Dynamic Sanctioning for Robust and Cost-Efficient Norm Compliance

Daniel Villatoro
IIIA-CSIC
Bellaterra, Spain
dvillatoro@iiia.csic.es

Giulia Andrighetto
European University Institute, Italy
ISTC - CNR, Rome, Italy
giulia.andrighetto@istc.cnr.it

Jordi Sabater-Mir
IIIA-CSIC
Bellaterra, Spain
jsabater@iiia.csic.es

Rosaria Conte
ISTC - CNR, Rome, Italy
rosaria.conte@istc.cnr.it

Abstract

As explained by Axelrod in his seminal work *An Evolutionary Approach to Norms*, punishment is a key mechanism to achieve the necessary social control and to impose social norms in a self-regulated society. In this paper, we distinguish between two enforcing mechanisms. i.e. *punishment* and *sanction*, focusing on the specific ways in which they favor the emergence and maintenance of cooperation. The key research question is to find more stable and cheaper mechanisms for norm compliance in hybrid social environments (populated by humans and computational agents). To achieve this task, we have developed a normative agent able to punish and sanction defectors and to dynamically choose the right amount of punishment and sanction to impose on them (*Dynamic Adaptation Heuristic*). The results obtained through agent-based simulation show us that sanction is more effective and less costly than punishment in the achievement and maintenance of cooperation and it makes the population more resilient to sudden changes than if it were enforced only by mere punishment.

1 Introduction

Theoretical, empirical and ethnographic studies have demonstrated that punishment in human societies promotes and sustains cooperation in large groups of unrelated individuals and more generally plays a crucial role in the maintenance of social order [Fehr and Gachter, 2002; Sigmund, 2007]. Several mathematical and computational models have been designed to explore how punishment - and in particular altruistic punishment - can emerge, and how it can be maintained [Jaffe and Zaballa, 2010; Boyd *et al.*, 2010; Boyd and Richerson, 1992; Helbing *et al.*, 2010]. And other models have been designed to explain how to choose the most effective punishment to regulate (e-)Institutions [Grossi *et al.*, 2007; Janssen *et al.*, 2010; Rauhut and Junker, 2009; de Pinninck *et al.*, 2007; Blanc *et al.*, 2005].

Although these studies have provided key insights to the understanding of punishment, they have largely viewed at this mechanism from the classical economic perspective as a way of changing people's conduct by increasing the costs of undesired behaviour [Becker, 1968]. The model of decision making advocated by this perspective is that of the rational actor influenced only by economic incentives.

We claim that this way of considering punishment is incomplete and not likely to mantain large-scale cooperation at least at reasonable costs for the social system. Instead, we argue that punishment is effective in regulating people's behaviour not only through economic incentives, but also for the normative request it asks people. In some situations, apart from imposing a cost for the wrongdoing, the punisher informs violators (and the public) that the targeted behaviour is not approved and that it violates a social norm, thus focusing their attention on that norm. We claim that when this happens, cooperation is more stable and the costs for achieving and maintaing it are lower than when only economic incentives are used.

In previous work, ([Giardini *et al.*, 2010; Andrighetto *et al.*, 2010]), the term *punishment* has been used to refer to the enforcement mechanism aimed at obtaining deterrence only by changing the costs and benefits of a particular situation; while we used *sanction* to indicate the mechanism aimed at changing people' conduct also by signalling to violators (and the public) that their behaviour is not approved and that it violated a social norm. Clearly, in real life situations there is often an overlap - even if very slight - between these two mechanisms; analysing punishment and sanction in isolation however allows us to explore the specific contribution of each of them to the achievement and maintenance of cooperation and possibly to design actions aimed to highlight and exploit such contributions.

Recently, researchers have conducted several laboratory experiments with human subjects designed to explore the norm-signalling effect of sanction, analysing what factors might impact the expressive power of this mechanism [Masclet, 2003; Noussair and Tucker, 2005], but to our knowledge, the work presented here is the first simulation study that focuses specifically on this topic. Simulation experiments allow us to isolate *in silico* punishment and sanction, verify their relative effects on cooperation, and perform what-if analyses that allow us to address policy design issues. Moreover, our simulation framework allows us to prove different adaptation heuristics for the monetary damages associated to sanction and punishment, in order to observe their effects on the dynamics of the establishment of cooperation.

In particular, in this paper we explore the hypothesis that cooperation is more stable and less costly for society if individuals are enforced by sanctions: this enforcing strategy has the effect of activating agents' normative motivation to cooperate, leading to a more durable cooperation than if agents are driven only by the instrumental motivation to avoid punishment. More specifically, the norm-signaling component of sanction allows social norms to be activated and to spread more quickly in the population than if it were enforced only by mere punishment. This normative elicitation has the effect of increasing pro-social behaviours and consequently cooperation within the population.

The article is organized as follows: in Sec. 2, punishment and sanction are analyzed and distinguished on the basis of the specific ways in which they work in order to obtain deterrence. In Sec. 3, we present a rich normative agent architecture, which allows agents to be influenced by punishment and sanction and to process the normative information communicated by the latter. Finally, in Section 4 some simulation results comparing the effectiveness of punishment and sanction in the achievement and maintenance of cooperation, and their relative costs for the system are presented and discussed. Future work and conclusions follow.

2 Punishment vs Sanction

As already stated in Section 1, we distinguish between two different enforcing strategies, punishment and sanction. On the one hand, we refer to punishment as a practice that consists in imposing a cost on 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 becomes a less attractive option. The effect of punishment is achieved by influencing the *instrumental* mind of the individual, by shaping his material payoffs. This approach to punishment is in line with the economic model of crime, also known as the rational choice theory of crime [Becker, 1968], 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.

On the other hand, we use sanction to indicate the enforcing strategy that apart from imposing a cost for the wrongdoing is also intentionally aimed at (1) *signalling* to the target (and possibly to the public) that his behaviour is not approved because it violated a social norm ([Giardini *et al.*, 2010; Hirschman, 1984; Xiao and Houser, 2005; Andrighetto *et al.*, 2010])[1] and (2) *asking* him to comply with it in the future. The sanctioner ideally wants the sanctioned to change his conduct not just to avoid the penalty but because he recognizes that there is a norm and wants to respect it. Sanction mixes together material and symbolic aspects and it is aimed at changing the future behaviour of an individual by influencing both its *instrumental* and *normative* mind. In order to decide how to behave, the individual will take into consideration not only a mere costs and benefits measure but also the norm.

Often the sanctioner uses scolding to reign in free-riders, or expresses indignation or blame, or simply mentions that the targeted behaviour violated a norm. Through these actions, he aims to focus people's attention on different normative aspects, such as: (a) the existence and violation of a norm; (b) the high rate of norm surveillance in the social group; (c) the causal link between violation and sanction: "you are being sanctioned because you violated that norm"; (d) the fact that the sanctioner is a norm defender. As suggested by works in psychology [Cialdini *et al.*, 1990; LaVoie, 1974], all these normative messages have a key effect in producing norm compliance and favouring social control as well. Even a strong personal commitment to a norm does not predict behaviour if that norm is not activated or made the focus of attention [Bicchieri, 2006; Cialdini *et al.*, 1990]. Furthermore, the more these norms are made *salient*, the more they will elicit a normative conduct.

Norm salience indicates to an individual how operative and relevant a norm is within a group and a given context [Andrighetto *et al.*, 2010]. It is a complex function (presented in Sec. 3.2), 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. 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 the agent.

We claim that both punishment and sanction favor the increment of cooperation in social systems, but sanction achieves cooperation in a more stable way and at a lower cost for the system. Cooperation is expected to be more robust if agents' decisions are driven not only by instrumental considerations but are also based on normative ones. Moreover, an individual that complies with the norm for internal reasons is also more willing to exercise a special form of social control as well, reproaching transgressors and reminding would-be violators that they are doing something wrong.

3 Normative Framework

In order to capture the specific dynamics of punishment and sanction and to test their relative effects in the achievement and maintenance of cooperation a simulation model has been developed. In this model, agents play a variation of the classic Prisoner's Dilemma Game (PDG)[2], where an extra stage has been included in the game: after deciding whether to cooperate or not, agents can also choose whether they want (or not) to punish or sanction the opponents who defected. The motivation behind the PDG 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

[1] Clearly, also punishment can have a norm-signalling effect as an unintended by-product, but only the sanctioner intentionally has this norm-defense goal.

[2] The PDG payoffs are the following: $P(C,C) = 3,3; P(C,D) = 0,5; P(D,C) = 5,0; P(D,D) = 1,1$. Where C stands for Cooperate and D for Defect

a number of characteristics with the PDG: dyadic encounters, repeated interactions, and one-shot games.

Each timestep of the simulation is structured in 4 phases, that are repeated for a fixed number of timesteps:

1. Partner Selection: Agents are randomly paired with their neighbors in the social network.

2. First Stage: Agents play the PDG, deciding whether to cooperate (C) or to defect (D).

3. Second Stage: Agents decide whether to punish/sanction or not the opponents who defected. Only agents who have recognized that there is a norm of cooperation governing their group (see Section 3.1) use sanction to enforce others' behaviours; otherwise punishment is used. Punishment works by imposing a cost to the defector, this way affecting its payoffs. On the other hand, apart from imposing a cost, sanction also informs the target (and possibly the audience) that the performed action violated a social norm, thus having an impact both on agents' payoffs and on the process of norm recognition and norm salience. If an agent decides not to punish/sanction and it is a norm-holder (i.e. an agent with a 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, these strategies are updated on the basis of their decisions, payoffs and social information acquired.

3.1 Agent Architecture

Our agent's architecture is inspired in the EMIL-I-A architecture [Andrighetto *et al.*, 2010], which allows agents to recognize which social norms are governing their group and to modify the degree of salience of these norms. However, in this work we have adapted the architecture and tuned the weights that update norm salience with data obtained from a combined research with experimental economists (normalization shown in Table 1), aimed at analyzing the interactions of human and virtual agents in an equivalent setting.

Our norm architecture has three important parts: the *norm recognition module*, the *salience meter*, and *decision-making*.

The *norm recognition module* allows agents to interpret a social input as a norm. In order for agents to recognize the existence of a norm, they have to hear at least *two* normative messages from consistent agents[3], such as "you should not take advantage of your group members by shirking" and observe *ten* normative actions compliant with the norm or aimed to defend it (i.e. cooperation, punishment and sanction, observed or received). Once these conditions are fulfilled, our agents generate a normative belief that will activate a normative goal (the normative drive) to comply with the norm. The decision-maker is fed with the normative goal and compared with other (possibly active) goals of the agent. 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 how salient a certain norm is. This norm salience meter makes norm compliance more stable and robust, enabling the agents to *dynamically* monitor the normative scene and to adapt according to it. 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 such as in Q-Learning). For example, in an unstable social environment, if the norm enforcement capability 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.

3.2 Strategy Update

In this model, agents have to take two decisions at two different stages: to cooperate or defect and to punish/sanction or not, and both of them are probability driven. These decisions are influenced by an aggregation of material and normative considerations. More specifically, the decision to cooperate or to defect is affected by the following drives:

(1) Self-Interested Drive (SID): it motivates agents to maximize their individual utility independently of what the norm asks. The SID is updated according:

$$SID_t = SID_{t-1} + (O \times \frac{R_t - R_{t-1}}{R_{Max} - R_{Min}})$$

where SID_t represents the self-interested drive at time t, O the orientation (+1 if the agent Cooperated and -1 if it Defected), R_t the reward obtained at time t, and R_{Max} and R_{Min} respectively the maximum and minimum reward that can be obtained. In the case where the marginal reward is zero, it is substituted by an inertial value with the same orientation as in the last variation. This way, the proportional and normalized value of the marginal reward obtained indicates how the agent would change its *cooperation probability*.

(3) Normative Drive (ND): once the cooperation norm is recognized, agents decisions are influenced also by the normative drive. The normative drive is affected by the norm salience: the more salient the norm is, the higher the motivation to cooperate. Salience is updated according to the formula below and the social weights described in Table 1:

$$Sal_t^N = Sal_{t-1}^N + \frac{1}{\alpha \times \phi}(w_C + O \cdot w_O + NPD \cdot w_{NPD} + P \cdot w_P + S \cdot w_S + E \cdot w_E)$$

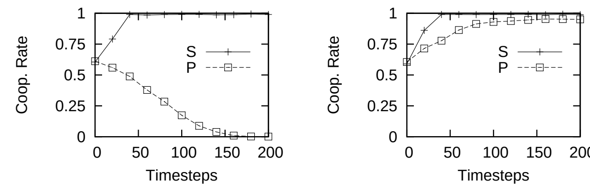
where Sal_t^N represents the salience of the norm N at time t, α the number of neighbors that the agent has, ϕ the normalization value, w_X the weights specified in Table 1, and finally O, NPD, P, S, E indicate the registered occurences of each cue. The resulting salience measure ($salience \in [0-1]$, 0 representing minimum salience and 1 maximum salience) is subjective for each agent thus providing flexibility and adaptability to the system.

The agents who cooperated can decide to punish/sanction defectors. As stated, only agents who have recognized that there is a cooperation norm regulating their group can sanction, otherwise they will just use punishment. As discussed in Section 2, 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

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

Social Cue	Weight
Norm Cooperation/Defection (C)	w_C = (+/-) 0.99
Observed Norm Compliance (O)	w_O = (+) 0.33
Non Punished Defectors (NPD)	w_{NPD} = (-) 0.66
Punishment Observed/Applied/Received (P)	w_P = (+) 0.33
Sanction Observed/Applied/Received (S)	w_S = (+) 0.99
Explicit Norm Invocation Observed/Received (E)	w_E = (+) 0.99

Table 1: Norm Salience Meter: Cues and Weights.



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

Figure 1: Effects of Punishment (P) and Sanction (S) on the Emergence of Cooperation.

pecuniary benefit from its acts. On the other hand, the sactioner is driven by a normative motivation: it sanctions to favor the generation and spreading of norms within the population. Given these differences, the probability governing the decision of punishing or sanctioning is modified by different factors[4] and they change in the following way:

(4) Punishment Drive: Agents change their tendency to punish on the basis of the relative amount of defectors with respect to the last round. If the number of defectors increased, agents' motivation to punish will decrease accordingly.

(5) Sanction Drive: Agents change their tendency to sanction on the basis of the norm salience. The more salient the norm is, the higher the probability to sanction defectors.

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

4 Experiments

One of the main objectives of this research is to study the achievement of cooperation in adverse situations, where defecting is the utility-maximizing 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 agents endowed with normative minds allowing them to process the signals produced by these two enforcing mechanisms separately.

As in this work we are not interested in analysing the emergence of norms, some agents already endowed with the cooperation norm are initially loaded into the simulation: we refer to them as initial norm's holders (INHs), and are initially loaded with the norm at a salience of $0,8$. 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 the consequent implementation in society.

All the following experiments have been run on two different topologies: a fully connected network (where agents have access to complete information) and a scale-free network (where agents only access local information). Despite obtaining the same final convergence results in both topologies, we have observed a delay in the scale-free networks, caused by the cascade effect of the spread of norms produced by the location of the INHs (better connected INHs produce faster convergence results).

4.1 Emergence of Cooperation

In this first experiment, we analyze the relative effects of punishment and sanction on the achievement of cooperation, paying attention to the amount of INHs and the different damages imposed with punishment and sanction. By comparing Fig. 1(a) and Fig. 1(b), we observe that different damages (i.e. the amount of punishment/sanction imposed to the target) affect the cooperation levels differently. As expected, agents' motivation to defect decreases in a much stronger way with a damage of 5 than with a lower damage of 3[5]. It has to be pointed out that, despite what happens when using punishment, in populations enforced by sanction (with a minimum amount of INHs, 10 in this experiment), cooperation is also achieved when imposing a lower damage, as can be seen in Fig. 1(a).

Both punishment and sanction directly affect the agents' SID, reducing their motivation to defect. However, these two enforcing mechanisms are effective in achieving deterrence only when the damage imposed is at least 3, as with a damage of 3 (Cooperation Payoff = 3, Defection Payoff = 5 - 3 = 2) cooperation is the utility-maximizing strategy.

As said in Sec. 3.2, agents' cooperation probability is affected by both the SID and the ND: sanction - thanks to its signalling component - influences the normative drive more than punishment. In order to obtain deterrence, punishment exploits the power of norms much less than sanction, that is why it needs to impose higher damages on its targets.

The amount of INHs produces an interesting result (although not reported with any figure due to space constraints): when the number of INHs is increased, emergence is achieved faster and it follows a distribution equal to the neighbors distribution (in the regular networks the emergence of cooperation increases linearly and exponentially in scale-free networks).

Besides allowing cooperation to emerge faster than punishment, another crucial advantage of sanction is its cost-

[4]Even though agents pay a cost to enforce defectors, this cost is not taken into account when they update their decisions to punish and sanction.

[5]These values have been chosen for experimentation as both 3 and 5 punishment damages turn the cooperative action into the utility-maximizing strategy. A damage of 3 produces a *slight* improvement for cooperaton (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).

efficiency. When sanction is used to achieve cooperation, an average reduction of the 32% of the social costs is obtained compared to punishment, as can be seen in the rows labeled as "Static" in Table 2. Sanction affects the normative drive of agents more than punishment, making compliance more robust and cheap.

4.2 Dynamic Adaptation Heuristic

In the experiments shown in Fig. 1(a) and Fig. 1(b), the damage that both punishers and sanctioners can impose on defectors in order to deter them from violating again is fixed for the entire duration of the simulation. But in some circumstances, this damage could be reduced, without lowering its deterring effect. In order to allow our agents to dynamically choose the right amount of punishment and sanction to impose, thus reducing social costs and wasting of resources, a *Dynamic Adaptation Heuristic* has been implemented. This heuristic works in the following way:

if $Defectors_{t-1} \leq Defectors_t$ *AND* $Defectors_t > ToleranceDef$
then Increase PunCost by Δ;
if $Defectors_{t-1} > Defectors_t$ *OR* $Defectors_t < ToleranceDef$
then Decrease PunCost by Δ;

Algorithm 1: Dynamic Adaptation Heuristic (DAH).

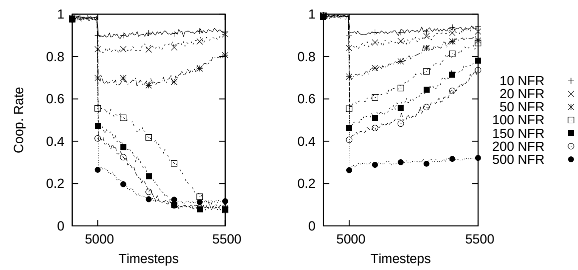
By keeping track of the amount of defectors in the previous timestep ($Defectors_{t-1}$), and comparing it with the actual amount of defectors, the imposed cost of punishment is adapted consequently with a Δ ($\Delta = 0.1$ in this work), thus obtaining an intelligent dynamic adaptation.

	Punishments	Sanctions	Ind. Cost	Social Cost
Static Pun	4.8748	0	5	24.374
DAH Pun	2.1142	0	5.1640	10,9177
Static Sanc	0.3364	1.2491	5	7,9275
DAH Sanc	**0.2932**	**0.8669**	**3.7031**	**4.2959**

Table 2: Performance of the Dynamic Adaptation Heuristic. Average values per timestep.

In order to test how the implemented DAH affects the performances of both punishment and sanction, an exhaustive exploration of the search space with different amounts of INHs (from 10 to 90, increasing at intervals of 10) has been performed. Results obtained for different amounts of INHs follow the same distribution, for the cases with and without DAH. In Table 2, the average performances of both punishment and sanction in the dynamic and static conditions are shown. The DAH allows society to significantly reduce the number of punishing (57% less) and sanctioning (27% less) acts with respect to the static condition.

However, when using DAH, the average individual cost of punishment is slightly higher than that of sanctioning. This is given by the cyclic dynamics produced by agents driven only by their strategic drives. When enforced by punishment, agents abandon their defecting strategy because they want to avoid the cost of punishment. The number of cooperators starts to increase and punishers decrease the punishment cost accordingly. However, this reduced punishment cost makes defection to become the utility maximizing strategy. Consequently, the number of defectors will increase again and the punishment cost accordingly. With this newly adapted punishment cost, defection will not be the optimal strategy anymore, reaching the initial situation again.



(a) Punishment. (b) Sanction.

Figure 2: New Free Riders introduced at TS = 5000.

Consequently, the social expenses using DAH are considerably reduced both when punishing (66%) and sanctioning (56%). The implemented heuristic allows society to intelligently reduce the social costs needed for the achievement and maintenance of cooperation.

4.3 Adapting to the Environment: Free Riders Invasions

This experiment is aimed to test the hypothesis that sanction makes the population more resilient to change than if it were enforced only by mere punishment. If suddenly a large amount of new defectors joins the population, we suppose that defectors will take longer to invade the population in which sanction has been used. In order to confront the relative speed of adaptation and degree of resilience of the populations enforced with punishment and sanction, we run simulation experiments in which after timestep 5000, new free-riders (from a minimum amount of 10 to a maximum of 500) have been injected in the populations[6].

Experimental results (see Fig. 2) show that a population enforced by (DAH) sanction is able to receive up to 200 new free riders and still to mantain a high level of cooperation, while when (DAH) punishment is used, only 100 new free riders make cooperation collapse. In the population enforced by sanction, a larger amount of cooperation norms have spread, this having a refraining effect on the decision of abandoning the cooperative strategy. Highly salient norms guarantee a sort of inertia, restraining agents to change their strategy to a more favorable one.

5 Conclusions and future work

The study of punishment and sanctions is a challenging topic in Multi-Agent systems [Grossi *et al.*, 2007; de Pinninck *et al.*, 2007]. Several authors ([de Pinninck *et al.*, 2007; Blanc *et al.*, 2005]) have tested the effect of punishment in regulating peer-to-peer simulated environments, showing that to solve

[6]Reported results are from experiments performed on Fully Connected Networks, as in Scale-Free Networks behave differently depending on the positioning of the newly introduced free riders, a result out of the scope of this work.

free-riding problems a constant and stable punishment system is necessary.

To our knowledge, this is the first work in which agents are endowed with rich cognitive architectures allowing them to be affected by the normative information associated to sanction and to dynamically gauge the amount of damage to impose on defectors. The simulation results presented in this paper clarify the relative ways in which punishment and sanction affect the emergence of cooperation. More specifically, these results verify our hypotheses that the signaling component of sanction allows this mechanism (a) to be more effective in the achievement of cooperation; (b) to make the population more resilient to environmental changes than if enforced only by mere punishment; (c) to reduce significantly the costs for cooperation to emerge. Our tasks for the future will consist of a detailed analysis of the topological effects in scale-free networks (including the strategic positioning of the Inital Norm Holders and free-riders) and the improvement of the Dynamic Adaptation Heuristic, by allowing a flexible adaptation of the Δ value.

Acknowledgments

This work was supported by the European University Institute; the Complexity-Net Project SEMIRA; the EC Sixth Framework Programme and European project COST Action IC0801 AT; the Spanish Education and Science Ministry [Engineering Self-* Virtually-Embedded Systems (EVE) project, TIN2009-14702-C02-01]; MacNorms Project [PIFCOO-08-00017]; the Generalitat de Catalunya [2005-SGR-00093]; 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

[Andrighetto *et al.*, 2010] Giulia Andrighetto, Daniel Villatoro, and Rosaria Conte. Norm internalization in artificial societies. *AI Communications*, 23:325–339, 2010.

[Becker, 1968] Gary S. Becker. Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217, 1968.

[Bicchieri, 2006] C Bicchieri. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, New York, 2006.

[Blanc *et al.*, 2005] Alberto Blanc, Yi-Kai Liu, and Amin Vahdat. Designing incentives for peer-to-peer routing. In *INFOCOM*, pages 374–385, 2005.

[Boyd and Richerson, 1992] 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.

[Boyd *et al.*, 2010] Robert Boyd, Herbert Gintis, and Samuel Bowles. Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare. *Science*, 328(5978):617–620, April 2010.

[Cialdini *et al.*, 1990] Robert B. Cialdini, Raymond R. Reno, and Carl 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.

[de Pinninck *et al.*, 2007] Adrian Perreau de Pinninck, Carles Sierra, and W. Marco 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.

[Fehr and Gachter, 2002] E Fehr and S Gachter. Altruistic punishment in humans. *Nature*, 415:137–140, 2002.

[Giardini *et al.*, 2010] 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.

[Grossi *et al.*, 2007] 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.

[Helbing *et al.*, 2010] Dirk Helbing, Attila Szolnoki, Matja Perc, and Gyrgy Szab. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New Journal of Physics*, 12(8):083005, 2010.

[Hirschman, 1984] Albert O Hirschman. Against parsimony: Three easy ways of complicating some categories of economic discourse. *American Economic Review*, 74(2):89–96, May 1984.

[Jaffe and Zaballa, 2010] Klaus Jaffe and Luis Zaballa. Cooperative punishment cements social cohesion. *Journal of Artificial Societies and Social Simulation*, 13:3, 2010.

[Janssen *et al.*, 2010] Marco A. Janssen, Robert Holahan, Allen Lee, and Elinor Ostrom. Lab Experiments for the Study of Social-Ecological Systems. *Science*, 328(5978):613–617, April 2010.

[LaVoie, 1974] Joseph LaVoie. Type of punishment as a determinant of resistance to deviation. *Developmental Psychology*, 10(2):181–189, 1974.

[Masclet, 2003] David Masclet. L'analyse de l'influence de la pression des pairs dans les quipes de travail. CIRANO Working Papers 2003s-35, CIRANO, June 2003.

[Noussair and Tucker, 2005] C.N. Noussair and S. Tucker. Combining monetary and social sanctions to promote cooperation. Open access publications from tilburg university, Tilburg University, 2005.

[Rauhut and Junker, 2009] Heiko Rauhut and Marcel Junker. Punishment deters crime because humans are bounded in their strategic decision-making. *Journal of Artificial Societies and Social Simulation*, 12(3):1, 2009.

[Sigmund, 2007] K. Sigmund. Punish or perish? Retaliation and collaboration among humans. *Trends in Ecology & Evolution*, 22(11):593–600, November 2007.

[Xiao and Houser, 2005] E. Xiao and D. Houser. Emotion expression in human punishment behavior. *Proc Natl Acad Sci U S A*, 102(20):7398–7401, May 2005.