

Foreword

The monograph you are about to read is a study on social phenomena. The study of social phenomena has been approached by many disciplines, each one following a different methodology, analyzing the phenomena from a different perspective and, sadly, usually ignoring the other disciplines that are also studying the same phenomena. In this monograph however, you will find a bunch of disciplines working together to analyze two very relevant topics: the formation of conventions and the behavior of humans in front of punishment. Using artificial intelligence as a hinge, Dr. Villatoro has been able to seat in the same table (can be this a consequence of his passion for cooking?) several disciplines to work together pursuing a common objective. So, if I had to use a single word to define the work presented in this monograph, I have no doubt that I would use the word “interdisciplinary”. The monograph you have in your hands is a very nice example of what can be accomplished when ideas and tools coming from different disciplines (namely artificial intelligence, experimental economics, game theory and cognitive science) are combined. This is especially clear in the second part of the monograph. Experimental Economics (using new technological tools provided by artificial intelligence) provides the “know how” to analyze the behavior of humans regarding two social phenomena: distributed punishment and the distinction between sanction and punishment. Cognitive Science provides a model that is able to explain the observed behavior in the Experimental Economics experiments. Finally artificial intelligence, and specifically agent based simulation, provides the tools to implement the cognitive model and put it into the mind of artificial entities that can participate in a recreation of the experiments performed previously in the lab with humans. The results of the simulation, that match those obtained in the lab with humans, serve to support the cognitive model. An interesting aspect of the interdisciplinary work presented here is that each discipline obtains an individual benefit of the join work, apart from contributing to the common objective. Is not the case that a discipline has to renounce to advance because it has to collaborate with other disciplines to achieve a common goal. Experimental economics can give an explanation to an observed behavior thanks to the cognitive model and the tools provided by artificial intelligence, cognitive science can validate a model about social phenomena thanks to experimental economics and agent based simulation, and finally artificial intelligence obtains a model that can be implemented and used in artificial entities so they behave like humans in these specific aspects. The work presented in the first part about the formation of conventions, subconventions and the mechanisms to dissolve them, is just the icing on the cake. This

kind of join work among disciplines is not as common as it should be. I'm sure you will find it very interesting and illustrative of what can be achieved by joining efforts among disciplines.

Bellaterra, November 2012

Jordi Sabater-Mir
IIIA - CSIC

Acknowledgements

Muchas horas y dedicación están invertidas en este documento, pero afortunadamente este camino no lo he recorrido solo. Aquí sólo nombrare a algunas de las personas e instituciones que han contribuido activamente a empujarme a lo largo de este camino.

Primero de todo, tengo que agradecer a mi director de tesis, Jordi Sabater por dos motivos. Primero de todo, agradezco que hace años confiara en un africano entrevistado via Skype y por haberme dejado explorar y crecer científicamente. Gracias a él, no sólo he aprendido el método e integridad científica, sino que me ha ayudado a madurar como persona, enfrentándonos juntos a momentos adversos y alegrías insospechadas. El segundo motivo por el que le estoy agradecido es por haberme concedido esta oportunidad junto al Señor Piñol: un gran compañero y amigo, con el que sufrimos el proyecto eRep. ¿Quién diría que tantas horas descifrando y programando, nos iba a llevar a brindar (una vez más) por la ciencia en el Danubio?

Y también son protagonistas de estos brindis las siguientes personas, a las que les estoy infinitamente agradecido y que me han hecho el camino más llevadero. Meritxell, guapa!, por ser nuestro ejemplo a seguir, profesional y personalmente. Angi, por tus ánimos, fuerza y compañerismo. Sr. Nin, por el humor y los cafes. Mari Carmen, por esa sonrisa inagotable y la perseverancia ejemplar. Peter (Dani Polak), por habernos acompañado siempre. A las chicas de Administración e Inma Moros, por haberme solucionado uno y mil problemas, siempre con una sonrisa en la cara. A Tito, por ser un perdedor en la mesa de pingpong pero un buen amigo y un profesional con el cual me siento orgulloso de haber trabajado. A Elena, por tantas cervezas juntos en tantos puntos del mundo intentando sobrellevar las penas de la ciencia. A Pablo Noriega, por haberme dejado compartir muchos ratos a su lado, ilustrándome con su elegancia, y por haberme ayudado una infinidad de veces con consejos más sabios de lo que él imagina. A Pedro Meseguer, por acortar las distancias entre doctorandos y científicos con su amistad. A Juan Antonio Rodríguez, por todos sus consejos en momentos clave. Además, quiero agradecer a todos y cada uno de los miembros del IIIA, ya que para mí es un orgullo haber sido parte de esta institución, y haber compartido todos estos años con sus miembros.

I cannot forget thanking the partners from the projects I have been involved in: David de la Cruz, Jordi Brandts, Héctor Solaz, Assumpció Vila, Jordi Estévez, Raquel Piqué, Manuela Rodríguez, Thijs, Wander, Tina, Stefan, Torsten, and the Yámanas.

I would like to thank Professor Sandip Sen, for hosting me in the University of Tulsa and sharing with me hours of work. I have learnt a lot from him and an important part of this thesis was accomplished thanks to him.

Infine, non potrò mai dimenticare i colleghi e amici di Roma. Vorrei ringraziare

Rosaria Conte per avermi accolto nel LABSS e avermi insegnato questo mestiere. Federica Mattei per i mille favori. Francesca Giardini, Gennaro di Tosto, Stefano Picascia, Francisco Grimaldo, Luca Tummolini, Walter Quattrociochi, Mario Paolucci, Federico Cecconi e Cristiano Castelfranchi per avermi ospitato. E non ci sono parole in italiano per ringraziare Giulia Andrighetto: mi ha dato la spinta di forza per finire questa tesi. La voglio ringraziare anche per avermi trasmesso la sua passione scientifica e per essere stata un'incredibile collega e una buonissima amica. *Grazies milles* cara mia!

I want to thank the Santa Fe Institute, for choosing me to participate in their summer school, giving me the chance to meet an incredible group of scientists in one of the best research labs of the world.

Quiero agradecer también a todas esas personas que habéis confiado en mí durante estos años y que sin vosotros esta tesis no se hubiera llevado a cabo: Júlia Cot, Tomás Fuentes, Adriá Serra, Jordi López, Dani Vico, Jordi Esteller, Carles Amagat, Charly Burgmann, Pedro Martín, Jose Ramón Lozano, Alejandro Parodi, Elvira Quero, María Elías, Cecily Roche, Javier Medina, Mariola Ruiz, Gabriela Petrikovich, Maite Segu, Anna Trillo, Estibaliz Puente, Víctor Cornet, Mario Bolívar, Ingrid Dolader, Ory Shoemaker y Alfonso Rodríguez (fling).

Para finalizar no puedo dejar de agradecer a mi familia. Vuestro apoyo incondicional ha sido fundamental para alcanzar esta meta, y sólo gracias a vosotros ha sido posible. A mis tías Puri y Toñi, y a mis primas de Barcelona, por haberme hecho sentir como en casa. A mi Tata por su buen humor y todo el cariño que siempre me manda; a mi prima Inés, por no fallarme nunca; a mi cuñada Cristina por su ánimo; a mis sobrinos Pablo, Julia y Jorge por confiar en mí; a mis hermanos Diego y Fernando, mis roles a seguir y por tenerlos siempre tan cerca; y finalmente, a mi madre, por transmitirme una parte de la fuerza, el empeño y la perseverancia que la caracteriza y a todos nos fascina; y a mi padre, por haber confiado en mí y apoyarme siempre, y por las sabias palabras que una vez me enseñó y me han ayudado a acabar la “selectividad”: *“Cabeza alta, paso firme y mucha mala leche”*.

Gracias a todos.

¡Por la ciencia!

Abstract

Social norms are one of the mechanisms for decentralized societies to achieve coordination amongst individuals. Such norms are conflict resolution strategies that develop from the population interactions instead of a centralized entity dictating agent protocol. One of the most important characteristics of social norms is that they are imposed by the members of the society, and they are responsible for the fulfillment and defense of these norms. By allowing agents to manage (impose, abide by and defend) social norms, societies achieve a higher degree of freedom by lacking the necessity of authorities supervising all the interactions amongst agents.

In this thesis we approach social norms as a malleable concept, understanding norms as dynamic and dependent on environmental situations and agents' goals. By integrating agents with the necessary mechanisms to handle this concept of norm, we have obtained an agent architecture able to self-police its behavior according to the social and environmental circumstances in which is located.

First of all, we have grounded the difference between conventions and essential norms from a game-theoretical perspective. This difference is essential as they favor coordination in games with different characteristics.

With respect to conventions, we have analyzed the search space of the emergence of conventions when approached with social learning. The exploration took us to discover the existence of Self-Reinforcing Structures that delay the emergence of global conventions. In order to dissolve and accelerate the emergence of conventions, we have designed socially inspired mechanisms (*rewiring* and *observation*) available to agents to use them by accessing local information. The usage of these social instruments represent a robust solution to the problem of convention emergence, specially in complex networks (like the scale-free).

On the other hand, for essential norms, we have focused on the "Emergence of Cooperation" problem, as it contains the characteristics of any essential norm scenario. In these type of games, there is a conflict between the self-interest of the individual and the group's interest, fixing the social norm a cooperative strategy. In this thesis we study different decentralized mechanisms by which cooperation emerges and it is maintained.

An initial set of experiments on Distributed Punishment lead us to discover that certain types of punishment have a stronger effect on the decision making than a pure benefit-costs calculation. Based on this result, we hypothesize that punishment (utility detriment) has a lower effect on the cooperation rates of the population with respect to sanction (utility detriment and normative elicitation). This hypothesis has been integrated into the developed agent architecture (EMIL-I-A). We validate the hypothesis by performing experiments with human subjects, and observing that behaves accordingly to human subjects in similar scenarios (that represent the emergence of cooperation).

We have exploited this architecture proving its efficiency in different in-silico scenarios, varying a number of important parameters which are unfeasible to reproduce in laboratory experiments with human subjects (because of economic and time resources).

Finally, we have implemented an *Internalization* module, which allows agents to reduce their computation costs by linking compliance with their own goals. Internalization is the process by which an agent abides by the norms without taking into consideration the punishments/sanctions associated to defection. We have shown with agent based simulation how the Internalization has been implemented into the EMIL-I-A architecture, obtaining an efficient performance of our agents without sacrificing adaptation skills.

Chapter 1

Introduction

1.1 Motivation

Social norms have been claimed to be one of the research topics in multi-agent systems to be investigated [Luck et al., 2005]. Social norms are part of our everyday life, and they have been of interest in several areas of research [Elster, 1989]. Social norms help people self-organizing in many situations, specially in open-environments where having an authority representative is not feasible. On the contrary to institutional rules, the responsibility to enforce social norms is not the task of a central authority but a task of each member of the society. From the book of Bicchieri [Bicchieri, 2006], the following definition of social norms is extracted:

“The social norms I am talking about are not the formal, prescriptive or proscriptive rules designed, imposed, and enforced by an exogeneous authority through the administration of selective incentives. I rather discuss informal norms that emerge through the decentralized interaction of agents within a collective and are not imposed or designed by an authority”.

Although this definition does not pretend to be a formal definition, it gives us the notion that social norms are used in human societies as a mechanism to improve the behavior of the individuals in those societies without relying on a centralized and omnipresent authority. In recent years, the use of these kinds of norms has been considered also as a mechanism to regulate virtual societies and specifically societies formed by artificial agents ([Saam and Harrer, 1999, Shoham and Tennenholtz, 1992, Grizard et al., 2006]), and in some cases, mixed populations (agents and humans).

The study of social norms within multi-agent systems has been framed inside the Computational Social Science community, which is *“in charge of collecting and analyzing data at scale that may reveal patterns of individual and group behaviours”* [Lazer et al., 2009]. The combination of both techniques (computer science and social science) makes special sense when we deal with multi-agent systems, as agents are social entities. The exploitation of the knowledge provided by the social sciences is becoming more useful with the integration of humans and agents (E.g. [Brito et al., 2009]).

Latterly, and due to the interest of the multi-agent community on social norms, the NorMAS movement has been founded [Nor, 2008]. This movement gathers an interdisciplinary community (mainly formed by computer scientists, social scientists, psychologists, and economists) around the topic of the norms in multi-agent systems. From one of the workshops organized by this community, the following definition of normative multi-agent system was agreed by consensus:

“A normative multi-agent system is a multi-agent system organized by means of mechanisms to represent, communicate, distribute, detect, create, modify, and enforce norms, and mechanisms to deliberate about norms and detect norm violation and fulfillment” [Boella et al., 2008].

The online emergence of social norms (how they are created at first time) in a decentralized way is one of the problems in the community that needs to be solved and some authors are trying to approach. As we are focusing on decentralized open virtual societies, different norms might emerge on different areas of the social topology. However, social norms are by definition ([Coleman, 1998]) more socially efficient when the whole population abides by them. We are interested in studying the spreading and acceptance of social norms, what Axelrod [Axelrod, 1986] calls *norm support*. Our understanding of norm support deals with the problem of which norm is established as the dominant when more than one norm exists for the same situation, or in those situations where the agents' self-interest is against the interest of the society, and therefore violating the norms is the strategy that provides the highest payoff.

There are basically two well-known approaches to study this emergence: the game-theoretical and the cognitive. The former studies the process by which agents calculate the cost-benefit ratio of the possible actions taking into consideration the existence of a norm [Villatoro et al., 2010] and the strategic behavior of the other agents involved in the interactions. On the other hand, the latter studies the emergence from the point of view of the beliefs, desire and intentions of the agents and their relation to normative beliefs [Andrighetto et al., 2010b].

After reviewing the literature on both research approaches, we agree with Young [Young, 2008] that there are three mechanisms by which norms emerge, that generalize all the factors previously treated in the literature:

- Pure Coordination: These are “social” phenomena, because they are held in place by shared expectations about the appropriate solution to a given coordination problem, but there is no need for social enforcement.
- Threat of social disapproval or punishment for norm violations.
- Internalization of norms of proper conduct.

The first mechanisms have been widely studied under the topic of convention emergence [Shoham and Tennenholtz, 1997a, Kittock, 1993, Delgado et al., 2003, Mukherjee et al., 2007, Sen and Airiau, 2007, Walker and Wooldridge, 1995] (e.g. the typical example of these kind of mechanisms are which side of the road should cars drive: both agents are benefited from following the coordination norm, otherwise, their utility is drastically reduced).

However, we have observed a general practice in that area of research which consists in establishing the system as converged when 90% of the population shares the same convention. Because of the definition of convention (a norm that establishes a focal point amongst the possibles to promote coordination when shared by all participants), we cannot consider 90% as an acceptable convergence rate and we perform an exhaustive study on that area. We focus on aspects like the strategy update rule used by agents or the topology that fixes their interactions to study full convergence (100%) in the system. We will observe how certain strategy update rules and topologies promote the existence of subconventions, that delay the emergence of a global convention. We propose socially-inspired methods to dissolve subconventions and allow the society to reach the desired full-convergence.

On the other hand, the second and third points are tightly related amongst them. Sanctions serving as a mechanism for norm emergence has been widely studied by social scientist [Coleman, 1998], psychologists [Bandura, 1976] and economists [Fehr and Gächter, 2002]. In this thesis we will study how different punishment technologies can reinforce the process of norm emergence, allowing the system to self-policy. We initially envisioned to donate agents with the necessary mechanisms to achieve *Multilateral Costly Punishment* [Posner and Rasmusen, 1999]. In order to be realistic with the parameter setting for the Multilateral Costly Punishment, we performed experimental economics experiments that proved us the existence of a stronger force behind the punishment: the normative message affecting a normative decision making. With the light thrown in those experiments, we decided to develop a cognitive agent architecture that not only performs a benefit-cost calculation with respect to the existence of norms, but it also considers the fact of the norm's existence as determinant in the decision making. In this part we will also prove the determinant difference between punishment and sanction, which is an important element for our developed architecture, and a state-of-the-art discovery in the behavioural economics literature. As we said previously, the third point affirmed by Young for norms to emerge (*Internalization*) is in a direct relationship with our research on punishment technologies. In our work [Andrighetto et al., 2007, Conte, 2009, Andrighetto et al., 2010b], we intend *Internalization* as the process where an agent becomes compliant with norms not because the fear of potential punishment (and utility detriment) but just because the mere existence of norms and the willing to abide by them. We will observe how this process ensures a more robust society against norm-violations, without sacrificing adaptation skills.

1.2 Contribution

This thesis contributes to the field of self-organized and normative multi-agent systems in three lines:

First - Exhaustive Exploration of the Dimensions in the Convention Emergence Problem

We present an experimental framework to analyze the dynamics in the process of

convention emergence, observing the effects of the different parameters configuring the model. Specially, we focus on the problem of the emergence and dissolution of subconventions. We have detected how these subconventions emerge mainly for two reasons: the strategy update rule promoting concordance with previous history (and culture maintenance) and topological conditions promoting endogamy. The conducted research took us to discover the existence of Self-Reinforcing Structures, mainly in Scale-Free Networks, producing subconventions to remain metastable. We propose socially-inspired mechanisms that dissolve these Self-Reinforcing Structures, allowing the society to achieve full convergence.

Second - Experimental Platform for Regulated Hybrid Experiments

In order to obtain experimental results about certain punishment technologies, a platform allowing to test the behavior of humans under certain conditions (achieved by the partners interactions) was necessary. Another requirement for the platform was a restricted behavioural set of available actions to agents and a pre-established interaction protocol. Because of the previous reasons, the paradigm of Electronic Institutions seemed the most adequate.

We have developed a friendly user web-interface for human subjects to interact with other subjects inside an Electronic Institution. With the usage of this platform we allow humans to interact from remote locations through a simple web interface. On the other hand, the electronic institution paradigm allows us, as experiment designers, to implement agents that perform in an specific way, locating human subjects in specific interesting experimental conditions.

This platform has been essential to perform behavioral economics experiments, introducing agents into them, and allowing subjects to participate remotely, advancing the state-of-the-art platforms (like Z-tree [Fischbacher, 2006]).

Third - EMIL-I-A Architecture for Self-Regulating Societies

We have developed a BDI agent architecture which distinguishes at a cognitive level the difference between *punishment* and *sanction*. Experimental results with human subjects confirmed the difference between punishment (utility detriment) and sanction (utility detriment plus norm elicitation) in the decision making of human subjects. This difference is incorporated precisely in our agent architecture (EMIL-I-A), producing effects on the decision making of the agents at several levels. Moreover, we test the performance of our EMIL-I-As in the same experimental conditions than with humans. These experiments show that EMIL-I-As produce the same dynamics than those showed by humans, obtaining an state-of-the-art agent architecture with punishment capabilities able to interact (accordingly) with humans.

Finally, we implement an *Internalization* mechanism that allows norm compliance to happen without the external enforcement of punishment, proving the adaptive capabilities of EMIL-I-A while saving in cognitive load.

1.3 Overview and Structure of the Thesis

This thesis is structured in nine chapters:

Chapter 2: We present a game-theoretical difference between conventions and essential norms. This difference is basic to structure the rest of the thesis. Then we make an analysis of the state of the art in the MAS literature in both fields. With respect to conventions, we review the different works performed on the emergence of conventions and how they treated different aspects of such process. On the other hand, we review the different works and mechanisms analyzed for the emergence of cooperation. We focus on those that are applicable for self-organizing societies, paying special attention at MAS with punishment technologies incorporated.

Chapter 3: In this chapter we present the general framework for the convention emergence problem. This framework allows us to study the different dimensions affecting the convention emergence, focusing in the achievement of full convergence. A detailed analysis of the search space of parameters allows us to discover that full convergence was not previously achieved because of certain strategy update rules and certain topological configurations promoting subconventions.

Chapter 4: In order to dissolve the identified subconventions, we propose socially-inspired techniques that also accelerate the process of convention emergence. With the usage of these instruments we discover the existence of Self-Reinforcing Structures, and propose a combined social instrument for dissolving the subconventions created in those areas of the topology.

Chapter 5: In this chapter we propose distributed punishment as a punishment technology that can produce norm emergence in situations represented by common good games. Experimental results with human subjects give us the intuition that other types of punishment may exist and affect differently to the decision making of the agents.

Chapter 6: We present the EMIL-I-A architecture, conceived to interpret the difference between punishment and sanction at a cognitive level. This architecture is constructed assuming the existence of normative beliefs and goals, which are orchestrated with the norm salience, representing the degree of activation in the social environment of the different recognized norms.

Chapter 7: In this chapter we prove the correctness of the EMIL-I-A architecture by simulating the same experiments where human subjects participated, obtaining similar results. Moreover, we exploit our agent architecture in different experimental conditions, understanding the dynamics of our agents when interacting in different environmental situations. Finally we introduce a Dynamic Adaptation Heuristic of the cost of punishment in order to achieve a more efficient sanction, profiting from its cognitive load, and maintaining high cooperation rates.

Chapter 8: This chapter discusses a cognitive mechanism that allows agents to internalize norms, and therefore, comply with them without external enforcement. By means of simulation in a P2P inspired scenario, we will observe how internalization allows agents to efficiently respond to their peer needs, adapting to environmental conditions and dynamically changing the normative scheme.

Chapter 9: We conclude our research of normative self-organized systems and sketch future lines of research to exploit the work presented in this thesis.

1.4 Related Publications

The following publications are a direct consequence of the development of the thesis.

- G. Andrighetto, D. Villatoro, F. Cecconi, R. Conte. Simulazione ad Agenti e Teoria della Cooperazione. Il Ruolo della Sanzione. Sistemi Intelligenti (forthcoming)
- G. Andrighetto, D. Villatoro, R. Conte. Norm Dynamics within agents. In B. Edmonds Dynamic View of Norms. Cambridge University Press (forthcoming)
- D. Villatoro, J. Sabater-Mir and S. Sen. Social Instruments for Robust Convention Emergence. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011). (In press).
- D. Villatoro, G. Andrighetto, R. Conte and J. Sabater-Mir. Dynamic Sanctioning for Robust and Cost-Efficient Norm Compliance. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011). (In press).
- G. Andrighetto, D. Villatoro. Beyond the Carrot and Stick Approach to Enforcement: An Agent-Based Model. Proceedings of the European Conference on Cognitive Science, New Bulgarian University, Sofia, 21-24 May 2011. Cognitive Science Society.
- T. Balke and D. Villatoro. Operationalization of the Sanctioning Process in Hedonic Artificial Societies. Proceedings of the AAMAS Workshop on Coordination, Organizations, Institutions and Norms (COIN @ AAMAS 2011).
- D. Villatoro, J. Sabater-Mir and S. Sen. Social Instrument for Convention Emergence. Proceedings of 10th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2011).
- D. Villatoro, S. Sen and J. Sabater-Mir. Exploring the dimensions of Convention Emergence in Multi-agent Systems. Advances in Complex Systems (ACS) Volume No.14, Issue No. 2 pp 201-227. (2011).
- G. Andrighetto, D. Villatoro, R. Conte and J. Sabater Mir. Simulating the relative effects of punishment and sanction in the achievement of cooperation. Proceedings of the Eighth European Workshop on Multi-Agent Systems (EUMAS10).
- R. Conte, G. Andrighetto, D. Villatoro. From Norm Adoption to Norm Internalization. (DEON 2010).
- G. Andrighetto, D. Villatoro, R. Conte. Norm internalization in artificial societies. AI Communications. Vol No.23, Issue No.4 pp.325-339. (2010)
- D. Villatoro, S. Sen and J. Sabater-Mir. Of Social Norms and Sanctioning: A Game Theoretical Overview. International Journal of Agent Technologies and Systems, Vol. 2, Issue 1, pp 1-15. (2010).

- A. Vila-Mitjá, J. Estévez, D. Villatoro and J. Sabater-Mir. Archaeological Materiality of Social Inequality Among Hunter-Gatherer Societies. Archaeological Invisibility and Forgotten Knowledge: Conference Proceedings, Łódz, Poland, 5th-7th September 2007. Archaeopress. 2010
- D. Villatoro, S. Sen and J. Sabater-Mir. Topology and memory effect on convention emergence. Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology. (IAT 2009).
- D. Villatoro, N. Malone and S. Sen. Effects of interaction history and network topology on rate of convention emergence. Proceedings of 3rd International Workshop on Emergent Intelligence on Networked Agents (WEIN'09 @ AAMAS).
- D. Villatoro and J. Sabater-Mir. Dynamics in the Normative Group Recognition Process. Proceedings of IEEE Congress on Evolutionary Computation (IEEE CEC 2009).
- I. Brito, I. Pinyol, D. Villatoro, J. Sabater-Mir. HIHEREI: Human Interaction within Hybrid Environments. Proceedings of 8th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2009) (*Best Student Demo Award*).
- D. Villatoro and J. Sabater-Mir. Group Recognition through Social Norms. Proceedings of 8th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2009).
- D. Villatoro and J. Sabater-Mir and S. Sen. Interaction, observance or both? Effects on convention emergence. Proceedings of Dotzè Congrés Internacional de l'Associació Catalana d'Intelligència Artificial (CCIA 2009).
- D. Villatoro and J. Sabater-Mir. Towards the Group Formation through Social Norms. Proceedings of the Sixth European Workshop on Multi-Agent Systems (EUMAS08).
- D. Villatoro and J. Sabater-Mir. Towards Social Norm. Proceedings of the Eleventh International Congress of the Catalan Artificial Intelligence Association (CCIA08).
- D. Villatoro and J. Sabater-Mir. Mechanisms for Social Norms Support in Virtual Societies. Proceedings of the Fifth Conference of the European Social Simulation Association (ESSA08).
- D. Villatoro and J. Sabater-Mir. Categorizing Social Norms in a Simulated Resource Gathering Society. Proceedings of the AAAI Workshop on Coordination, Organizations, Institutions and Norms (COIN @ AAAI08).
- D. Villatoro and J. Sabater-Mir. Norm Selection Through Simulation in a Resource-Gathering Society Proceedings of 21st European Simulation and Modelling Conference (ESM07).

- J. Sabater-Mir, I. Pinyol, D. Villatoro and G. Cuni. Towards Hybrid Experiments on reputation mechanisms: BDI Agents and Humans in Electronic Institutions. Proceedings of 12th Conference of the Spanish Association for Artificial Intelligence (CAEPIA07).