$$

If we write p_c^t the proportion of chains c at time t in the population and define the relative unsatisfiability by:

$$\hat{F}_c^t = p_c^t . F_c - \sum_{c' \neq c} p_{c'}^t F_{c'}^t(c)$$

the master equation of metamimetic games can then be rewritten :

$$\Delta p_c^t = \hat{F}_c^t$$

This equation defines dynamical processes that belong to the class of replication by imitation (see for example Weibull 1995). However, they correspond to none of the processes already studied, which are more or less are equivalent to the replicator dynamics. It can be shown easily from these equations (see annexes) that the discrete replicator dynamics (see for example Hofbauer & Sigmund 1988, Weibull 1995) is a particular case of metamimetic dynamics where there is only one meta-rule, which imply that metamimetic dynamics are not reducible to replicator dynamics.

We expect that the study of these meta-dynamics as defined above will be of high interest to understand social dynamics. In this perspective, there would be a lot to say to link metareflexive mimetic systems with existing theories of social cognition. First, things are surely not as black and white as our description of metamimetic chains and their evolution. However, we think it is a good starting point. Second, if metamimetic games are to be a part of a general model of social cognition, we will need to specify their articulation with existing theories of inference, memory and learning.

In particular there would be a lot to say about the relation between metareflexive mimetism, which is backward looking, and fictious play, which is forward looking (see for example Young 2001), in the process of decision-making. We can reasonably think that both processes are complementary, acting in parallel, and can be unified in a single framework. Such an approach is desirable but is beyond the scope of the present work. We let this for future research. Nevertheless, we can notice that in the framework of standard game theory, agents don't change their goals and preferences during their life, which was stressed to be one of the limits of the theory (Bowles 2001, Gintis 1998, Henrich et al. 2001, Lessourne 2004). On the contrary, evolution of individual motivations is the most important feature in metareflexive mimetic systems (condition C-III). Consequently, we might expect that the cross fertilization between metareflexive mimetism and game theory will provide an interesting approach to the modeling of the evolution of preferences.

For the end of this paper, we will abstract away from these complications to focus on a first example of metamimetic game that is an extension of the spatial prisonner's dilemma. This will help us to have a first intuition of what kinds of dynamics we can expect to find.

IV The Spatial Dilemma Game in metamimetic framework

To give a first idea of the properties of *metareflexive systems* we will see here an example of a metamimetic game with an extension of the spatial prisonner's dilemma game (Nowak and May 1992). First of all, it is important to stress that metareflexive mimetism could be seen as a way for embedding other models, or as a plug-in to add to existing ones. It is not aimed at replacing any existing theory. Consequently the aim here *is not* to give a fully integrated model of human behavior, but to show how we can embed a given model in the metamimetic framework. Here, it will obviously not solve the original limitations of the model of Nowak and May – exogenous network, absence of individual learning, and pure copying mechanism without inference processes – but we will see that the dynamical properties of metamimetic games and in how, in the particular model of the spatial prisonner's dilemma, it can account for sustainability of a certain levels cooperation under all range of the parameters studied.

The problem of emergence of cooperation and its sustainability is an active domain of research with still a lot of open questions (Hammerstein 2002), especially in the domain of human cultural evolution. The standard model for cooperative interactions is the prisonner's dilemma, which we described in section II-a. Nowak and May (1992) proposed a spatial version of the prisonner's dilemma game based on memory less agent guided by payoffs-biased imitation. The interest of such spatial game is that it illustrates the spreadingof behaviors among populations, revealing very interesting phenomena.

There has been a lot of developments based on the spatial prisonner's dilemma game (Brauchli 1999, Duràn 2003, Lindgren & Johansson 2002, Nakamaru 1997, Nowak et al. 1994), studying the sustainability of cooperation under various conditions. These studies often consider more sophisticated agents than those of the original model. Coming back to the model of Nowak and May will thus enable us to see the pure effect of reflexivity, without any further sophistication. Moreover, as we will see, the spatial settings of Nowak and May are the worst situation for cooperation since agents cannot discriminate between neighbors.

a. The model of Nowak and May 1992

The original model of Nowak and May (1992) considered memoryless players choosing between two simple strategies: always cooperate (*C*) or always defect (*D*). These players are placed at the nodes of a two dimensional toric grid, the game is organized in rounds and agents are interacting repeatedly in discrete time. Each round, agents play a prisonner's dilemma game (as describe in II-a) with the same strategy (*C or D)* against each of their eight closest neighbors (players in the eight adjacent cells) plus themselves. Between two rounds, players compare their payoffs with those of their eight neighbors. In case one neighbour is strictly more successful than all other neighbors (themselves included), they adopt its strategy (*payoff biased imitation)*.

This kind of settings illustrates situations where individual's interests are in conflict with the collective interest and where you cannot interact selectively with your neighbors. Cooperation benefits to all your neighbors and defection punishes all of them.

It is consequently the worst case for cooperation. This illustrates common pool resources dilemma (Ostrom 94). For example, agents can be farmers pumping underground water diminishing the stocks of their neighbors. In a drought period, if agents do not respect water restrictions they will dry-out the soils. Then, only those who will have made water stocks will still have water left. But the collective interest is that every body pumps water in reasonable quantities, so that underground water stays available all the time.

Nowak and May studied this spatial game with the particular conditions $S=0$, $P=0$, $R=1$ and $T>1$. Although this is not a prisonner's dilemma game ($P=S$) they assumed that their finding were not qualitatively altered if $P=\varepsilon$ with ε positive but significantly below unity. In this case, the dynamics reported is two folds (more details will be found in Nowak and May 1993, Nowak et al. 1994): for $p<1.8$ or $p>2$, the dynamics converges almost always toward a stable state, where for $p<1.8$ cooperator generally are predominant, while for $p>2$ defectors are generally predominant. The most interesting regime is for $1.8<p<0.2$ where we can observe dynamic pattern of cooperator areas and defector areas evolving in a chaotic way, but keeping on the long run the rate of cooperation constant at a level around *31.2%.* It is this kind of dynamics that made the success of this model since it exhibits complex patterns with heterogeneous population, and consequently a certain amount of cooperation, in an area where the social dilemma was quite important (the advantage of defection almost twice the advantage of cooperation). However, the standard parameters for the prisonner's dilemma game are $T=5$, $R=3$, $P=1$ and $S=0$ which in the notations of Nowak and May correspond to $b=5/3$ and ε=1/3. How robust are their findings in this case? We did a computational study with the multi-agents plateform designed to simulate metareflexive mimetic systems. We first make sure that we could correctly reproduce Nowak and May results (see web appendix http://chavalarias.free.fr/metamimetism.htm). Then we did a computational study for $b=5/3$ and ε=1/3. Our results are clear (Figure 6)

1- In the case $b=5/3$ and ε=1/3, the chaotic regime disappears.
2- If we avoid the particular value of b for which there are ties, the dynamics converges toward a static state mostly cooperator for $b<1.33$ and a static state mostly defector for $b>1.33$.
3- This threshold falls at *1.1* if we do not consider self-interaction.

An extensive study of the behavior of spatial games with the settings of Nowak and May for a large area in the parameters space of the matrix of the game can be found in Hauert 2001. It confirms the fact that chaotic dynamics cannot be found in prisonner's dilemma game (except for very low strength of social dilemma). Consequently, the interesting pattern emergence described by Nowak and May is specific to the game they choose which has definitely not the same dynamics than a standard spatial prisonner's dilemma. By passing, payoffs-biased imitation is not sufficient for the sustainability of cooperation in spatial prisonner's dilemma game with such settings.

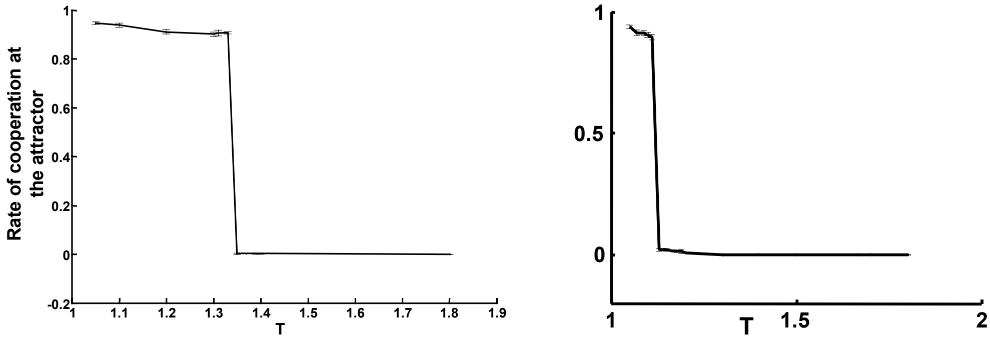

Figure 6 : Evolution of the rate of cooperation in funtion of *T* at the attractor in Nowak and May model for *P=1/3*, *R=1 et S=0.* Initial conditions where random uniform with *50%* of cooperators. Left *:* The game with self-interaction. Right: the game without self-interaction. The chaotic regime has disappeared. The final state is either almost all *D* or all *C*.

b. The metareflexive prisoner

We will now study a game that is the natural metamimetic extension of the precedent model. The general settings will be the same except for the fact that the matrix of the game will be a prisonner's dilemma game; *we will not consider self-interaction* and the mimetic dynamics will be those presented in section *II-a*. We will consider mimetic rules that are modifiable traits (condition *C-II, metacognition*) and update reflexively (condition *C-III, reflexivity*). To keep the same agent structure as Nowak and May, we will take agent with a cognitive bound of *1 (figure 7)*. We thus get the simplest structure for a metareflexive agent. The two categories of modifiable traits of the game are the behavior and the meta-rule. Consequently we have *K*(T_{mod})*={'behavior,'meta-rule'}*. Payoffs are in the category *other*

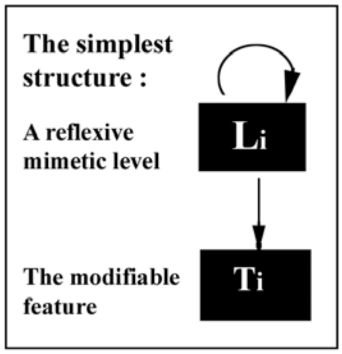

Figure 7 : the simplest structure for a metareflexive agent : one behavioral level an one meta-level.

traits. An agent is then characterized by a set $(r_0,r_1,g) \in R_0 \times R_1 \times \Re$ and its metamimetic chain is (r_0,r_1). For sake of readability, we will write these chains $c=(b,r)$. Agents will perceive all kinds of traits *i.e.* the meta-rules, the behaviors and the payoffs of their height neighbors. As in Nowak and May, we will take $R_0=\{C,D\}$. As for the set of metamimetic rules, we will consider the simplest rules that can be built from the perceived traits. We can relate these rules to the most frequent rules considered in literature:

1. ***Maxi rule***: "copy the modifiable trait of your neighbour with the maximum payoffs" (*i.e.* the payoffs biased rule of Nowak and May). This rule characterizes selfish agents.
2. ***Mini rule***: "copy the modifiable trait of your neighbour with the minimum payoffs". This rule can be related to altruist or generous rules.
3. ***Conformist rule*****:** "copy the modifiable trait used by the majority of agents".
4. ***Non-conformist rule***: "copy the modifiable trait used by the minority of agents"
5. ***Random rule***: "copy a modifiable trait at random"

We have $R_1=\{Maxi, Mini, Conformist, Non\text{-}Conformist, Random\}$.

We would like to insist on the fact that these rules apply to *categories* of modifiable traits and can be used for the update of a trait of any category. For example, a conformist agent will copy the most common behavior when updating its trait of category 'behavior', and will copy the most common metamimetic rule in neighborhood when updating its trait of category 'metarule'.

c. Computational Study

We will now give some computational results. The idea is to study the evolution the different proportions of rules and behavior with time under various environmental conditions. As in Nowak and May's article, we consider here parallel updating but we checked that the dynamics is qualitatively the same in continuous time. Making them endogenous could in fact elude the problem of the update frequencies constants, which is a natural operation in our framework. The idea is to say that the frequency of the reflexive update and the update frequency of the modifiable trait are part of the description of the rule and consequently, are copied with the rule, eventually with some errors. This is an interesting option but for sake of clarity, we will discard it in this paper. In the following simulations, we will then consider parallel updating: at each period, each agent updates reflexively its meta-rule and then updates its behavior. Since agents' cognitive bound is *1*, agents will have no choice other than full reflexive updates when revising their meta-rules.

The complete description of the algorithm used for these simulations can be found in appendix. We will present computational results in the following way :

α - *A* detailed study for a the standard setting of the prisonner's dilemma game: *T=5, R=3, P=1* and *S=0,* and an initial rate of cooperation (*Inicoop)* of *30%*

β - *A s*tudy of the influence of the advantage of defection on cooperation as determined by the matrix and the initial rate of cooperation, on the dynamics outcomes (Cartography of attractors).

***α* - A particular case**

The first results presented here have been done with the five rules mentioned above. Population size is *10 000* agents. As initial conditions, we took a uniform distribution for metamimetic rules, and a level of cooperation of *30%* uniformly distributed in the population. The first striking phenomenon is that the system reaches very quickly its unique attractor (fig. *8 & 9*), which is mostly static (only a few number of oscillators). This means that at the attractor, the unsatisfiability of most agents is zero. They thus perform repetitive behavior.

Moreover, the path to this attractor is mostly the same along the different simulations, with a very low variance on the distributions of imitation rules and strategies. Here, graphs show the evolution until *100* periods, but we checked the stability of the attractor for large time scales. This attractor is heterogeneous for both distributions of imitation rules and strategies, with formation of stable clusters of cooperators (*cf. fig 10 & 11*). As we saw in last section, cooperation wouldn't have been sustainable if there were only *maxi*-agents. Here the rate of cooperation increases from *30%* to *42%* during the simulation.

We can also notice that the *random* rule completely disappeared at the attractor. This emphasizes a characteristic of meta-dynamics: top-level metamimetic rules should be their own preferential trait at the attractor. Since this is never the case for the random rule, which chooses indifferently among rules present in the neighbourhood, it is bound to disappear. Metamimetic dynamics allow stochastic rules but at intermediate meta-levels, which are not permitted here. Consequently, in the following simulations, we won't consider *random*-agents any more. In the same way, we can check that at the attractor the situation make sense for agents when compared with the goals they have: *non-conformist* are actually locally in minority, *conformist* are actually locally in majority, *maxi* and *mini* agents have interlaced populations and are locally the most (resp. the least) successful agents (*cf. fig 10*).

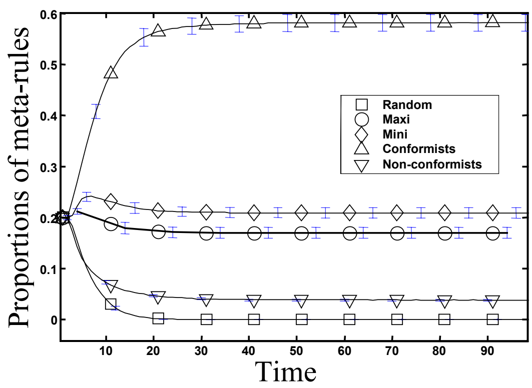

Figure 8 : Statistics on the evolution of of metamimetic rules for the game defined by *T=5, R=3, P=1* and *S=0*. Initial conditions are a uniform distribution of metamimetic rules and 30% of cooperators. The distribution of rules quickly converges to the attractor where *conformists (*up triangles*)* dominate. The *Random rule* disappears. Error bars represent the standard deviation (*50* runs, *10 000* agents each).

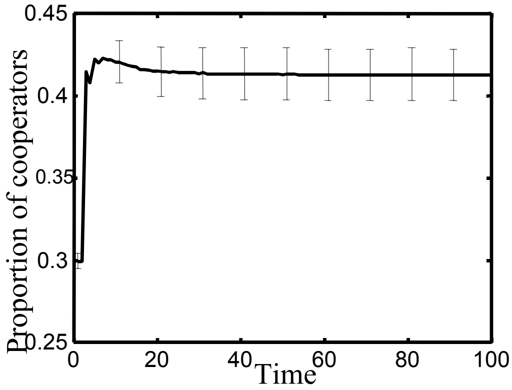

Figure 9: Statistics on the evolution of cooperation. The rate of cooperators increases from *30%* until population reaches a level of *41%* of cooperators**.** Error bars represent the standard deviation (*50* runs, *10 000* agents each).

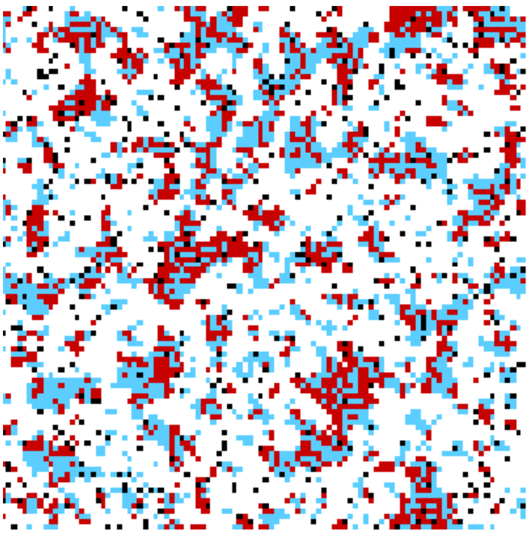

Figure 10: **The spatial distribution of metamimetic rules.** Each small square represents an agent. (here *10 000* agents on the toric grid). At the asymptotic state, *non-conformist* agents are actually locally in minority (black squares), *conformist* agents are actually locally in majority (white clusters). *Mini (*blue or light grey*)* and *maxi* (red or dark grey) agents have interlaced populations.

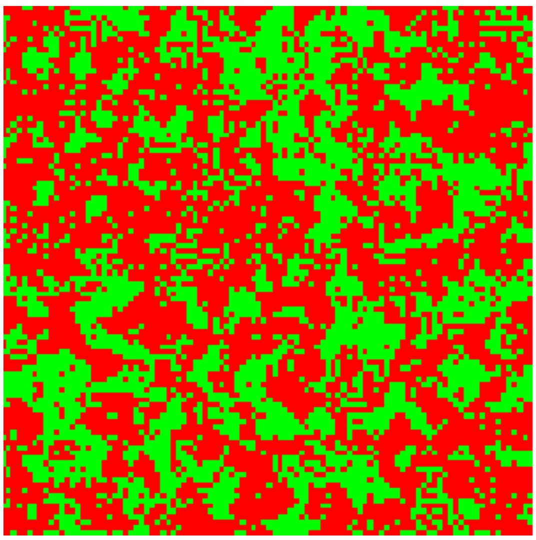

Figure 11: **The spatial distribution of behaviors at the attractor.** We can observe the emergence of sustainable clusters of cooperators (green or light grey).

β - Cartography of attractors

To study the influence of the initial rate of cooperators (*IniCoop)* and the strength of the temptation for defection on the dynamics, it is more convenient to consider a matrix of the game described by only one parameter. It is well known that two parameters suffice to describe the whole set of distinct games. The problem now is to select a subset of this *2D* space that would nevertheless generate all the interesting dynamics. For this purpose, we will take a parameterisation frequently used in the social dilemma literature associated with the following payoffs matrix:

Agent B → ----------- ↓ *Agent A*	*C*	*D*
C	*(1-p,1-p)*	*(0,1)*
D	*(1,0)*	*(p,p)*

Table 1 : The matrix of the game

We will assume that $0<p<0.5$ so that the condition $T>R>P>R$ is satisfied. The condition $T+S<2R$ is violated (we have equality) but it doesn't have noticeable consequences on the dynamics. This condition expresses the fact that there is no way that two players could share their rewards so that it would be more advantageous for them to have distinct behaviors (*D* and *C*) than to cooperate. Since in our game, such a share is impossible, this condition is not very important.

In this spatial game, the payoffs of an agent *A* with a neighborhood V_A is thus (where b_B is the behavior of agent *B)*:

- If *A* played *C :* $g_A(C) = \sum_{B \in V_A \wedge b_B = C} (1-p)$
- If *A* played *D :* $g_A(D) = \sum_{B \in V_A \wedge b_B = C} 1 + \sum_{B \in V_A \wedge b_B = D} p$

We can check that we have, $g_A(C) < g_A(D)$ whatever the composition of neighbourhoods. A quick study of this setting in the case with only *maxi*-agents will reveal quite similar dynamics to those presented in *IV.1.* with a threshold at 0.2 : under initial random conditions of *50%* of cooperators, for *p<0.2* cooperators invade, for *p>0.2,* defectors invade. We will now see what happens when agents are metareflexive.

We did a study similar to the case presented in paragraph IV.1-α, for *Ini Coop* varying between 5% and 95% and parameter *p* varying between *0.1* and *0.45.* The same qualitative properties were observed concerning the attractors but the full description of these results falls beyond the scope of this article and will be found in further papers (see web appendix for the full graph). We will just give here a rough description of the metamimetic dynamics observed.

Behavioral level : The rate of cooperation at the attractor is plotted on Figure 12. We can see that this rate is always above *9,5%* and above *40%* in the majority of the cases. Attractors at the behavioural level depend heavily on *IniCoop* for low *p* but are almost independent of *IniCoop* for *p>0.2*. On the contrary, *p* has always a great influence on these attractors. Even if there is no space to develop the point here, it is noteworthy that the high level of cooperation for most of the parameters space is a very interesting result in the perspective of the emergence of cooperation.

Meta-rules level : Here *Inicoop* has even less influence on the meta-rules than ithas on behaviors (see web appendix for more details). The proportions of *conformist* agents decrease when *p* increases, while the opposite phenomenon happens with *maxi* and *mini* agents. If *conformist*agents are always the population with the highest density (cf. Figure 13), there is a significant proportion of *maxi* and *mini* agents for *p>0.2* (more than *20%* for each population). On the contrary, the proportion of *non-conformists* is not sensitive to *p* and is almost constant along both axes (it seems to be a function of the topology of the social network). In further work, we will demonstrate that the process of clusters formation is more favourable to *maxi* and *mini* as *p* increases. This means that the relative unsatisfiability of *maxi* and *mini* decreases with *p.* At the attractor, most agents perform repetitive behavior without changing anything at their behavioural level or meta-level. However, few agents, at the border of clusters, keep changing one of these two modifiable traits. We will see why in next section.

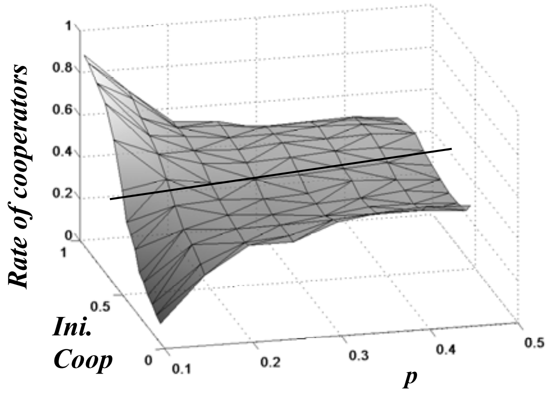

Figure 12 : Dependence of the rate of cooperators at the attractor (100 time steps) in function of parameters *p* and the initial rate of cooperators *Ini Coop.* The rate is always above *9.5%.* Cooperation is sustainable under all the conditions studied. The line represents the set of simulations corresponding to the area of parameters used in **Figure 13**

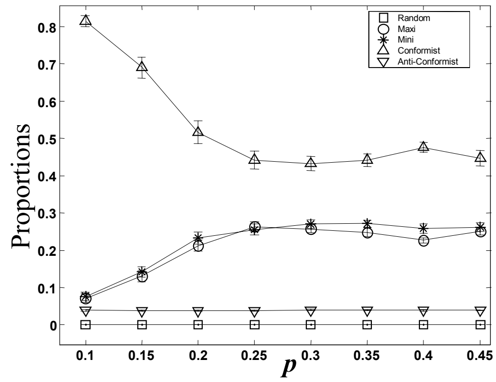

Figure 13 : Evolution of the distribution of metamimetic rules in function of *p,* for an initial rate of cooperation of *50%.* We can see that *conformists (*up triangles*)* are predominant under all condition but *maxi* (circles) and *mini* (stars) agents do better in area concerned with high *p* than in those concerned with low *p*. Each point has been obtained with 10 runs. Error bars represent the standard deviation.

d. At the border of social groups

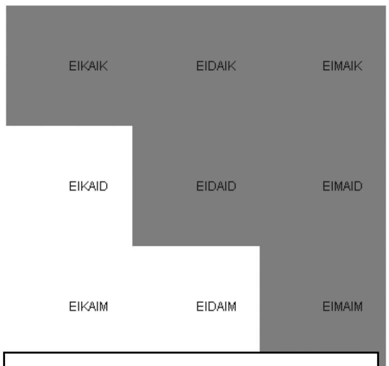

Figure 14 : Neighbours of Edaid. Light grey agents are

To understand why some agents are perpetually unsatisfied at the attractors and keep oscillating between several meta-rules, we will take the example of a particular agent, Eidaid that hesitates perpetually between the *conformism* rule and the *maxi* rule. This agent is at the border between a *conformist* area and a *maxi* area (*cf.* figure 8). All her neighbours are defectors. It is easy to see why Eidaid can't keep the conformist rule: the majority of its neighbours are *maxi-agents.* Then, each time Eidaid becomes *conformist,* it will update its rule to *maxi* the next period. On the other hand, a study of payoffs distribution shows that the most successful agent in Eidaid's neighbourhood is Eidaim, a *conformist* agent*.* The reason is that Eidaim has the chance to have an *non-conformist* neighbour, which is playing *C*. Thus, each time Eidaid adopts the *maxi* rule, Eidaid will copy Eidaim at the next period, and become *conformist* again. Endlessly. In the case, we can say that Eidaim is frustrated. We can see here that the topology of the social network is crucial for this kind of phenomena. It is precisely because Eidaid wants to imitate some neighbours without having the same neighbourhood and thus, the same information that it always hesitates between the two rules. The fact that this kind of configurations can only happen at the border between two areas (of homogenous behaviors or meta-rules) with the fact that the whole population is strongly structured explains why there are so few frustrated agents.

e. Failing to imitate

To address the question of endogenously fixed distribution of mimetic rules, we should be able to consider systems starting from any initial distribution of metamimetic rules, and study its evolution. At this stage, we have to introduce a noisy component in the system since otherwise, the evolution of systems starting from a homogeneous type's distribution would be trivial. We studied systems with noise at level ε as modelled for example in Young *2001*. We considered that agents imitate according to their rule with a probability $1-\varepsilon$, and adopt a random modifiable trait among all the possibilities with a probability ε. This noise could represent errors at the level of inference, copying, decision or action.

To study the dependence of the initial distribution of rules on the final distribution we did several simulations with five different initial conditions for the distribution of mimetic rules: a uniform distribution, and the four homogeneous distributions (only *maxi*, only *mini*, etc.) The influence of the noise level was as presented above. We present here results of simulations with a low level of noise, after *2000* periods. Again, agents are memoryless and perceive only the current period payoffs. Parameters are: $\varepsilon=0.005$, $p=0.3$ and *Ini Coop=50%,* (cf. Figure 15). We can see that the final distribution of rules is the same for the different initial conditions: around *80%* of *conformist* agents, about *10%* of *non-conformists* and about *5%* of *maxi* and *mini.* The study of the level of cooperation across initial conditions reveals that this level varies between *51%* and *60%*. The reason is that the final distribution of meta-mimetic rules is mostly conformist, which mean that the systems at the sensitive to initial conditions. On the other hand, the evolution of cooperation in first periods heavily depends on the composition of population at the meta-level. Consequently, we have a path-dependent dynamics on the behavioural level. The poor performances of *maxi* and *mini* compared to same simulation without noise (fig 8-a) can again be explained qualitatively by the unsatisfiability. Noise increases the absolute unsatisfiability of all rules since it introduces uncertainty, which causes the agents to take wrong decisions more often. However, *conformists* and *non-conformists* are less sensitive to noise than *maxi* and *mini* since their rules for imitation are indexed on densities (aggregated data), which are more stable under noise than last period payoffs of single agents. Consequently noise has less impact on their unsatisfaction than on those of *mini* and *maxi*. The fact that the final proportions of metamimetic rules in these simulations do not depend on initial conditions on these proportions suggests that proportions of metamimetic rules are a property of environmental conditions (*p and* ε). Future studies will show that this is actually the fact: the final distribution of meta-rules in this game do not depends on the initial distribution of meta-rules or behaviors. It is a function of *p* and noise levels at the different levels of metamimetic chains.

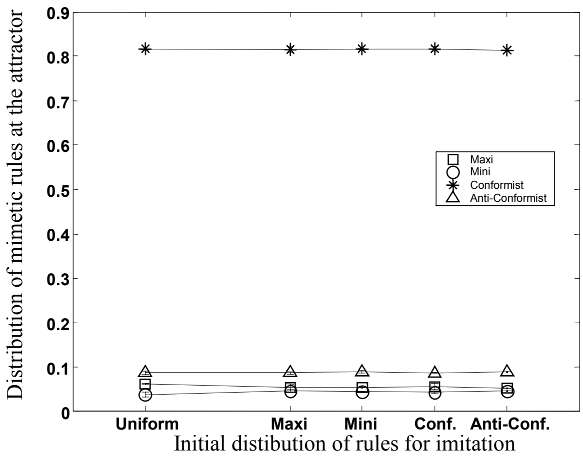

Figure 15 : Influence of initial conditions in noisy games: The distribution of metamimetic rules at the attractor (period *2000)* in function of the initial distribution of rules. Initial conditions on behaviors are *50%* of cooperators. The first point, *Uniform*, stands for a uniform distribution of rules at the beginning of the simulation. Other points are for homogeneous initial distributions with one of the four types. We can see that final distributions are the same.

Conclusions

As heuristic for the modelling of human social systems, several scientists proposed to focus on models that include human specific cognitive capacities. The reason is that only such models should be able to explain the huge gap of complexity in social structures between animal's and human's societies. Following this heuristic, we proposed a schematic representation of reflexivity, a property that is well known to be a specificity of human cognition, in the framework of mimetic systems. This led us to the notion of *metareflexive mimetic systems*. To conciliate metareflexive mimetic systems with the requirement of bounded rationality, we defined what we have called *meta-mimetic games*. Those games have the particular properties that, first mimetic rules can be their own meta-rules, second, we get a meta-dynamics on mimetic rules without the need to specify any other evolutionary process like, for example, a replicator dynamics. We further show with a first computational study around the spatial prisonner's dilemma game, that these meta-dynamics exhibit strong attractors with heterogeneous population and patterns emergence. In particular, we have seen that in the case of the spatial prisonner's dilemma, cooperative structures emerge and are sustainable in all the domain of parameters studied even though agents are memoryless with non-selective strategies (*all C* or *all D*). This is due to the fact that in meta-mimetic games, players are not stuck to a single goal like payoffs maximisation, which generates the dilemma, but can change their goal under social influence. In this way, they can locally collectively get out of the dilemma.

This is only a first approach of metareflexive mimetic systems than opens the door for the endogenization of mimetic rules and other kinds of human's traits (behaviors, update time frequencies, preferences, etc). However, we have tried to show that metareflexive dynamics are a fascinating field of investigations with rich spatio-temporal patterns concerning the traits studied (here imitation rules and behaviors) and numerous possibilities of extensions. We expect future works to take several directions:

- 1. As already mentioned, there is a lot of work to do in order to link this framework to existing theories related to human behaviour *i.e.* inference, memory and learning. *Inference* deals with the way people extract information from their environment and in particular, how theyinfer the rules others are usingwhat are the error rates during these processes and how it could be formalized. Our framework is therefore closely linked to this topic. In effect, we might expect that some rules are easier to infer or are less error prone, which will have as a consequence the increase in the satisfiability of their users, and thus their proportion in population. *Memory* can be used mainly to increase the space of the rules considered by increasing the number of event they are build on. For example, we considered here memoryless agents that can establish judgment only on the current round. It could be advantageous for maxi-agents for example, to consider the averaged maximum payoffs on a given number of rounds, this maximum number being bound by the memory size of the agent. In that case, the time window would be part of the description of a maxi-rule and therefore will be endogenous. *Learning* is perhaps the domain where research can be the most exciting. The most sophisticated learning method will never provide the learning criteria; it is just not the scope of learning. On the other hand, metamimetic games provide a way to consider endogenous goals, formed on what agents can perceive, but the way agents can improve their behaviors with individual learning to achieve these goals is not modelized. The integration of both conceptions in a same framework would then enable to study the all chain of the decision processes.

- 2. We saw that environmental conditions were crucial to determine the dynamics and especially the level of noise in the system. Yet, it is desirable to give particular attention to the modeling of noise at the different levels of the cognitive processes and see its influence on the dynamics. It is here a quite challenging program that will surely give interesting evidences of the structuring power of noise in these particular dynamical systems.

- 3. In a cultural evolution perspective, we might ask what is the evolutionary advantage of metareflexive mimetic systems and their impact on cultural evolution. This would assure us that the reflexivity is not just a gadget disconnected from the evolutionary paradigm. The first study here suggests that reflexivity enables high level of cooperation in population, giving an evolutionary advantage to groups of metareflexive agents. This is only a preliminary work and the study must be continued. In particular, we could see if the emergence of reflexivity as described here is a plausible step in a scenario for the evolution of human societies. This may leads to new perspectives in the framework of the gene-culture co-evolutionary theory.

Acknowledgments :

The author would like to thank T.K. Ahn, Jean-Pierre Dupuy, Jean-Louis Dessalles, Wander Jager, Marco Janssen, Nicolas Jonard, Jean Petitot, Lucien Scubla, Richard Topol, and Gérard Weisbuch for helpful discussion and comments.

A particular thanks to Paul Bourgine, my PhD director, who is always enthusiastic on the topic of ideas about ideas.

This paper was supported by the French National Center for Scientific Research (CNRS) and the research program MEL/OHLL (Modeling of the Emergence of Language).

Appendix I : Methodology

The algorithm used for the simulations presented in this paper is the following:

Set up of the game:

- Give a value for p (here $.1 \leq p \leq .45$).
- Neighbourhood composed by the eight adjacent cells. *Toric grid.*

Initial Conditions:

- Give the spatial distribution of imitation rules.
- Give the spatial distribution of behaviours.

At each period, for each agent:

- The imitation rule is used to update itself. An agent changes her rule if there are some neighbours *strictly* more successful than her. For example, if agent *A* had the *Conformist* type and if the majority of her neighbours have turned to *Maxi* since last round, *A* will adopt the *Maxi* rule.
- The imitation rule (eventually new) is used to update the behaviour. If *A*, a *Maxi* agent, played *C* last round but a *D*-player did strictly better than all *A*'s neighbors (*A* included), *A* will become a *D*-player.
- The agent plays with her height neighbours.
- The new payoffs of agents are computed by summation of the height scores of the two-players games.

Appendix II: The metamimetic game and the replicator dynamics

Proposition : *The discrete replicator dynamics is equivalent to a metamimetic game on a complete graph with a single meta-rule.*

Proof :

This is straightforward since what is important is that the master equation of metamimetic games and the discrete replicator dynamics equation both belong to the category of balance equations. However, it points out some differences between the two kinds of dynamics. The standard form of discrete replicator dynamics for a population of strategies *(1,..,n)* with proportions $\sigma(t)=\left(p_1^t,\ldots,p_n^t\right)$ can be written (Hofbauer and Sigmund 1988):

$$p_i^{t+1}=\frac{a+f_i(\sigma_t)}{a+\bar{f}(t)}p_i^t$$

with $\hat{f}(t)=\sum_{i=1}^{n}p_i^t f_i(\sigma_t)$.

It can be rewritten :

$$\Delta p_i^t=\left(\frac{a+f_i(\sigma_t)}{a+\hat{f}(t)}-1\right)p_i^t$$

Let us consider a metamimetic game on a complete graph with a single rule *r* and a set of behaviors *(1,..,n)*. The metamimetic chains can then be named after the associated behavior. Let's consider for the rule *r* the following stochastic metamimetic rule: an agent *A* will imitate a neighbor *A'* with a behavior *j* with a probability proportional to $\frac{a+f_j(\sigma_t)}{a+\hat{f}(t)}$. Since each neighborhood contain the whole population (the graph is complete), we then have $\forall i,j: F_i^t(j)=\frac{a+f_j(p_j^t)}{a+\hat{f}(t)}p_j^t$.

We replace this relation in the master equation:

$$\Delta p_c^t=-p_c^t.F_c+\sum_{c'\neq c}p_{c'}^t.F_{c'}(c)$$

$$\Delta p_i^t = -p_i^t . \sum_{j \neq i} \frac{a + f_j(p_j^t)}{a + \hat{f}(t)} p_j^t + \sum_{j \neq i} p_j^t \frac{a + f_i(p_i^t)}{a + \hat{f}(t)} p_i^t$$

$$\Delta p_i^t = -p_i^t . \frac{1}{a + \hat{f}(t)} \sum_j p_j^t (a + f_j(p_j^t)) + \frac{a + f_i(p_i^t)}{a + \hat{f}(t)} p_i^t \sum_j p_j^t$$

$$\Delta p_i^t = \left(\frac{a + f_i(p_i^t)}{a + \hat{f}(t)} - 1 \right) p_i^t \; .$$

The discrete replicator dynamics translated in terms of metamimetic dynamics is thus equivalent to the particular case of a metamimetic game with a single meta-rule that can be formulated by “imitate neighbors at random proportionally to their fitness”. ■

In particular, we can see that metamimetic dynamics are not akin to replicator dynamics as soon as there is more than one meta-rule.

References

Alvard, M. S., The Adaptive Nature of Culture, *Evolutionary Anthropology*, 136-149, 2003

Arthur B. W., Inductive reasoning and bounded rationality, *Am. Econ. Assoc. Papers and Proc* 84, 406-411, 1994

Axelrod R., The Dissemination of Culture: A Model with Local Convergence and Global Polarization, Axelrod R., *Journal of Conflict Resolution*, 41, 203-26., 1997

Bandura A., *Social Learning Theory,* Englewood Cliffs, NJ: Prentice-Hall, 1977

Bourgine P., What is Cognitive Economics, in *Cognitive Economics*, Bourgine P, Nadal J-P (ed.),Springer, 2004

Bowles S., Gintis, H., Evolution of reciprocal preferences, *SFI working paper*, 2000

Bowles S., Gintis H., The Origins of Human Cooperation, Peter Hammerstein (ed.) *The Genetic and Cultural Origins of Cooperation*, Cambridge: MIT Press., 2003

Boyd R., Richerson P.J., *Culture and the Evolutionary process*, Chicago, IL: University of Chicago Press., 1985

Brauchli K., Killingback T. & Doebeli M., Evolution of Cooperation in Spatially Structured Populations, *J. theor. Biol*. 200, 405-417, 1999

Byrne R.W., Russon A.E., Learning by Imitation: a Hierarchical Approach, *Behavioral and Brain Sciences*, 1998

Cavalli-Sforza L.L., Feldman M.W., *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton Univ. Press 1981.

Damasio R. A., *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, éditions Odile Jacob, 1999

Donald M., *Origins of the Modern Mind*, Donald M., Cambridge, MA: Harward University Press, 1991

Duràn O., Mulet R., Evolutionary Prisoner's Dilemma in Random Graphs, arXiv:cond-mat/0305353 v2 21 May 2003, 2003

Fehr E., Fischbacher U., The nature of human altruism, *Nature*, Vol 425, october, 2003

Feldman M. W., Laland K. N., Gene-culture coevolutionary theory, *TREE* Vol 11 n°11 Nov 96 Elsevier, 1996

Galam S., Sous les chemises de la symétrie, *Pour la Science*, 1998

Gans E., Mimetic Paradox and the Event of Human Origin, *Anthropoetics* - The Electronic Journal of Generative Anthropology Volume I, number 2, 1995

Gintis H., *Game Theory Evolving*, Princeton Univ. Press, 2003

Girard R., *Mensonge romantique et Vérité romanesque*, Girard R., Grasset, 1961

Girard R., *Des choses cachés depuis la fondation du monde*, Girard R., Ed. Grasset, 1978

Hauer C., Fundamental clusters in spatial 2x2 games, *Proc. R. Soc. London,* B 268, 761-769, 2001

Hammerstein P., Genetic and Cultural Evolution of Cooperation, *Dahlem workshop report*, Hammerstein (ed.), 2002

Henrich J., Boyd R., The Evolution of Conformist Transmission and the Emergence of Between-Group Differences, *Evolution and Human Behavior* 19: 215–241, 1998

Henrich J. & al., 'Economic Man' in Cross-cultural Perspective: Behavioral Experiments in 15 Small-scale Societies, *American Economic Review*, 91:73-78, 2001
Hofbauer J., Sigmund K., *The Theory of Evolution and Dynamical Systems*, London Mathematical Society, Cambridge University Press, 1988

Henrich J. and Gil-White F. J., The Evolution of Prestige, *Evolution and Human Behavior* 22:165-196, 2001

Janssen M., Jager W., An integrated approach to simulating behavioural processes: A case study of the lock-in of consumption patterns, *Journal of Artificial Societies and Social Simulation* vol. 2, no. 2, <http://www.soc.surrey.ac.uk/JASSS/2/2/2.html>, 1999

Kaniovski Y. M., Kryazhimskii A. V. and Young H. P., Adaptive Dynamics in Games Played by Heterogeneous Populations, *Games and Economic Behavior* 31, 50-96, 2000

Lindgren K., Evolutionary Dynamics in Game-Theoretic Models, "Talk presented at the workshop "The economy as an evolving complex system II"", Santa Fe Institute", 1995

Nakamaru N., Matsuda H. & Iwasa Y., The Evolution of Cooperation in a Lattice-Structured Population, *Journal of Theoretical Biology*, 184, 65-81, 1997

Nowak, M. A., May R. M., The Spatial Dilemmas of Evolution, *Int. J. of Bifurcation and Chaos*, Vol.3, n°1, 35-78, 1993

Nowak M. A., May, R.M., Evolutionary Games and Spatial Chaos, Nowak M. A., May, R.M., *Nature*, 359: (6398) 826-829, 1992

Nowak, M. A., Bonhoeffer, S. & R. M. May, More Spatial Games, *Int. J. of Bifurcation and Chaos*, Vol.4, n°1, 35-56, 1994

Orléan A., The evolution of imitation, In Patrick Cohendet, Patrick Llerena, Hubert Stahn et Gisèle Umbhauer (eds.), *The Economics of Networks*, Berlin, Heidelberg et New-York, Springer-Verlag, Chap 15, 325-339, 1998

Ostrom E., Gardner R. & Walker J., *Rules, Games, and Common-Pool Resources*, University of Michigan Press, 1994

Rendell L., Whitehead H., Culture in whales and dolphins, *Behavioral and Brain Sciences* 24 (2):(2001), 2001

Smith J.D., The Comparative Psychology of Uncertainty Monitoring and Metacognition, *Behavioral and Brain Sciences*, Volume 26, Issue 3, 261-283, 2003

Sperber D., Metarepresentaitons in an evolutionry perspective, *Metarepresentations: A Multidisciplinary Perspective*. Oxford UniversityPress, 2000, pp.117-137, 2000

Tomasello, M., *The Cultural Origins of Human Cognition*., Harvard University Press., 2000

Weibull J. W., Evolutionary Game Theory MIT press, 1995

Wilson E.O., *Sociobiology: the new synthesis*., Harward University Press, Cambridge, 1975

Young H. P., *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton Univ. Press., 2001

Zelazo P. D., Frye D. and Rapus T., Age Related Dissociation Between Knowing Rules and Using Them, *Cognitive Development*, 11, 37-63, 1996