HAL open science

Boundedly Rational Agents Playing the Social Actors Game - How to reach cooperation

Christophe Sibertin-Blanc, Joseph El Gemayel

To cite this version:

Christophe Sibertin-Blanc, Joseph El Gemayel. Boundedly Rational Agents Playing the Social Actors Game - How to reach cooperation. Web Intelligence - WI and Intelligent Agent Technologies - IAT 2013, Nov 2013, Atlanta, United States. pp. 375-382. hal-01147310

HAL Id: hal-01147310

https://hal.science/hal-01147310v1

Submitted on 30 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 12980

To link to this article : DOI :10.1109/WI-IAT.2013.135
URL : http://dx.doi.org/10.1109/WI-IAT.2013.135

To cite this version : Sibertin-Blanc, Christophe and El-Gemayel, Joseph *Boundedly Rational Agents Playing the Social Actors Game - How to reach cooperation*. (2013) In: Web Intelligence - WI and Intelligent Agent Technologies - IAT 2013, 17 November 2013 - 20 November 2013 (Atlanta, United States).

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Boundedly Rational Agents Playing the Social Actors Game

How to reach cooperation

SIBERTIN-BLANC Christophe, EL GEMAYEL Joseph
IRIT, University of Toulouse
Toulouse, France
sibertin@ut-capitole.fr, joseph.elgemayel@gmail.com

***Abstract*—In any organizational setting, each of the participating actors adopts a quite steady behavior with respect to others. This behavior is not always in line with what might be expected given the role of the actor, although it is quite cooperative and most often contributes significantly to the proper operation of the organization. Within the framework of the SocLab approach for the modeling and analysis of organizations, we propose a model of the bounded rationality that social actors could implement in the process yielding the joint adaptation of their behaviors.**

Keywords*— *organizational behavior; cooperation; actors game; regulation; bounded rationality; organization; reinforcement learning; SocLab.

I. WHY AND HOW TO COOPERATE?

This article considers organizations, and more generally social *systems of organized action*, viewed as social constructs:

- established for some purpose, and thus aimed at achieving some goal(s),
- including individuals and resources,
- provided with rules about the handling of the resources by the individuals, intended to serve the achievement of the organization's goals,

where each of these elements is more or less precisely determined and recognized. This encompasses well-structured organizations (firms, institutions, ...), collaborative contexts (citizens associations, family, ...) as well as diffuse systems of collective action (public policy arena, partnered relationships, ...).

In such contexts, the members are found to exhibit a more or less steady behavior with respect to each other, and the occurrence of this well-established phenomenon – the regulation – is necessary for the existence of the organization: it allows each one to anticipate the behavior of others and so ensures the co-ordination required for the achievement of the organization's aims, and so preserves its raison d'être.

According to the Sociology of Organized Action, also called Strategic Analysis [2], this regularity originates from the strategic nature of actors' behaviors: each one has some goals and his behavior is driven by his believes on the best way to achieve these goals. As the goals of an actor embeds personal aims into his organizational role(s), the latter do not fully explain his behavior, more so as the concrete application of any organizational rule involves a contextual interpretation, which is at the discretion of actors. Therefore, the behaviors of social actors most often deviate from the rules while being, as we shall see later, generally much oriented toward cooperation.

Viewing the regulation of an organization as a phenomenon that emerges from interactions between individual processes, the concern of this article is how actors succeed in the stabilization of their respective behaviors and why they stabilize as the do. So, it proposes an algorithm intended to model the actors' rationality while accounting the facts that they must jointly adapt their behaviors to the ones of all others, they are greatly unaware of the objectives and means of others, and most social behaviors are not deliberate. From the organization point of view, it proposes a model of the regulation process as resulting from the interplay between the actors' rationality.

This question is addressed within the framework of the SocLab approach for the analysis of social relationships within organizational settings, which is extensively presented in a companion paper [1]. This approach is supported by a software platform that allows the user to edit models of organizations, to study the properties of models with analytic tools, and to compute by simulation the behaviors that the members of an organization could adopt the ones with respect to others [24]. Focusing on the social dimension of organizations, the simulation algorithm considers rather the conduct of an actor *vis-a-vis* others, his attitude which is assessed in terms of *cooperativity,* than his substantive contribution to the realization of the organization's goals. As far as one agrees with its theoretical fundaments, the SocLab platform looks like a tool for organizational diagnoses and the analysis of scenarios regarding evolutions of organizations. It can be used also for the design and analysis of virtual organizations having no direct reference to reality, intended to the study of theoretical properties of organizational configurations featuring particular characteristics.

The purpose of this article is to present the SocLab simulation algorithm. Section II outlines the SocLab modeling of the structure of organizations that leads to consider the interactions among actors as a specific game we call the *social actors game*. Two models of simple organizations are also presented. Section III states requirements for a simulation algorithm regarding the expected results (the emergence macro-level view) and the actors' competences (the micro-level

view). Section IV presents the actors' bounded rationality algorithm while providing, as much as possible, its social science underpins, while section V gives an overview of results. The last section discusses the originality of the algorithm and related works.

II. THE SOCLAB MODEL OF SOCIAL ORGANIZATIONS

The SocLab framework aims at formalizing and slightly extends the Sociology of Organized Action [2]. For space limitation we just outline the SocLab model of the structure of organizations, especially the elements that found the social actors game. The reader will find in [1] a comprehensive presentation and discussion of this model and how to model and study a concrete organization.

A. The Structure of Organizations

The core of this meta-model is represented graphically as a UML Class Diagram in figure 1. Accordingly, the structure of an organization is constituted of a set of *Actors* and a set of social *Relations* linked by the *Control* and *Depend* associations.

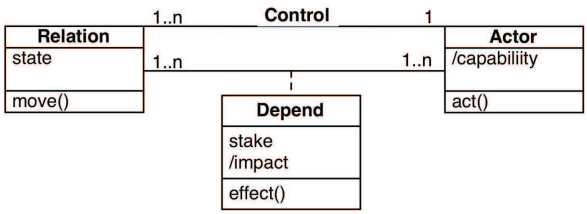


Figure 1. The meta-model of the structure of organizations

A relation is the matter of the recurrent interactions between an actor who controls the availability of a specific resource and actors who need this resource to achieve their goals. *Resources,* as the media for interactions among actors, must be understood in a very general sense as something useful or necessary for actors' activities: material objects, factual, procedural or axiological believes and knowledge, or even personal attitudes. Considering the concept of *role*, a relation may also be viewed as the playing of a role, be it defined by the organization or caught by the actor.

The *state* of a relation corresponds to the controller actor's policy regarding the management of the resource; its range is the *Space of behaviors* of the relation and its value measures to what extent the actor is cooperative (positive values) or uncooperative (negative values). When an actor *acts()*, he *moves()* the states of the relations he controls toward more or less cooperation, and he is the only one to be able to do so.

An actor depends on the relations regarding the resources he needs to achieve his goals. How much an actor depends on a relation is determined by the necessity of the resource for the achievement of his goals and the relative importance of these goals. This level of dependency is represented by the *stake* attribute of the depend association. Each actor has 10 points of stakes that he distributes on the relations he depends on, including the ones he controls. For each actor dependent on a relation, there is an *effect()* function that, depending on the relation's state (*i.e.* the controller actor behavior), determines how well the actor can make use of the underlying resource. Negative effects correspond to hampering or preventing the use of the resource while positive effects correspond easing its use. The *impact* of a relation upon an actor is the value of the *effect()* function applied to the state of the relation weighted by the actor's stake.

The aggregation of the impacts that an actor receives from the relations he depends on determines his *capability*, or action capacity. The capability measures how much the actor has the means to achieve his goals, so that to obtain a high level of capability constitutes a *meta-goal* that drives his behavior. If impacts are aggregated by the sum operator, when the organization is in the state $s = (s_{r1}, \dots, s_{rm})$, where $R = \{r_1, \dots, r_m\}$ is the set of relations, the capability of actor a is defined as

$$capability(a, s) = \sum_{r \in R} stake(a, r) * effect_r(a, s_r)$$
$$= \sum_{r \in R} impact(r, a, s_r).$$

The aggregation (once again by the sum operator) of the actors' capabilities defines the *global capability* of the whole organization as

$$capability(s) = \sum_{a \in A} capability(a, s)$$

The model of an organization includes others elements such as constraints between relations or solidarities between actors, allowing to deal with essential dimensions of organizations [1].

B. The Actors Game

Such an interaction setting defines an *actors game* where each actor plays against the whole of other actors of the organization. Let's define arbitrarily the range of the Space of behaviors of a relation r as $SB_r = [-10; 10]$, and the codomain of effect functions also as [-10; 10] (the ranges of value do not matter since only comparisons and proportions make sense).

At each step of the game, every actor has the possibility to move the values of the states of the relations he controls, and this change of the game's state modifies the capability of all actors. Let $(s_{r1}, \dots, s_{rm})$ be a state of the organization and $(c_{r1}, \dots, c_{rm})$ be moves such that $(c_r + s_r) \in SB_r$ and c_r is chosen by the actor who controls r. Once each actor has chosen such an action, the game goes to a new state defined by

$$Transition: [-10; 10]^m \times [-10; 10]^m \rightarrow [-10; 10]^m$$
$$(s_{r1}, \dots, s_{rm}), (c_{r1}, \dots, c_{rm}) \mapsto (s_{r1} + c_{r1}, \dots, s_{rm} + c_{rm})$$

The game ends when a stationary state is reached. In such a state, each actor no longer modifies his behavior because the level of capability he actually obtains satisfies him. Therefore, the organization is regulated and can operate in this way.

The actors game is a game in the sense of von Neumann and Morgenstern [3], where the capability of each actor serves as utility function. However, it differs from games that are considered in economics. The end of an actors game is not at all to find how to optimize the discounted accumulation of rewards, i.e. the production and distribution of wealth. Its end is to find a state providing each actor with an acceptable reward, i.e. a lasting mode of operation for the organization.

C. *The Trave-Tour case*

For space limitation, we present only simple examples of models of organizations (models of real organizations can be found in [1], [4] or [5]).

Travel-tour is a tour operator having two agencies, TRO1 and TRO2, located at Trouville. Agnes is a secretary attached to TRO1 agency. She has a temporary employment and works part-time at TRO2, so that she has to move between the two agencies. These last months, the results of the TRO1 agency increase, while the ones of TRO2 agency stay stable, or even decrease. To foster TRO1 agency, the Regional Executive proposes to regularize the situation of Agnes with a firm employment and to affect her exclusively to TRO1 agency. Accordingly, Agnes will have a firm contract and will be relieved to split her work in two parts, while the TRO1 Director will have a full-time secretary at his disposal.

However, each one refuses the proposal and has very good reasons to do so. A deeper analysis reveals that TRO2 agency is more inventive in designing travel packages, while TRO1 agency has an efficient commercial staff. Being aware of the TRO2's activity, Agnes provides information to the Director, enabling TRO1 to skillfully finalize the TRO2's ideas. As for Agnes, for personal reasons she has not as short-time objective getting a steady job. Moreover, she is very cool in her working relations with the other employees of TRO1, and she greatly appreciates that none of the TRO1 and TRO2 directors has the possibility to exert a precise control on her work. Thus the change in the organization would increase the control of the director on Agnes's activities (what she does not wish) and the Director would loose his source about TRO2 (what he does not wish).

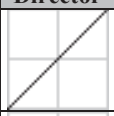
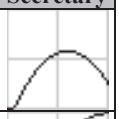
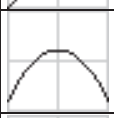
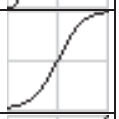
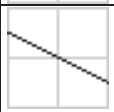
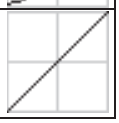
	Stakes		Effect functions	
	Director	**Secretary**	**Director**	**Secretary**
give information	7.0	**2.0**		
control the work	**2.0**	6.0		
renew the contract	**1.0**	2.0		

Figure 2. The stakes and the effect functions of the dependence association between relations (in line) and actors (in columns). The functions' x-axis corresponds to the relation's state (*i.e.* the cooperativity of its controller actor), the y-axis to the effect on the actor of the relation being in this state.

Figure 2 shows a model of the structure of this organization for highlighting the reasons of refusals. It includes only the two actors with three relations, which would be drastically changed by the Regional Executive's proposal: the information given by the secretary (vanished), the control of the work by the Director (enhanced) and the contract's renewal (vanished). We only comment the shape of the effect functions of the 'give information' relation. Positive values of the state of this relation correspond to a high quantity and quality of information delivered by the Secretary, negative value to disinformation and null values to silence. For the Director, he would take full advantage of positive information and be greatly disturbed by false ones. For the Secretary, the best is to talk about her work in the TRO2 agency in casual way (state = 2). Giving more information makes her uncomfortable and she risks to be detected, and giving false information even more. Details about this model may be found in [6].

Table I shows, for some noteworthy states of the organization, the state of relations and the capability of actors. This organization shows the typical features of a Prisoners' Dilemma: for each one, the better is to trust the other and to renounce to his self-reward in order to get the more valuable collaboration from the other.

TABLE I. IN COLUMNS, THE STATES CORRESPONDING TO: THE MAXIMUM CAPABILITY OF THE WHOLE ORGANIZATION AND OF EACH ACTOR; THE WORST STATE; SIMULATION RESULTS. IN ROWS, THE STATE OF RELATIONS AND THE CAPABILITY OF ACTORS.

	Maximum capability			Worst Global	Simulation results		
	Global	Director	Secretary		Value	Percentage	Deviation
give information	10.0	10.0	2.0	-10.0	10.0		0.0
control the work	7.0	0.0	10.0	-10.0	10.0		0.0
renew contract	10.0	-10.0	10.0	-10.0	3.33		8.89
Director	59.2	**79.0**	-7.0	-81.0	52.3	84.3%	4.44
Secretary	65.1	-28.0	**82.6**	-97.8	56.5	85.5%	17.78
GLOBAL	**124**	51.0	75.6	**-178**	108.8	94.9%	

D. *The Free-rider case*

This artificial model consists of four actors, who control one relation each. A1 controls R1 and depends on the three relations (R2, R3, R4) controlled by the other actors (respectively A2, A3, A4) who, in turn, depend on R1. Figure. 3 shows the amount of stakes and the shape of effect functions of this model: A1 confronts the three actors A2, A3, and A4, who are independent of each other. This structure leads naturally one of the 3 actors to take advantage of the cooperation of 2 others by defecting in order to maximize his capability.

Table II shows the capability of actors for remarkable configurations of the system. Eight configurations correspond to Pareto optima while the last configuration (C9) corresponds to a Nash equilibrium (all relations in state -10) that turns out minimize the global capability of the system. The maximum capability of the system (C1) is achieved by the cooperation of every actor (all relations in state +10), at the expense of the maximum that each actor could pretend (C8, C2, C3, C4 respectively). Note that the maximum of A1's capability (C8) requires the cooperation of the three other actors without being rewarded by A1. The configurations C5, C6 and C7 correspond

to states that maximize the capability of two actors and strongly penalize A1.

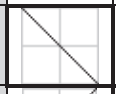
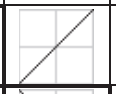
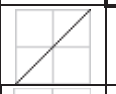
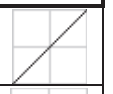
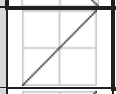
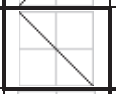
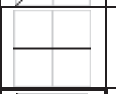
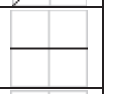
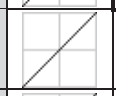
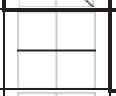
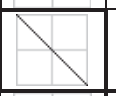
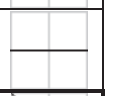
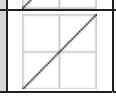
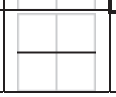
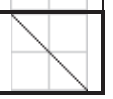
	R \ A	A1	A2	A3	A4
Stakes	R1	**1**	9	9	9
	R2	3	**1**	0	0
	R3	3	0	**1**	0
	R4	3	0	0	**1**
Effect functions	R1				
	R2				
	R3				
	R4				

Figure 3. The stakes and effect functions of the dependence association between relations (in rows) and actors (in columns). The functions' x-axis corresponds to the relation's state and the y-axis to the effect on the actor of the relation being in this state.

TABLE II. ACTORS' ACPABILITIES (IN COLUMNS) FOR NINE STATES (IN ROWS) OF THE FREE-RIDER ORGANIZATION: THE STATE THAT MAXIMIZES THE GLOBAL CAPABILITY (C1); THE FOUR STATES THAT MAXIMIZE THE CAPABILITY OF EACH ACTOR: A1 (C8), A2 (C2), A3 (C3), AND A4 (C4); THREE STATES WHERE TWO ACTORS DO NOT COOPERATE: A2 AND A3 (C5), A2 AND A4 (C6), A3 AND A4 (C7); THE NASH EQUILIBRIUM (C9).

		Actors' Capability					Occurrence percentage in simulation results
		A1	A2	A3	A4	Global	
Configurations	C1	80	80	80	80	**320**	8 %
	C2	20	**100**	80	80	280	30 %
	C3	20	80	**100**	80	280	25.3 %
	C4	20	80	80	**100**	280	34 %
	C5	-40	**100**	**100**	80	240	0.3 %
	C6	-40	**100**	80	**100**	240	1.7 %
	C7	-40	80	**100**	**100**	240	0.7 %
	C8	**100**	-100	-100	-100	-200	0 %
	C9	-80	-80	-80	-80	**-320**	0 %

III. WHAT A RATIONALITY FOR ACTORS GAME PLAYERS?

To compute how the actors of an organization are likely to behave in respect with others, we just have to provide them with a rationality allowing to play the social actors game. We seek a model of the way a social actor determines his behavior within an organizational setting that is funded in social sciences. It is required on the one hand to yield the observed facts when all the actors of an organization jointly behave in this way (macro level) and, on the other hand, to be plausible regarding the point of view of individuals (micro level). The intended emergent outcomes of a simulation process where all actors apply this process are the following:

- The simulation ends in a stationary state after a reasonable number of steps. However, simulations may last long if it is difficult for the actors to find how to cooperate – e.g. consider a circular organization with 5 actors where each one depends on his predecessor –, or if no cooperation is beneficial – e.g. a null-sum game (but real organizations are not constant-sum games). Terminal states are equilibria, not in the Nash sense where no one has advantage to change his behavior, but in the social control sense where a change is quickly sanctioned by others. As the difficulty to find a state that features a given property we mean a low frequency of this property in the space of all the states of the organization.
- At the final state, every actor has a quite high level of capability. This "high level" should be considered in proportion (i.e. as (value - minimumValue) / (maximumValue - minimumValue)) rather than in value, because actors have disparate ranges of capability. Moreover, most organizations include structural conflicts so that there is no state providing every actor with a high level of capability, a fortiori his maximum level of capability. The settlement of conflicts depends on the power of the resources of every actor to defend and put forward his own situation. Globally, we expect to obtain states near of Pareto optima while being equitable without excessive gaps between the capabilities of the actors.
- At the final state, the organization's global capability (the sum of all actors' capabilities) is quite high. However, if the discovery of such a state is very difficult, actors can be discouraged and satisfy with a low level of capability. Notice that almost all real organizations include conflicts among actors, so that the maximum global capability seldom coincides with the maximum capability of one of the actors. Any stationary state results from a compromise.

We do not know the mind process that drives the behavior adopted by a social actor. However, we intend to propose a model that is plausible regarding the cognitive and psycho-social points of view and thus features the following characteristics.

Social actors have goals (that are not necessarily limited to own selfish interests) and to be rational is to pursue the achievement of these goals. To get a high level of capability is obviously the first goal, in fact a meta-goal whose achievement provide the means needed to achieve concrete goals. This may be qualified as his *individual goal*. We assume that, when an actor takes part in an organization, he has some reasons to prefer to belong to that organization than the converse. Therefore, he is interested in the persistence of this organization that requires that it operates somehow well, and this may be qualified as his social goal. It is well known that an organization cannot operate when people just "work to rule". The proper operating of an organization requires that people implement positively the rules, that is to say they cooperate with each other. Thus, in most organizational settings, it is beneficial for actors to cooperate, and they do to the extent this is beneficial, both for their individual goal – in order to get the

reciprocal cooperation of others – and for their social goal – to ensure the continuation of the organization.

The rationality of the actor's decision process does not imply it is a conscious deliberative process, since even young babies adapt their behaviors to expectations of their entourage [7]. This search is also not an optimization process. It would be illusory since, as already noted, real social actors games have no socially feasible optimal state. Rationality must not be confused with optimization [8] and, from the cognitive point of view, there are many evidences that the Bayesian maximization of the subjective expected utility of Homo Economicus is not a realistic description of human decision-making, see among many others [9], [10], [11] and [12].

The sociology of organized action is in harmony with the Simon's paradigm of bounded rationality [13] that fully satisfies the organizational context we are considering: in addition to their cognitive limitations and the accounting of the cost of decision-making, actors have very few information about the actual state of affairs (due to the opacity that prevails in social relationships, we consider that actors are aware just of their own situation); the alternatives of the choice, i.e. the possible behaviors, are not given and their discovering is a part of the decision process; they are uncertain of the effect of their actions, because there are many incertitudes about the reactions of other actors. According to Simon's bounded rationality, an actor does not search for an optimal solution but just a satisfactory, or "satisfycing", alternative. To this end, his decision process driven by an aspiration level and he "*searches until an alternative is found that is satisfactory by the aspiration level criterion*". The value of this aspiration level is proper to each one and determined by his believes about the society within which he lives, his past individual experiences: *"the existence of a satisfactory alternative is made likely by dynamic mechanisms that adjust the aspiration level to reality on the basis of information about the environment*" [13].

Thus, our model of a boundedly rational social actor includes an ***ambition*** variable[1] that serves as a threshold for the actor to decide whether he is satisfied by the current state of the organization: an actor will be satisfied if and only if his capability exceeds his ambition. The ambition of an actor is initialized at his maximum capability and evolve according to the context; it is gradually reduced when the actual capability is below and upgraded when it is above. At the level of the whole game, ambition provides us with the convergence criterion for simulation of an actors game: a stationary state is reached when the capability of every actor exceeds his ambition.

At each step of the game, each actor perceives the new value of his capability and compares it with the value at the previous step. He can know whether it has increased or decreased and relate this change with his previous action. This is the only information that he gets at each step of the game.

As a more global assessment of his position, we assume that an actor is able to evaluate the ***gap*** between his ambition and his capability. We will see in the following section that the gap is an important driver of the actor's decision process. It is accounted in proportion, i.e. as (ambition - capability) / (ambition - minCapability), where minCapability is the absolute minimum capability of the actor. In this way, each actor reasons within his own scale of values. This will erase the disparities between the amplitudes (in value) of actors' ranges of capability and favor a smoother convergence of simulation processes.

In the following section, we detail our algorithm of the rationality of social actors outlined in tables III and IV. A preliminary version of this algorithm is presented in [5].

IV. A MODEL FOR THE RATIONALITY OF ACTORS GAME PLAYERS

The knowledge about the most suitable behavior to adopt in an organizational setting is acquired by the participants, as shown by the way a newcomer in a group progressively adjusts his behavior. Most usual knowledge are acquired by experience and the paradigm of self-learning by trial and error [15], which is supported by very firm neurobiological foundations [16], fits our requirements. The learning can produce the necessary knowledge for case-based reasoning [17], by the construction of a base of rules of the kind (situation, action, strength) where:

- *situation* is the list of the impacts got by the actor from the relations he depends on, as they were at the step of the creation of the rule;
- *action* is a list of changes in the states of the relations controlled by the actor, chosen at random at the creation of the rule;
- *strength* is a numerical evaluation of the suitability of the rule; it is initialized at 0 and reinforced at each application of the rule depending on the resulting increase or decrease of the actor's capability.

There is a single rule base for all the relations controlled by the actor, instead of a base for each relation, because the behavior of an actor is a coherent whole. His behavior vis-à-vis any of the relations he controls is not independent of his behavior towards others.

At any step of a simulation, a rule is applicable if it has a positive strength and its situation is close, according to an Euclidean distance, to the current situation of the actor. The closeness of two situations is defined by an individual parameter, called the ***discriminality,*** that allows to account for a cognitive trait of each actor and determines his ability to discriminate between situations. It ranges from 1 to 5: 1 makes rules to be applicable in any situation; 5 allows the actor to distinguish between very bad/bad/neutral/good/very good situations. If there is no applicable rule, a new one is generated; else one of the three rules having the higher strength is selected. Rules having a negative strength are forgotten.

The application of a rule consists in shifting the state of each relation of the value registered for that relation in the rule's action. If the actor controls relations r1, …, rk being in states s_{r1}, … s_{rk} and the rule's action is (c_{r1}, .., c_{rk}), the relation will become in states $s_{r1} + c_{r1}$, … $s_{rk} + c_{rk}$. This definition of actors' actions assumes that a new behavior always results of a

[1] According to Durkheim for example, human beings are really able to evaluate their ambition: "*Chacun, dans sa sphère, se rend vaguement compte du point extrême jusqu'où peuvent aller ses ambitions et n'aspire à rien au-delà"* [14, p. 23].

motion relative to the previous behavior that is chosen with regard for (the evolution of) the current situation. The state change of a relation may be positive (toward a more cooperative behavior), negative (toward less cooperativity) or neutral. The size of a change is selected at random within a 0-centered interval of bounds +/- action_range discussed below.

The spaces of behaviors of relations are bounded because whatever behavior is not feasible so that the ranges of actors' behaviors are limited. As any natural process, the actors game includes initial and limit conditions. Instead of an intricate and pointless management of rules' actions, we may distinguish between the performance of an action (the action component in the selected rule) and its actual effect (the change in the relations' states). If a rule has been positively reinforced enough so that its application leads the state of a relation to cross one of its bound, one may just set the state at the bound:

$s_r' = s_r + c_r$;
if($10<s_r'$) then $s_r'=10$;if($s_r'<-10$) then $s_r'=-10$.

This amounts to keep the rule because it has been beneficial – and possibly continues to do so on other relations – but to consider only its direction, not its intensity.

When an actor applies a rule `r` at time `t`, the variation of his capability at time `t+1` records the effect of the rule on the relations he controls. On the relations controlled by other actors, the effect of the rule on the variation of the actor's capability appears only at time `t+2`, once others have reacted to the effect of `r` on their own capability. Thus the actor's capability variation at time `t` must be assigned in proportion of the actor's ***autonomy*** at the strength of the rule applied at time `t-1` and the remaining part at the strength of the rule applied at time `t-2`. The autonomy of an actor corresponds to the proportion of his (potential) capability that depends on himself. It may be defined as the proportion of his stakes on the relations he controls. So, in the Travel-tour case, the autonomy of the Director is 0.3 and the one of the Secretary is 0.2.

An essential parameter of any learning is the tradeoff between the exploration of the environment and the exploitation of already acquired knowledge [18]. If the exploration is excessive, the process difficultly reaches a stable state because still another alternative is tested; if the exploitation is excessive, the process retains the first acceptable solution and possibly ignores better solutions. So, each actor manages an ***ExpR*** variable whose value 1 corresponds to full exploration and value 0 to full exploitation.

ExpR decreases (from its initial value 0.9 to 0.1) according to the gap between the actor's ambition and capability, so that more the actor's capability is far below his ambition, more ExpR is high and he searches for something else. The gap being evaluated in proportion, as explained above, actors having a wider range of capabilities do not necessarily need more time steps for making their ambition to join their capability. This improves the possibility that actors have similar values for their respective ExpR variables, which favors a smooth convergence of simulations, since an intemperate exploration by an actor can destabilize the learning dynamics of other actors.

TABLE III. THE GLOBAL LOOP OF THE REGULATION PROCESS.

```
                    // Initialization
foreach relation r:
    s_r = 0  // or any other value
foreach actor a:
            // compute structural constants
    autonomy = (∑_{r ∈ R / controls (r) = =a} stake (a,r)) / (∑_{r ∈ R} stake (a,r))
    minCapability = min_{s ∈ S} (capability (a, s))
    closenessThreshold, slope, …
            // initialize learning variables
    capability_0 = ∑_{r ∈ R} impact(r, a, s_r)
    ambition_0 = maxCapability = max_{s ∈ S} (capability (a, s))
    ExpR_0 = 0.9

repeat              // The Simulation Loop
    foreach actor a:                    //all actors see the same world
            a.action = a.selectAction ( )
    foreach actor a:                    //they don't act in turn
            performAction (a.action)
Until (foreach actor a: a.ambition_t ≤ a.capability_t)
```

TABLE IV. THE SELECTACTION() FUNCTION OF THE ACTORS' LEARNING ALGORITHM.

```
            // Perception of capability and updating gap
capability_t = ∑_{r ∈ R} impact(r, a, s_r)
deltaCapability = capability_t - capability_{t-1}
gap = (ambition_{t-1} - capability_t) / (ambition_{t-1} - minCapability)
            // Updating ambition
if (ambition_{t-1} > capability_t)
    if ((gap < (11 – tenacity) / 10) && (deltaCapability = = 0))
            ambition_t = ambition_{t-1} – ((1 – ExpR_{t-1}) *
                                (ambition_{t-1} – capability_t + 1)) / 100
    else    ambition_t = ambition_{t-1} – ((1 – ExpR_{t-1}) * gap / 50
else        //a is already satisfied
    ambition_t = ambition_{t-1} + ((capability_t – ambition_{t-1}) / 100)
            // Updating exploration rate
ExpRIns = 0.1 + (0.8 / (1 + e^{slope * (gap - abscissa)}))
ExpR_t = ExpR_{t-1} * ExpR_{t-1} + (1 – ExpR_{t-1}) * ExpRIns
            // Updating action_range
action_range = 2 * ExpR_t
            // Updating strength of last and penultimate selected rules
            // SR_t stands for the Rule Selected at time t
SR_{t-1}.strength_t = (1 - ExpR_t) * SR_{t-1}.strength_{t-1}
                            + ExpR_t * autonomy *deltaCapability
SR_{t-2}.strength_t = SR_{t-2}.strength_{t-1}
                    + ExpR_t * (1 - autonomy) * deltaCapability
            // Forgetting bad rules
if (SR_{t-2}.strength_t < 0) RuleBase.remove(SR_{t-2})
            // Selecting the set M of rules applicable at time t
    M.clear ( )
    foreach rule R in RuleBase
     if (distance(R.situation, CurrentSituation < closenessThreshold)
                    M.add(R)
            // Selecting an action
if (M.isEmpty())
    SR_t = (CurrentSituation(), (atRandom( )*action_range), 0)
    RuleBase.add (SR_t)
else
    SR_t = ChooseOneOfThreeRulesWithMaximumStrength (M)
return (SR_t.action)
```

ExpR also varies depending of an actor's individual parameter, his ***tenacity,*** that determines to what extent he is prone to explore or to exploit. This parameter enables to account for a psychological trait of individuals. It also enables to account for the involvement level of the actor into the organization. More his participation to the organization is important for him, more his tenacity should be high. The value of ExpR follows the curves given in Figure 4, with the application of a learning-rate to regularize the variation.

The *learning-rate* of a learning process, usually denoted α, determines how much new information is integrated in the already acquired knowledge. If V is a learned variable which is updated at each time step according to information I acquired at this step, we have $V_{t+1}:=(1-\alpha).V_t + \alpha.I_{t+1}$. A high learning-rate corresponds to an important acquisition of information that just is sought by a high level of exploration. Hence, our algorithm attributes the following effects to high values of the ExpR variable:

- The *action_range* variable, that determines the size of the relations' shifts in the action of new rules, is larger so that the changes in the actor's behavior are greater.
- A high learning-rate for the strength of rules and the ExpR variable itself.
- The diminution of the ambition variable is smaller, so that the actor does not renounce to explore the potentialities of the game. However, if the gap is very small and the capability does not vary, one assumes that the actor is near to become satisfied and a limit condition is applied.

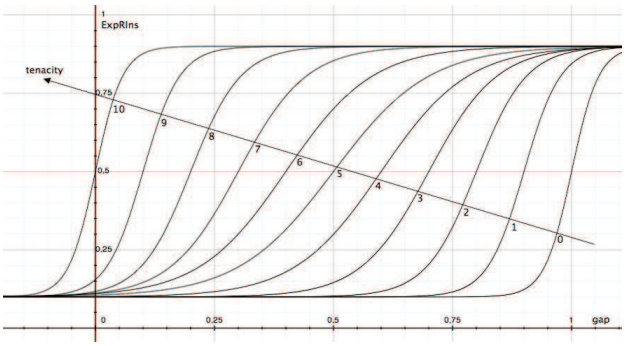


Figure 4. For each value of the tenacity from 0 to 10, the curve for the update of ExpR depending on the actor's gap (see the equations in TABLE IV). With tenacity = 10, the actor explores (ExpR > 0.5) until he his satisfied (gap ≤ 0); with tenacity ≤ 3, the actor exploits as soon as his gap is below 0.3.

V. SIMULATION RESULTS

The reliability of such a simulation algorithm relies first on its foundation in social science and secondly on the correctness of the results when it is applied to organizations which are simple enough for exhibiting a well-understood functioning. Then, one may trust the knowledge brought by the simulation results for poorly understood organizations. For space limitation we cannot show to what extent simulations results comply with the requirements stated in section III. We just briefly comment results about the simple models presented in section II.

Regarding the Travel-Tour case, the three right hand columns in Table I give the averaged result of 300 simulations, with tenacity 4 for the Director and 6 for the Secretary (she is a bit more involved in the game), and discriminality 1 for both. All simulations have converged after 2140 steps in average (minimum 412 and maximum 18184). The deviation of the state of "renew the contract", which causes the deviation of the capability of both actors (specially the secretary), is high because the feedback is small so the director difficultly evaluates the impact of his actions: his stake and the effect function's amplitude are low. The global capacity is fulfilled at 94% of its maximum value, say about 85% of both actors' maximum capabilities. More details about the results are given in [6].

Regarding the Free-Rider case, the results concern 300 simulations where the tenacity is 5 and the discriminality is 1 for all actors and the number of steps 3680 in average. All simulations have exactly converged toward one of the C1 to C7 configurations in the proportion given in the right hand column of Table II. Regarding concrete social organizations, the convergence of simulations toward distinct modes is not uncommon: it indicates that the organization could operate in different ways, even if only one of them actually occurs.

Figure 5 shows a sensitivity analysis of the actors' capabilities depending on the tenacity of A1. To the detriment of longer simulations up to 50000 steps, A1 succeeds in constraining other actors to cooperate, so that all simulations converge toward the C1 configuration. However, in some organizations, a too high level of the tenacity of one (or several) actor prevents the convergence because there is no means for the tenacious actor to satisfy his steadily high ambition level. In most models, discriminality is not a sensitive parameter.

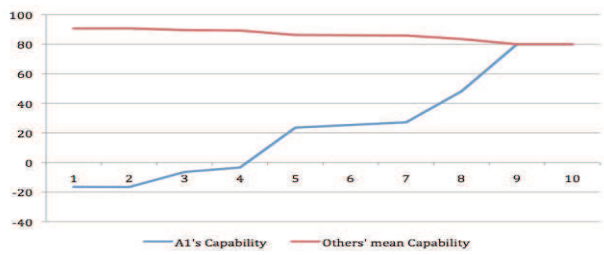


Figure 5. The variation of actors' capabilities depending on the A1's tenacity.

VI. DISCUSSION

When people are engaged in a system of collective action, their normal behavior is to cooperate: working to rule and hindering are commonly assessed as reprehensible behaviors. This article reproduces this empirical observation and shows, by a plausible model of how human beings adapt their behaviors the one another within organizational settings, that cooperation occurs even among actors who are essentially self-interested, have few information and use limited cognitive resources. However, the structure of the organizational

relationships and the individual traits make that they are more or less inclined to cooperate and that finding how to cooperate can be difficult. The basic condition is to take the risk to cooperate and assume that others will reciprocally cooperate, and to have a realistic ambition. Then, actors will cooperate more so as the structure of the organization makes it beneficial – i.e. the gap between the maximum and the minimum global capability is large – and quite easy – i.e. states supplying high capabilities are not too sparse.

Many variants of this algorithm may be studied. For example, one may consider that the actor's goals are incommensurable. Then, the actor seeks to improve every component of his situation, more or less independently from others, instead of their aggregation in a mere sum. We have found no significant enhancement of the results by this way. The same holds for a more sophisticated management of the rule base, for instance the aging of rules to forget those which have not been activated since a long time, or the recording of bad rules to not generate them again. On the other side, providing actors with more information such as the possibility to compare their capability with the one of others or worrying about the global capability of the organization opens new perspectives for dealing with the feelings and emotions of actors.

At first glance, playing the actors game seems to fall in the domain of Multi-Agent Reinforcement Learning (MARL): the actors have to jointly find a best solution for the game (see *e.g.* [19, 20] for comprehensive surveys). However, the actors game differs from the games that are considered in this field of research. The agents of an actors game do not have to build a strategy, *i.e.* what to do in each case, they have to find a state of the game. In other words, they do not try to maximize the discounted sum of their rewards, but to find a behavior, *i.e.* the state of the relations they control, that will ensure them a satisfying reward. With a null discount factor, the Bellman's equation becomes trivial and the assumptions of the Q-Learning approach, which is the underpinning of most MARL algorithms, are not fulfilled.

In most algorithms for distributed problem solving, the agents either exchange information on their respective state or action in the course of the search or they have in advance knowledge about properties of the environment. Here, we only assume that an actor knows his level of autonomy and his maximum and minimum levels of capability, so that he has no means to build a model of another actor or of the whole game. As for algorithm related to bounded rationality [11], they are mainly heuristics to supplement the lack of information or of cognitive capacity or they address very different questions such as classification or form recognition, see e.g. [21, 22, 23]. The aspiration adaptation theory of Selten [9] copes with the Simon's aspiration level of boundedly rational agents, but they are assumed to be endowed with much more cognitive capacities and information about the consequences of their choices.

ACKNOWLEDGMENT

This work has been partially funded by the EMOTES project of the French ANR (ANR 11 EMCO 004 03).

REFERENCES

[1] C. Sibertin-Blanc, P. Roggero, F. Adreit, B. Baldet, P. Chapron, J. El Gemayel, M. Mailliard, and S. Sandri, "SocLab: A Framework for the Modeling, Simulation and Analysis of Power in Social Organizations", Journal of Artificial Societies and Social Simulation (JASSS), In Press [16(4)?], 2013. http://jasss.soc.surrey.ac.uk/

[2] M. Czozier and E. Friedberg, Actors and Systems: The Politics of Collective Action. The University of Chicago Press, 1980. [L'acteur et le système. Contraintes de l'action collective, Paris, Seuil, 1977].

[3] O. Morgenstern and J. Von Neumann, The Theory of Games and Economic Behavior. 3rd ed., Princeton University Press, 1953.

[4] F. Adreit, P. Roggero, C. Sibertin-Blanc, C. Vautier, "Using SocLab for a Rigorous Assessment of the Social Feasability of Agricultural Policies", International Journal of Agricultural and Environmental Information Systems, 2 (2), june 2011.

[5] J. El-Gemayel, C. Sibertin-Blanc, P. Chapron, "Impact of the tenacity upon the Behaviors of Social Actors". Advances in Practical Multi-Agent Systems, Quan Bai and Naoki Fukuta (Eds) pp. 287-306, Studies in Computational Intelligence 325, Springer, oct 2011.

[6] C. Sibertin-Blanc, F. Adreit, "The Travel-tour case study". CoMSES Computational Model Library, 2013. Retrieved from: http://www.openabm.org/model/3792/version/1

[7] B. Agnetta, P. Rochat, "Imitative games by 9-, 14-, and 18-month-old infants." Infancy, 6 (1), pp. 1-36, 2004.

[8] R. Boudon, "Limitations of Rational Choice Theory". American Journal of Sociology, Vol. 104, No. 3 pp. 817-828, 1988.

[9] R. Selten, "Aspiration adaptation theory", Journal of Mathemacitcal Phychology, 42, pp. 191-214, 1998.

[10] D. Kahnemann, P. Slovic, and A. Tversky, Judgement under Uncertainty, Heuristics and Biases, Cambridge University Press, 1982.

[11] A. Rubinstein. Modeling Bounded Rationality. MIT Press, 1998.

[12] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath, "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", The American Economic Review 91(2) 2001.

[13] H. A. Simon, "Theories of bounded rationality", In Decision and organization, C.B. McGuire & Roy Radner (Eds), North-Holland, 1972.

[14] E. Durkheim, Le Suicide, Etude de sociologie. Livre Deuxième : Causes sociales et types sociaux, chap. 5, II, 1897.

[15] R.S. Sutton and A.G. Barto, Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998.

[16] M. Domjan, The Principles of Learning and Behavior, 6th Edition. Wadsworth Publishing Company, Belmont, CA, 2009.

[17] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", AI Communications, Vol. 7 Nr. 1, pp 39-59, March 1994.

[18] J. March, "Exploration and exploitation in organizational Learning", Organization Science, vol. 2, Nr 1, pp. 71-87, February 1991.

[19] Y. Shoham, R. Powers, and T. Grenager, "Multi-agent reinforcement learning: A critical survey", Comput. Sci. Dept., Stanford Univ., Stanford, CA, Tech. Rep. [Online], 2003. Available: http://multiagent.stanford.edu/papers/MALearning_ACriticalSurvey_2003_0516.pdf

[20] L. Busoniu, R. Babuska, and B. De Schutter, "A Comprehensive Survey of Multiagent Reinforcement Learning". IEEE Transactions on systems, man, and cybernetics – part C, vol. 38, no. 2, 2008.

[21] P. Jehiel, "Learning to play limited forecast equilibria", Games and Economic Behavior, vol. 22, pp. 274-298, 1998.

[22] A. Dal Forno and U. Merlone, "A multi-agent simulation platform for modeling perfectly rational and bounded-rational agents in organizations", J. of Artificial Societies and Social Simulation, 5 (2), 2002.

[23] S. Schuster, "An algorithm for the Simulation of Bounded Rational Agents", MPRA Paper 15942, Univ Library of Munich, Germany, 2009.

[24] http://soclabproject.wordpress.com.