Simulating the relative effects of punishment and sanction in the achievement of cooperation

Giulia Andrighetto[1,2], Daniel Villatoro[3], Rosaria Conte[1], Jordi Sabater-Mir[3]

[1] LABSS, Institute of Cognitive Science and Technologies (ISTC), CNR
via S. Martino della Battaglia 44, 00185 Rome, Italy
[2] European University Institute (EUI)
Florence, Italy
[3] Artificial Intelligence Research Institute (IIIA)
Spanish Scientific Research Council (CSIC)
Bellatera, Barcelona, Spain

Abstract. As specified by Axelrod in his seminal work "An Evolutionary Approach to Norms", punishment is a key mechanism in a self-regulated society to achieve the necessary social control and to impose certain norms. In this paper, we distinguish between punishment and sanction, focusing on the specific ways in which these two different mechanisms favor the emergence of cooperation and the spreading of social norms within a social system. To achieve this task, we have developed a normative agent able to recognize and impose on defectors either punishment and sanction, and have implemented an proof-of-concept simulation model to test our hypotheses.

1 Introduction

Theoretical, empirical and ethnographic studies about punishment in human societies have demonstrated that this behavior promotes and sustains cooperation in large groups of unrelated individuals and more generally plays a crucial role in the maintenance of social order [16, 26, 9]. Moreover, in recent years, several mathematical and computational models have been designed to explore the various hypotheses concerning these issues, checking their implications in simplified and prototypical experiments [22, 20].

Although these accounts have unraveled the relevance of punishment in human societies, they have often overlooked that it actually consists in a complex behavioral repertoire, in which, as suggested by [17], it is useful to disentangle at least revenge, punishment and sanction.

In this paper, we will explore, by means of cognitive modelling and agent based simulation, the specific ways in which punishment and sanction favor the achievement of cooperation and the spreading of social norms in social systems populated by autonomous agents (for a detailed cognitive analysis of these two phenomena, see [17]). As both punishment and sanction are costly behaviors, both for the enforcer and the target, a well-designed enforcement system should combine high efficacy in discouraging cheating with limited costs for society.

Punishment and sanction are both mechanisms aimed at changing the behaviors of others, in order to make them abstain from future violations. Because of this similarity, these two phenomena are often mistaken one for another and considered as a *single* behavior. We claim that punishment and sanction are different behaviours and that can be distinguished on the basis of their mental antecedents and of the way in which they aim to influence the future conduct of others.

On the one hand, punishment is a practice consisting in imposing a fine to the wrongdoer, with the aim of deterring him from future offenses. Deterrence is achieved by modifying the relative costs and benefits of the situation, so that wrongdoing turns into a less attractive option. The effect of punishment is achieved by increasing individuals' expectations about the price of non-compliance.

This view of punishment is in line with the one supposed by the economic model of crime (see [6]) and with the approach adopted by experimental economics (see, [28], for a review of this approach).

On the other hand, sanction works by imposing a cost, as punishment does, and in addition by *communicating* to the target (and possibly to the audience) both the existence and the violation of a norm [17, 21, 32]).

The sanctioner ideally wants to induce the agent to comply with the norm not just to avoid punishment, but because he recognizes that there is a norm and wants to observe it for its own sake. Despite punishment, sanction aims to change the future behaviour of an agent by influencing both its *instrumental* and *normative* mind. In order to decide how to behave, the agent will take into consideration the norm and not only a mere costs and benefits measure.

We argue that norm compliance will be more robust if agents are enforced by sanction: where people have internal motivations to follow the norms, the frequency of compliance in the population will be higher than if people observe the norm only instrumentally (when it is in their interest to do so). Sanction are powerful social tools allowing norms and institution to be viable and robust across time.

More specifically, our hypotheses are that (1) sanctioning is characterized by a signaling function that possibly has the effect of generating or reinforcing intrinsic motivations to comply with the norm, favoring the achievement and stabilization of cooperation in a social system; (2) this signaling component allows norms to spread more quickly in the population, making it more resilient to change than if enforced only by mere punishment; and (3) sanctioning combines high efficacy in discouraging cheating with lower costs for society as compared to punishment.

In order to test the first hypothesis, we plan to conduct a series of experiments with natural subjects adopting a game-theoretical framework, while in this paper, we focus of the last two hypotheses, using agent-based simulation in order to test them.

The rest of the article is organized as follows: in Section 2, we introduce a theory of mental dynamics of norms, in order to provide some basic concepts. In Section 3 punishment and sanction will be analyzed and distinguished on the basis of the specific way in which they work in order to obtain deterrence. In Section 4.2, we present a rich normative agent architecture, allowing agents to be influenced by punishment and sanction and to process the normative information communicated by the latter. Finally, some simulation results aimed to compare the effectiveness

of punishment and sanction on the achievement of cooperation and norm spreading are presented and discussed (see Section 4). Future work and conclusions follow.

2 The mental path of norms

In order to understand how punishment and sanction work and which are their relative effects in the achievement of cooperation and the spreading of social norms, some preliminary notions should be clarified.

Building on Ullmann-Margalit's definition of norm [29] as *a prescribed guide for conduct which is generally complied with by the members of society*, we [1, 10] define a norm as a behavior that spreads through a given society to the extent that the corresponding prescription spreads as well, giving rise to a shared set of *normative beliefs* and *goals*. A normative belief is a mental representation, held to be true in the world, that a given action is either obligatory, forbidden or permitted for a given set of individuals in a given context. On the other hand, a normative goal is an internal goal *relativised* [4] to a normative belief: it is the will to perform an action *because* and *to the extent* that this is believed to be prescribed by a norm.

What leads agents endowed with normative beliefs and goals to execute them, especially since, by definition, norms prescribe costly behaviours? This usually happens by reference to external enforcement[5]. Agents calculate the relative cost and benefits of compliance and violation and then decide how to behave. But ideally, a norm wants to be complied with because it is an end in itself ([31]). This usually happens when a norm is internalized, i.e. when the norm addressee observes it independent of external sanctions and rewards. As discussed in [2], norm's salience is one of the main factors favouring norm intenalization: the more salient a normative mental representation is, the more it will elicit a normative behaviour [7, 32, 11].

Norm salience indicates to an agent how operative and relevant a norm is within a group and a given context [2]. It is a complex function, depending on several social and individual factors. On the one hand, the actions of others provide information about how important a norm is within that social group, in particular it depends on: (1) the amount of compliance and the cost people are willing to spend to comply ([11]); (2) the surveillance rate, the frequency and intensity of punishment ([19]) and the enforcement typology (private or public, 2nd and 3rd party, punishment or sanction, etc.) (see [24]); (3) the efforts and costs spared to educate the population to a certain norm; (4) the visibility and explicitness of the norm (see [11]); (5) the credibility and legitimacy of the normative source ([27]).

On the other hand, norm salience is also affected by the individual sphere, it depends on the degree of entrenchment with beliefs, goals, values and previously internalized norms of an agent ([15])[6].

[4] A goal is relativized when it is held because and to the extent that a given world-state or event is held to be true or is expected [12].

[5] See [13], for a fine grained analysis of different reasons behind norm compliance.

[6] It has to be pointed out that the norm salience can also gradually decrease: for example, it happens when the agent realizes that norm violations do not receive any

Norm compliance is expected to be more robust if agents' decisions are driven not only by instrumental considerations but are also based on normative considerations. Moreover, in many circumstances, an agent that complies with the norm for internal reasons will also exercise a special form of social control, getting others to comply with the norm, reproaching transgressors and reminding would-be violators that they are doing something wrong. Norm defence is extremely important in the spreading and stabilization of norms over a population of autonomous agents. As Axelrod [4] suggests, lowering the temptation to violate the norm might be not enough. Even in groups in which most people comply with the norm, if no one has an incentive to punish the remaining violators, the norm could still collapse. On the other hand, if agents are driven to honour norms for internal normative reasons, they are likely to defend it.

3 Punishment vs Sanction

As already said in the Sec. 1, we distinguish between two different enforcing strategies, punishment and sanction, often considered as a single behaviour[7].

On the one hand, punishment consists in imposing a cost to the target for the wrongdoing, to deter him from future offenses. This mechanism is aimed to change the future behavior of the agent influencing his strategic reasoning: the agent weighs up the relative costs and benefits of compliance and violation before deciding how to behave [23]. This approach to punishment is in line with the economic model of crime, also known as the rational choice theory of crime ([6]), claiming that the deterrent effect of punishment is caused by increasing individuals' expectations about the price of non-compliance. A rational comparison of the expected benefits and costs guides criminal behaviors and this produces a disincentive to engage in criminal activities.

This view of punishment has been criticized by several scholars stating that it considers *citizens just as consumers with unchanging or arbitrarily changing tastes in matters civic as well as commodity-related behavior* ([21]). Moreover, there is a large set of empirical evidences showing that punishment can increase cooperation also if it is purely *symbolic* and merely expresses social disapproval, without any material consequences for the punished individual ([25]).

We refer to this kind of punishment, mixing together material and non material aspects, as sanction, suggesting that it is aimed at inducing the target to abstain from further offenses not only to avoid the negative incentive, but in order to *respect* the norm.

This mechanism works by imposing a cost, as punishment does, and by intentionally communicating to the wrongdoer (and possibly to the audience) both the existence of a norm and that the performed action violated that norm. This communication can be achieved in several ways, for example (a) by transmitting

punishment or if the normative beliefs stay inactive for a certain amount of time, this meaning that the norm is not very active in the population anymore.

[7] Emotions, playing a significant role in these mechanisms, will not be investigated at this stage.

an evaluation such as "what a rude person!", (b) by showing indignation or blame, or (c) by explicitly mentioning the norm.

Sanction aims to focus agents' attention on several normative aspects, such as: (a) the existence and violation of a norm; (b) the high rate of norm surveillance in the social group; the causal link between violation and sanction: "you are being sanctioned because you violated that specific norm"; (c) the fact that the sanctioner is a norm defender.

All these normative messages, intentionally carried by sanction, have an impact on norm's activation and salience, different from that exerted by punishment. Because of its intentional signalling aim, sanction is more likely interpreted as a mechanism aimed at enforcing and defending a norm. On the other hand, punishment has not the same unambigous normative interpretation: it can be interpreted either as a norm-defense (thus having an impact on norm salience) or as an act of self-defense or just as a neutral act (having no impact on norm salience).

Given these differences, our hypotheses are that both punishment and sanction favor the increment of cooperation in social systems, but they have a different impact on the generation and spreading of social norms within the population and they involve different costs for the social system.

More specifically, as said in the *Introduction*, our hypotheses are that: (1) sanction is characterized by a signaling component allowing norms to spread more quickly in the population, favoring the achievement and stabilization of cooperation in a social system and making it more resilient to change than if enforced only by mere punishment; and, (2) sanctioning combines high efficacy in discouraging cheating with lower costs for society as compared to punishment.

In the following Sections, an agent based simulation aimed to test these hypotheses are presented and some preliminary results are discussed.

4 Simulation model

In order to verify the above mentioned hypotheses, we have developed a simulation model aimed to capture the specific dynamics of punishment and sanction and to test their relative effects in the achievement of cooperation and in the spreading of social norms. In our model, agents play a variation of the classic Prisoner's Dilemma (PD), where we included an extra stage to the game: after deciding whether to cooperate or not, agents can also choose whether they want (or not) to punish or sanction their opponent. The motivation for implementing the PD game is due to our long-term research goal aimed at studying enforcing technologies in virtual societies, and more specifically in environments like P2P scenarios or web-services markets. These types of scenarios share a number of characteristics with the PD game: dyadic encounters, repeated interactions, and one-shot games.

We assume that agents are located in a social network, which determines a fixed interaction topology [8]. Each timestep of the simulation is structured in 4 phases, that are repeated until convergence is reached (or for a fixed number of timesteps). More specifically, these phases consist in:

[8] Agents can only observe and interact with their direct neighbors.

1. Partner Selection: Agents are paired with other agents randomly chosen from their neighbors.[9].
2. First Stage: Agents play a PD game, with the following payoffs: $P(C,C) = 3,3; P(C,D) = 0,5; P(D,C) = 5,0; P(D,D) = 1,1$.
3. Second Stage: Agents decide whether to punish/sanction or not their opponent. Only defectors can be punished/sanctioned. If an agent decides not to punish/sanction and it is a norm-holder (i.e. an agent with an highly salient norm of cooperation stored in its mind), it can send an educational message to its opponent.
4. Strategy Update: As agents have mixed strategies[10], these strategies are updated on the basis of their decisions and of the social information acquired.

In the following paragraphs, we explain the options that agents have during the second stage.

4.1 Sanctioning or Punishing?

As already said in the Introduction, the main aim of this work is to study the differences between punishment and sanction in the achievement of cooperation and in favoring the spreading of social norms in a social system. We refer to *punishment* as the mechanism that only affects the target's payoffs, with a deterrent intention.

On the other hand, a *sanction* is aimed either to affect the target's payoffs (deterrent effect) and to signal the existence and violation of a norm (educative effect), communicating normative messages.

Only after having recognized and adopted the cooperation norm, agents use sanction against defectors, thus defending the norm; otherwise they will just punish them.

4.2 Agent Architecture

Despite the vast majority of simulation models of social in which heterogeneous agents interact according to simple local rules, in our model all the agents are endowed with normative architectures, allowing them (a) to influence each other by direct communication and by the use of punishment or sanction, (b) to recognize norms, (c) to generate new normative representations and to act on the basis of them. We based our architecture on a simplified version of EMIL-I-A [2], maintaining the norm salience process and deactivating the norm internalization module.

Our norm architecture has three important parts: the *norm recognition module*, the *salience meter*, and the *decision making*. The *norm recognition module* allows agents to recognize that a norm exists. In order for agents to recognize the existence of a norm, they have to listen by consistent agents [11] at least *two* normative messages

[9] This policy might lead to unpaired agents in certain configurations of topologies for each timestep, but the randomness of the partner selection ensures that all agents interact: those with higher degree will be more likely to interact than those with lower degrees.

[10] Differently to a pure strategy, mixed strategies have a probability with which an certain action will be chosen.

[11] An agent is consistent if when choosing to punish, he has before cooperated in the PD.

and observe or receive *ten* normative actions (i.e. cooperation, punishment and sanction, seen or received). [12]. Once these conditions are fulfilled, agents generate a normative belief, that will be stored in their normative board[13] and will activate a normative goal. The decision-maker is fed with the normative goal and compared with other (possibly active) goals. It will choose which one to execute (on the basis of their salience) and will convert it into a normative intention (i.e. an executable goal).

The *salience meter* indicates to the agent the social and individual salience of a certain norm. Salience makes norm compliance more stable and robust and enables the agents to *dynamically* monitor the functioning of the normative scene and to adapt according to it[14].

For example, in an unstable social environment, if the norm enforcement suddenly decreases, agents having highly salient norms are less inclined to violate them. A highly salient norm is a reason for which an agent continues to comply with it even in the absence of punishment. It guarantees a sort of inertia, making agents less prompt to change their strategy to a more favorable one. Vice versa, if a specific norm decays, our agents are able to detect this change, ceasing to comply with it and adapting to the new state of affairs.

4.3 Strategy Update

In this model, agents have two take two decisions at two different stages. These decisions are driven by the mixed strategies - cooperation and punishment - which are probability driven. Both the probabilities (cooperating at the first stage and being a punisher/sanctioner at the second stage) are obtained from an aggregation of agents' drives.

More specifically, agents have three drives that guide their decisions at the first stage:

(1) Self-Interested Drive: it motivates agents to maximize their individual utility independently of what the norm dictates. The self-interested drive is updated according to (a) the calculation of the marginal reward obtained during the last timestep, and (b) the actual action taken. A proportional and normalized value of the marginal reward obtained dictates how an agent would change its *cooperation probability*. Depending on the action taken, this value would move the agent's probability towards defection (if defected) or towards cooperation (if cooperated). For example, if defecting an agent improved its payoff of three units wrt the last

[12] These values are provisional and need to be fine tuned by experimentation with human-subjects.

[13] The normative board is a portion of the long-term memory of the agent where normative beliefs are stored, ordered by salience.

[14] It is interesting to note that this mechanism allows agents to record the social and normative information, without necessarily proactively exploring the word (e.g. with a trial and error procedure).

timestep, therefore, it will change its probability of cooperating with an intensity relative to 3 and towards defection[15].

(2) Social Drive[16]: Agents are influenced by what the majority of their neighbors do.

(3) Normative Drive: once recognized, agents decisions are affected by norms. The normative drive is affected by the norm's salience, and it would change the agent's strategy towards cooperation with a proportional value of it.

Therefore, and due to its close relation, in this section we also analyze how salience is updated according to the social information described below (we include a + sign if it makes the salience increase, and a - sign if it makes it decrease):

- *First Stage Cooperators Observed*: Neighboring agents who *cooperated* (+).
- *First Stage Defectors Observed*: Neighboring agents who *defected* (-).
- *Non Punished Defectors*: The amount of neighboring agents that defected at the first stage and were not punished at the second stage (-).
- *Consistent Punishment Observed*: Neighboring consistent agents who punished a defector, or, neighboring agents who have been punished by a consistent agent (+).
- *Consistent Sanction Observed*: Neighboring consistent agents who sanctioned a defector, or, neighboring agents who have been sanctioned by a consistent agent (+).
- *Consistent Educative Messages Observed*: Neighboring consistent agents who have sent an educative message, or, neighboring agents who have received an educative message from a consistent agent (+).
- *Consistent Punishment Received*: The amount of punishment a certain agent receives from a consistent agent (+).
- *Consistent Sanction Received*: The amount of sanction a certain agent receives from a consistent agent (+).
- *Consistent Educative Messages Received*: The amount of educative messages a certain agent receives from a consistent agent (+).

An aggregation of this social information influences and modifies the norm's salience. The weights' values used in the aggregation calculation (interpreted from [11] [17]) are the same used in [2]. From this aggregation function we will obtain a salience value that goes from $0 - 1$, meaning 0 that a norm is not salient, and 1 that a norm is completely salient.

Moreover, the punisher and the sanctioner are driven by different motivations. The former punishes in order to induce the future cooperation of others, thus expecting a future pecuniary benefit from its act ([23]). On the other hand, the sactioner is driven by a normative motivation: it sanctions in order to favor the generation and spreading of norms within the population. Its goal is to deter

[15] In case the marginal reward is 0 (this and last timestep reward are the same), agents would change their strategy with an inertial value in the same direction it last changed its probability.

[16] Even though we model this drive at the theoretical level, we have decided not to include it in the actual platform yet in order to have clearer results.

[17] The intuitive justification for the usage of these values is that of associating a higher weight to those social cues that are highly related to normative motivations and lower weights to those which would have selfish motivations.

agents from violating again because they come to share the norm and they ground their decision on it. We have modelled these drives, guiding agents' decision to punish/sanction or not in the following way:

- *Punishment Drive*: Agents change their tendency to punish on the basis of the relative amount of defectors wrt the last round.
- *Sanction Drive*: Agents change their tendency to sanction on the basis of their norm's salience.

Therefore, we can see how the mixed strategies are affected by agents' decisions and by social information. Eventually, these mixed strategies tend to extreme values (full cooperation or full defection, and complete punishment or no punishment), thus meaning that the system has converged.

4.4 Experimental Design

In order to analyze the differences of both types of punishment on the agents' decision making, we have performed an exhaustive experimental analysis. To reduce the search space (and save computation costs), we have prefixed some parameters (that we believe do not affect the results obtained): a total population of 100 agents, located in a fully connected network[18].

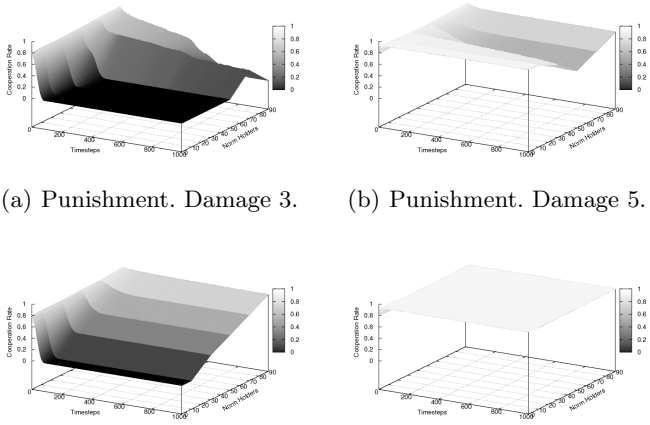
In the following experimental section, we contrast the results obtained in simulations where only punishment is allowed against situations in which both punishment and sanction are allowed. Only after having recognized the existence of the cooperation norm, agents will sanction defectors, thus defending the norm; otherwise they will just punish them. All the norm-holders are initially loaded with the norm and its salience at $0,8$. The initial cooperation probability for all agents is $0,8$ and a punishment probability of $0,5$. The amount of norm holders varies in each simulation, and they are specified in each figure.

We want to remind the reader that in this work we are not interested in analysing the emergence of norms, therefore some agents are initially loaded into the simulation with the cooperation norm (the norm's holders). If no agent had the norm, they would have to start a process of norm emergence that would include the recognition of an anti-social behavior, the identification of a possible solution and then the implementation in society.

4.5 Emergence of Cooperation

One of the main objectives of this research is to study the achievement of cooperation in adverse situations, where defecting would be the best strategy for the agents. In order to observe the relative effects of punishment and sanction in our artificial scenario, we exploit the advantages of having a agents endowed with normative minds allowing them to process the signals produced by these two enforcing mechanisms.

[18] Different social networks of interaction would definetly produce different dynamics in the system that at this moment we are not interested in analyzing. We refer the reader to [30].



(a) Punishment. Damage 3. (b) Punishment. Damage 5.

(c) Sanction. Damage 3. (d) Sanction. Damage 5.

Fig. 1. Effect on Cooperation with initial HP = 0.8.

In the first experiment, we pay attention to the relative effects of punishment and sanction on the achievement of cooperation.

In Fig. 1, we show the dynamics of punishment and sanctioning. The x-axis represents the timesteps of the simulation, the y-axis represents the average cooperation rates and the z-axis the initial amount of norm holders.

By observing the results, we notice that different damages (i.e. the amount of punishment/sanction imposed to the target) affect differently the cooperation rates. As expected, with a damage of 5 the agents' motivation to defect decreases in a much stronger way that when using a lower damage of 3[19]. It has to be pointed out that, despite what happens when using punishment, in populations enforced by sanction (and in which there is at least an initial amount of 70 norm's holders [20]), cooperation is also achieved when imposing a damage of 3.

Here follows the explanation of this result. Both punishment and sanction directly affect the self-interested drive of the agent, reducing its motivation to defect. However, these two enforcing mechanisms are effective in acheiving deterrence only when the damage is above 3. But, as we said in Paragraph 4.3, the probability of cooperating or not is also driven by the normative drive - affected by the norm's salience - and sanction has a stronger effect in increasing norm's salience than punishment.

[19] These values chosen as both 3 and 5 punishment damages turn the cooperative action more attractive in terms of payoff. A damage of 3 produces a *slight* improvement of the cooperative action (Payoff = 3) over the defection (Payoff 5 - 3 = 2). On the other hand, a damage of 5 produces a *stronger* difference between cooperation (Payoff = 3) and defection (Payoff = 0).

[20] In future versions of this work we will study the right proportion of norm's holders and their correct location in the topology to obtain a stronger effect on cooperation.

Moreover, a higher amount of norm holders speeds up the recognition process of the non-norm holders. Again, once the cooperation norm is recognized, the agent's cooperation probability is also affected by the normative drive and not only by self-interest motivations.

4.6 Norms Spreading

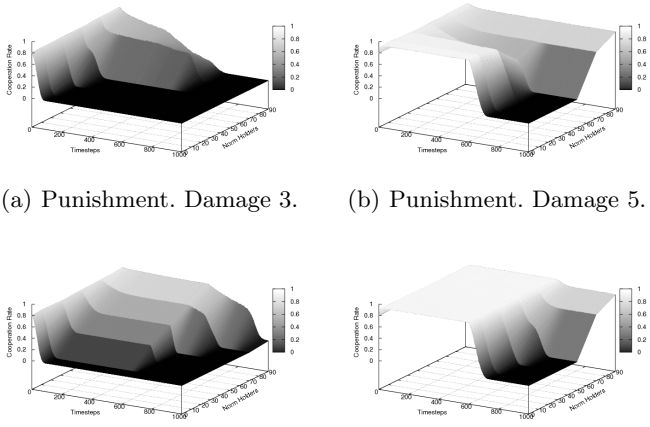
With the proof-of-concept experiment presented in the last paragraph, we not only tested the effects of punishment and sanctioning on the achievement of cooperation, but also, how the initial amount of norm holders affects the dynamics of cooperation. Analysing the same experimental data presented in the previous section, we now observe the relative effects of sanctions and punishments on the recognition of the cooperation norm. As said in Sec. 1, our hypothesis is that, thanks to its signaling nature, sanction allows norms to spread more and more quickly in the population, making it more resilient to change than if enforced only by mere punishment.

By paying attention to the average norm's salience per agent, we appreciate that sanction produces a stronger effect on the spread of norms than punishment, thus verifing our hypothesis. This phenomenon is mainly caused by the design of our normative architecture, where sanctions impact in a stronger way the agents' recognition and salience of norms.

Moreover, this is a self-reinforcing process: once an agent has recognized a norm, it will start sanctioning, getting others to comply with the norm, reproaching transgressors and reminding would-be violators that they are doing something wrong. This process affects and reinforces the sanctioner's own salience, but also the norm's salience of its neighbours.

4.7 Costs of Punishment

Our second hypothesis is that sanctioning combines high efficacy in discouraging defectors with lower costs for society as compared to punishment. From the data, we observe that in the two situations where both punishment and sanction allow cooperation to emerge (i.e. imposing a damage of 5), sanctions are 20,93% better in terms of the amount of sanction occurrences and the cost for the society wrt the situation where punishments are used. In other words, when using sanction, the amount of sanctioning acts and consequently the associated costs are decreased by 1/5. This is an interesting result that confirms our hypothesis, that the use of sanctions reduces the social costs of the achivement of cooperation as it impacts deeper on the salience. Moreover, (and even though the data obtained with Damage 3 are not comparable amongst them because of the different effects on the emergence of cooperation), we can observe that - from a system designer point of view - knowing the right value of the punishement/sanction to apply (5 better than 3 in this case) would reduce the costs (up to a 52% less), obtaining the same result on the emergence of cooperation. The reason for this phenomena is because a sanction of 5 would produce a stronger deterrent effect in the self-utilitarian drive of the agents than the sanction of 3, which needs to be repeatedly applied to achieve the same results.



(a) Punishment. Damage 3. (b) Punishment. Damage 5.

(c) Sanction. Damage 3. (d) Sanction. Damage 5.

Fig. 2. Effect on Cooperation with initial HP = 0.8 and with a block in the punishment.

4.8 Experiment: What happens when punishing/sanctioning is not possible?

This experiment is aimed to test the hypothesis that sanction, having a higher impact on the recognition and salience of the cooperation norm, makes agents more resilient to change than if enforced only by mere punishment. Our hypothesis is that if defection turns into an attractive option, for example because it becomes very unlikely that defectors are discovered or even more because there is no social control, we suppose that defectors will take longer to invade again the population in which sanction has been used. In this population, as observed in Section 4.6, a larger amount of cooperation norms have spread, this having a refraining effect on the decision of abandoning the cooperative strategy. In order to test this hypothesis, after the timestep 600 of the simulation, we deactivated the possibility to punish/sanction defectors. We motivate this experiment as a failure in the sanctioning system of any virtual environment.

In Figure 2, we can observe that, when suddenly control stops, agents enforced by sanction will continue to comply with the norm for a longer period compared to agents enforced by punishment. The explanation of this phenomena is again in the close relationship between sanctions (executed, observed and received) and their impact on norm's salience, as it still motivates them to comply with the norm even in the absence deterrent fines. One of the main advantages of this inertial effect of sanction is that policy makers and system designers can take advantege of this delay in order to restablish the state of the system.

5 Conclusions and future work

The study of punishment and sanctions is a challenging topic in Multi-Agent systems [18, 14, 5]. [14] shows a running example of punishment in a simulated environment, representing an interaction scenario regulated by a Prisoner's Dilemma (PD), where the interactions are restricted by a social network. Similar to this work, [8] uses a PD game in a peer-to-peer routing scenario, where punishment consists of blocking-time. Unfortunately, none of these works provide a rich cognitive architecture that would allow agents to process (and include into its reasoning and decision making) also the intrinsic signalling with which a punishment might be accompanied.

The proof-of-concept simulation presented in this work has served us to test our normative theory of punishment and the designed EMIL-I-A architecture. Our simulations results show the relative ways in which punishment and sanction affects the emergence of cooperation. More specifically, these results seem to verify our hypotheses that the signaling component of sanction allows this mechanism (a) to be more effective in the achievement of cooperation; (b) norms to spread more in the population, making it more resilient to environmental change than if enforced only by mere punishment; (c) to reduce significantly the social cost for cooperation to emerge. From a technical point of view there are a number of challenges to achieve in the short term.

First, we want to include an evolutionary mechanism allowing agents to dynamically calculate the right amount of punishment to impose in order to obtain deterrence. Endowing agents with this evolutionary mechanism might produce the same compliance results, eliminating wast and considerably reducing the social cost associated to the norm defence.

Second, now that the relationship between the self-interested and the normative drive has been analysed in detail, we are also interested in observing the dynamics introduced by the social drive. Could, in the right conditions, the behavior of peers motivate agent's behavior? Experiments with real human subjects [3] suggest so.

Moreover, we plan to perform several improvements to better understand the dynamics of sanctioning. The introduction of the Public Good Game would allow us to contrast our results with those obtained by experimental economists.

Finally, we plan to understand what are the differences between second and third party punishments. We hypothesize that by allowing agents to evolve their sanctioning strategies from a second to a third party fashion, social costs will be significantly reduced as the norm defence coverage will be increased.

Acknowledgments This work was supported by the Spanish Education and Science Ministry [Engineering Self-*Virtually-Embedded Systems (EVE) project, TIN2009-14702-C02-01]; Proyecto Intramural de Frontera MacNorms [PIFCOO-08-00017] and the Generalitat de Catalunya [2009-SGR-1434]; the European Science Foundation EUROCORES Programme TECT, funded by the Italian National Research Council (CNR); the EC Sixth Framework Programme and European project COST Action IC0801 Agreement Technologies. Daniel Villatoro is supported by a CSIC predoctoral fellowship under JAE program. We also thank the CESGA and Rede Galega de Bioinformatica for the technical support.

References

1. G. Andrighetto, M. Campennì, R. Conte, and M. Paolucci. On the immergence of norms: a normative agent architecture. In *Emergent Agents and Socialities: Social and Organizational Aspects of Intelligence. Papers from the AAAI Fall Symposium*, 2007.
2. G. Andrighetto, D. Villatoro, and R. Conte. Norm internalization in artificial societies. *AI Communications. (In press)*, 2010.
3. S. Asch. Opinions and social pressure. *Scientific American*, 193(5):31–35, 1955.
4. R. Axelrod. An evolutionary approach to norms. *The American Political Science Review*, 4(80):1095–1111, 1986.
5. A. L. C. Bazzan, S. R. Dahmen, and A. T. Baraviera. Simulating the effects of sanctioning for the emergence of cooperation in a public goods game. In *AAMAS '08*, pages 1473–1476, Richland, SC, 2008. International Foundation for Autonomous Agents and Multiagent Systems.
6. G. S. Becker. Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217, 1968.
7. C. Bicchieri. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, New York, 2006.
8. A. Blanc, Y.-K. Liu, and A. Vahdat. Designing incentives for peer-to-peer routing. In *INFOCOM*, pages 374–385, 2005.
9. R. Boyd and P. Richerson. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13(3):171–195, May 1992.
10. M. Campennì, G. Andrighetto, F. Cecconi, and R. Conte. Normal = normative? the role of intelligent agents in norm innovation. *Mind & Society*, 8:153–172, 2009.
11. R. B. Cialdini, R. R. Reno, and C. A. Kallgren. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6):1015–1026, June 1990.
12. R. Cohen, P and J. Levesque, H. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261, 1990.
13. R. Conte and C. Castelfranchi. *Cognitive and social action*. University College of London Press, London, 1995.
14. A. P. de Pinninck, C. Sierra, and W. M. Schorlemmer. Friends no more: Norm enforcement in multi-agent systems. In M. Durfee, E. H.; Yokoo, editor, *Proceedings of AAMAS 2007*, pages 640–642, 2007.
15. E. L. Deci and R. R.M. The "what" and "why" of of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry*, 2000.
16. E. Fehr and S. Gachter. Altruistic punishment in humans. *Nature*, 415:137–140, 2002.
17. F. Giardini, G. Andrighetto, and R. Conte. A cognitive model of punishment. In *COGSCI 2010, Annual Meeting of the Cognitive Science Society 11-14 August 2010,*, pages –. Portland, Oregon, 2010.
18. D. Grossi, H. M. Aldewereld, and F. Dignum. Ubi lex, ibi poena: Designing norm enforcement in e-institutions. In *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, pages 101–114. Springer, 2007.
19. D. F. . K. J. Haley. *Genetic and cultural evolution of cooperation*, chapter The strategy of affect: emotions in human cooperation, pages 7–36. Cambridge, MA: The MIT Press, 2003.
20. D. Helbing, A. Szolnoki, M. Perc, and G. Szab. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New Journal of Physics*, 12(8):083005, 2010.

21. A. O. Hirschman. Against parsimony: Three easy ways of complicating some categories of economic discourse. *American Economic Review*, 74(2):89–96, May 1984.
22. K. Jaffe and L. Zaballa. Co-operative punishment cements social cohesion. *Journal of Artificial Societies and Social Simulation*, 13:3, 2010.
23. D. M. Kreps, P. Milgrom, J. Roberts, and R. Wilson. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2):245 – 252, 1982.
24. D. Masclet. L'analyse de l'influence de la pression des pairs dans les quipes de travail. CIRANO Working Papers 2003s-35, CIRANO, June 2003.
25. C. Noussair and S. Tucker. Combining monetary and social sanctions to promote cooperation. Open Access publications from Tilburg University urn:nbn:nl:ui:12-377935, Tilburg University, 2005.
26. E. Ostrom, J. Walker, and R. Gardner. Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, 86(2):404–417, 1992.
27. A. Sacks, M. Levi, and T. Tyler. Conceptualizing legitimacy: Measuring legitimating beliefs. *American Behavioral Scientist.*, In Press.
28. K. Sigmund. Punish or perish? retaliation and collaboration among humans. *Trends Ecol Evol*, 22(11):593–600, 2007.
29. E. Ullman-Margalit. *The Emergence of Norms*. Clarendon Press, Oxford, 1977.
30. D. Villatoro, S. Sen, and J. Sabater. Topology and memory effect on convention emergence. In *Proceedings of the International Conference of Intelligent Agent Technology (IAT)*. IEEE Press, 2009.
31. G. H. von Wright. *Norm and Action. A Logical Inquiry*. Routledge and Kegan Paul, London, 1963.
32. E. Xiao and D. Houser. Emotion expression in human punishment behavior. *Proc Natl Acad Sci U S A*, 102(20):7398–7401, May 2005.