1. Identification of project and participants

PROJECT FULL TITLE: Simulating the Emergent Impact of Regulations Across cultures

Project acronym: SEMIRA
Project number: 101
Period covered: from: 1/1/2011 to: 31/12/2012
Duration of project: 2 years
Project coordinator: Gert Jan Hofstede

Name, title and organization of the representative of the project’s coordinator:

Assoc. Prof. Gert Jan Hofstede (dr ir), Wageningen University

Tel: +31 (0)317 4 84630
Fax:
E-mail: gertjan.hofstede@wur.nl

Project website address: www.semira.wur.nl

Identification of project participants / beneficiaries (includes official national registration number)

(1) Gert Jan HOFSTEDE, Wageningen University, Netherlands (999981634)
(2) Frank DIGNUM, Utrecht University, Netherlands
(3) Rosaria CONTE, CNR, Rome, Italy
(4) Ana PAIVA, INESC-ID, Lisbon, Portugal (PT 504 547 593)
2. Short description of activities and intermediate results

Below is the graphical organisation of the SEMIRA project into work packages, along with a summary of the projects meetings held so far. The partners are based in The Netherlands, Italy and Portugal, which are the three countries whose experiences after the smoking ban in bars are studied in this project.

The rest of this section describes the work accomplished in each work package. Note that the work packages are presented in the most logical, rather than numerical, order.
D2: Initial model for the interaction of regulations, emerging norms and culture.

Model components for SEMIRA agents

![Diagram of model components for SEMIRA agents]

Our task for this work package was to decide how normative reasoning and cultural factors could be integrated in a dynamic model of social norms.

Human social life is immensely complicated, but modelling a small set of its most salient features can allow AI models to mimic human interactions, and hence also the emergent normative consequences of multiple such interactions. Cross-disciplinary cooperation is essential for this undertaking to be successful.

In order to model the influence of culture on the evolution of social norms, it is first important to realise that processes at distinct timescales are implicated. The longest timescale is that on which culture itself evolves: centuries and longer. The next important timescale is that in which social norms can emerge in a group; this can be over days or months. The shortest timescale relevant for SEMIRA is the moment-to-moment interactions in a bar environment. For SEMIRA the representation of culture in our agents shall be fixed throughout all simulations, in large-scale models the social norms shall evolve, and in a virtual environment scenario the social norm shall be fixed, see three layers in Figure 1.

Long-term Evolution: Culture

Our representation of culture is based on the Hofstede Dimensions of Culture. We published a paper Mc Breen et al. (2011) outlining one possible such approach. It is a framework in which social norms can evolve against a background of deeper, culture-dependent, behavioural influences. These dispositions to perform certain behaviours in a particular relational context, we refer to as *meta-norms*. Individuals have instincts and beliefs about the rules for appropriate behaviour in typical situations. Meta-norms are shared in a culture and constant over the timescales of both our Large-Scale-Model (LSM), and Virtual
Environment Model (VEM). They encode the behaviour that is learnt from one’s earliest days of how one should behave.

Meta-norms are ‘relationally operationalized values’. They have pre-conditions regarding the relational context for when they apply, and post-conditions that then prescribe a certain appropriate behavioural intention. Meta-norms hence influence the creation of social norms based on cultural values. These are very often unconscious values, but they are values that are very effective in predicting behaviour.

Medium-term Evolution: Practices

Practices are the more malleable expression of cultural values. Some practices are prescribed by social norms, such as “not smoking in the presence of children”, and others are simple conventions, such as driving on the left or right. We are primarily interested, not in the detail of practices, but in their social meaning. Socially meaningful practices are enacted in rituals, see below.

We consider the defining feature of any social situation to be its Moral Circle. That is the group of people present who influence the behaviour and interpretation of an individual. Much behaviour is guided primarily by the group context. We hope to use Moral Circles to better represent social behaviour. In an implemented system the Moral Circle helps the agents to decide who matters in a given context and which behaviours are appropriate. The existing social norms of a society, as well as the Moral Circle configurations within which members of that society interact, condition the adoption of new social norms.

Real-time interactions: Rituals

The virtual environment developed as part of the SEMIRA project represents enacted behaviours in real time, within a bar. A crucial element of social behaviour in such situations is shared attention. That is, when those present are aware that they are in a shared experience that has social meaning. This we call a Ritual, and we conceive of such rituals as a basic unit of group interaction. Both the physical context and who is present are crucial influences of such interactions.

In order for our computer agents to be seen as believable characters by human users, we expect that users need to see that agents are aware of the social impact of their actions. When users can observe and interact with these agents, it is important that the embodied virtual agents show believable behaviour. This is a requirement in order to elicit meaningful reactions from those human users, regarding the social appropriateness of the agents’ behaviours. The reactions of the users can validate the cultural appropriateness of the agents’ behaviours. Cross-cultural comparisons of user reactions allow us to validate our implemented theories of social behaviour, and to examine the plausibility of culture as a driving force in the adoption or rejection of a new regulation in different countries.

Two conceptual papers on modelling culture and norms have been published, a third on norms and emotions in a scenario will be discussed in section D.6.
D3: Collecting data on smoking ban use case for validation

There were two major sources of data exploited when investigating the situation around smoking in public places both before and after the introduction of the smoking bans. The first was the numerous Eurobarometer reports related to tobacco use throughout the EU over the last seven years, and the second was the International Tobacco Control project. Both sources used randomly selected representative samples within each country surveyed. These data cover not only the three principal countries studied in this project but other European countries as well. This larger comparison made it apparent the Dutch are an outlier in European terms with regard to attitudes to the smoking ban in bars and cafes. Notable results of this research include:

- Italy began to legally restrict smoking in non-hospitality sectors much earlier than the other two countries, but these laws weren't widely followed at the time. Later data shows that the smoking ban in bars was quite strictly followed in Italy, and the ban in other places began to be followed once the bar ban was introduced.
- The support for banning smoking in bars before the introduction of the ban (I>P>NL), strongly predicts the current levels of smoking in bars in the three countries (NL>P>I). The figures in the Eurobarometer Survey of 2009 indicate that the percentages of respondents reporting witnessing smoking inside a bar were NL: 87, P: 39, I: 13. In both The Netherlands and Portugal there is a co-existence of smoking and non-smoking bars, with smoking bars being more common in The Netherlands.
- The number strongly supporting the ban increases significantly in all 3 countries after the ban comes into force. The Dutch, however, had a significant backlash against the workplace-smoking ban, once the hospitality sector ban came into force.
- The Dutch are slightly more aware of the dangers of passive smoking, but Dutch smokers are the most likely to expose others to Environmental Tobacco Smoke (ETS) in the home. They also have the most relaxed attitude to drinking while pregnant and smoking in the presence of pregnant women. The hierarchy of the frequency of drinking while smoking is (NL>P>I).
- The International Tobacco Control project involves the Netherlands and 20 other countries, but does not include Italy and Portugal. It shows that support, even among smokers, for the bans in bars and restaurants increases once the ban is introduced. However, this effect is weaker for bars than for other catering venues, and particularly weak in Dutch bars.

The three countries chosen for this project are representative of the breadth of different experiences within the EU regarding the smoking ban, with Italy and The Netherlands near the extremes.
D1: Initial agent model for individual decision making based on enriched BDI.

In this work package, we aimed at creating embodied agents that not only react and behave in a believable way in face-to-face interactions, but also do so according to their specified cultural and normative parameterization. As such we needed to consider not only a generic set of basic components that the agent’s architecture may embed (such as a reactive component, a planner, an emotional component) but also, specific requirements associated with the need to behave in a group, according to some rules, interacting with the other in a specific culturally shaped manner, etc.

A general diagram of the agent architecture developed is shown in Figure 1. It resulted from the integration of cultural aspects of human behaviour in an emotional agent architecture that used the traditional BDI paradigm. The chosen emotional architecture for the integration was FAIIMA (Dias & Paiva 2005, Lim et al 2008) that follows the OCC model of emotions.

![General Diagram of the Cultural Agent Architecture](image)

**Figure 2** – General Diagram of the Cultural Agent Architecture. (Affect derivation means the emotional reaction of an agent to an event.)

In general terms, the architecture works in the following manner. When an event is perceived, it passes through a Symbol Translator that translates the meaning of the event according to the culture’s predefined symbols, using a simple association mechanism. For instance, when an agent performs a “thumbs-up” gesture, in one culture it can be associated to an “approval” meaning, while in another culture, it can be associated to an “insult” meaning.

After being translated, the event is appraised in order to determine the emotional response of the agent. The idea that emotions are elicited by subjective evaluations (appraisals) of events or situations was first introduced by Magda Arnold. In the proposed architecture, the
cultural parameterisation of the agent is used to determine what is blameworthy and what is praiseworthy, which in turn generates moral emotions such as pride or shame. For further details on this process, refer to Mascarenhas et al. (2010). Norms are also linked to this appraisal. For instance, the violation of a highly salient norm is deemed to be blameworthy by the agents (more details in section D6).

For the agent to decide the next action it should perform there are two layers. The Deliberative layer consists of a continuous partial-order planner, which is constantly revising the agent’s plans and selecting an action to achieve its current intention, which may be to achieve an individual goal, to follow a cultural ritual or to fulfil a social norm. In contrast, the Reactive Layer, allows the agent to trigger fast reactions in response to a particular emotion (e.g. crying when feeling distress or frowning when feeling reproach for someone who violated a norm). Once an action is selected for execution, the action is performed in the virtual world through the agent’s effectors (devices used to produce a desired change in an object in response to input).

In order to model the notion of a moral circle (see section D.2), it is necessary that agents are capable of engaging in ritual activity with one another. In the proposed architecture rituals were implemented in a similar manner to how goals are implemented, yet with significant differences. Plan recipes used in traditional BDI architectures inspired the model for Rituals, but with a fundamental difference: traditional plans are based on technical activities (the focus is in the end result); whilst rituals are based on ritual activities with social meaning (the focus is in the sequence of steps). As such, a ritual has a set of roles associated with it and each role has one or more steps that must be performed following any specified ordering constraints. For more details on how Rituals have been implemented, please refer to Mascarenhas et al. (2009).
D4: Design Virtual Bar environment

The simulated agents of the FAtiMA model, which have been introduced in section D.1, require an adequate graphical realization for users to observe and interact with the agents.

We designed a virtual bar environment, depicted in Figure 3, using a realization (graphics and physical) engine for the creation of interactive 3D content, called Unity3D. The bar’s components (tables, chairs, walls…) and the characters were modelled, animated and rendered using Autodesk 3ds Max and exported to Unity. Our virtual bar environment consists of a set of assets that can be used to create different scenarios. These assets include objects, such as the characters, textures, animations and scripts... Scenarios are created by including assets and tuning their properties, such as, the initial number and position of characters, their roles, the actions they can perform, the animations linked to those actions etc.

However, since realization engines (including Unity) do not provide the appropriate abstraction level required by the agents, we use a simulation environment to integrate all FAtiMA’s processes. Therefore, by keeping the realization engine and the simulation environment decoupled, we foster reusability. For our simulation environment, we used the ION framework that allows us to link a character to a FAtiMA process (the character’s mind). Moreover, ION provides coherent access to information by changing the simulation state synchronously. This ensures the mediation of conflicts, offers both active and passive gathering of information, and allows dynamic changes in the simulation’s behaviour.
D6: Experiment with Virtual Bar environment

We generated a scenario using virtual agents with FAiMA and Unity3D, building on the work done in D.1, D.2 and D.4, to generate emotions that result from the evaluation (appraisal) of actions that are perceived to cause the fulfilment or violation of norms. This scenario occurs in a bar where there is a prohibition to smoke, as described in an introductory text. No ashtrays are present in the bar, and the only visual cue related to the norm is a no-smoking sign on the wall. The user’s is sitting opposite two friends at a table and there is also an elderly person sitting at another table. One of the user’s friends is a non-smoker and the other characters are smokers.

The overall plot is very simple: the smoker-friend talks for a while, then lights a cigarette and starts smoking (because its goal to smoke is always more important than the norm). The user is then presented with the choice to (a) continue the conversation, (b) mention the non-smoking sign and (c) ask the friend to stop smoking. The smoker then reacts to the choice and the scenario ends.

We made two versions of this scenario. The only difference between them was the belief, shared among the agents, regarding the salience of the no-smoking norm. In the first version, (the low-salient version) the salience of the norm is set very low. Since the salience is lower than its personal goal of smoking, the elder character smokes during the entire scenario. When the friend starts smoking the non-smoker character perceives that there is a norm violation and appraises the event as blameworthy. However, the blameworthiness is so low that it is not enough to exceed the threshold for triggering a reproach emotion, thus no emotional expression is made (see Figure 4-left).

Figure 4 - In the low-salience version (left image) the non-smoker does not react emotionally, while in the high-salient version (right image) the non-smoker reacts with a frown expression and the character in the background looks offended.

In the second version, (the high-salient version) the salience of the norm is very high. The importance of the smoking goal of the elder is now smaller than the norm, so he does not smoke. The smoker-friend still smokes since the norm, though important, is not as important as his goal to smoke. When the non-smoker friend perceives this norm violation, he appraises the action as very blameworthy, feels a strong reproach emotion, and reacts with a frown expression towards his friend. A negative react of the elder character is also portrayed (see Figure 4-right).

The scenario was evaluated with a small sample of users. The users saw one scenario and then filled in a questionnaire about their perceptions of the scenario seen. User’s perceptions of how important the norm was for the characters were significantly correlated with the
perception of the non-smoker character seated opposite the user being upset and being angry. Similarly, the perception of how acceptable it was for the characters to smoke inside the bar was significantly correlated with the belief in the non-smoker character being upset, being angry, and also being offended.

Although preliminary, the results obtained suggest that users were able to perceive a relationship between the emotions generated by our model, and the salience of the norm for the characters in the scenario. This is an important result because it indicates that generating these kinds of emotions from the norms specified in a multi agent environment, can help users to better understand the social context the agents are simulating.

The results suggest that users did relate the differences in the versions to the importance of the norm. As future work we plan to extend the model by introducing enforcing mechanisms, such as, punishments and sanctions. We also plan to conduct further tests using the model. For instance we would like to see if characters are less believable when they do not have an emotional response after a norm, which the user knows to be highly salient, is violated.

**D5: Design large-scale agent simulation for norm-culture-regulation interaction &**

**D7: Experiment with large scale simulation**

Work with large-scale agent-based models (containing hundreds of agents) has been aimed at exploring functional aspects of norms along three interrelated research directions:

1. Agents’ motivations, where cultural factors are implemented through the concept of values and their ordering.
2. Normative punishment, where negative and positive reinforcement of agents interactions through communication affects the salience of a norm, ultimately determining agents’ compliance with the norm.
3. Group dynamics, focused on the study of group composition and its effect on smoking behaviour.

These models examine complementary aspects of the development of a no-smoking norm in bars and employ varying levels of analysis. Model 1) looks at multiple bars, and separate environments within a bar premises, 2) currently models a single bar, but examines more closely the communication between agents and 3) abstracts from the particular events within bars to focus on the dynamics of friendship groups and the choice of bar type.

**Model 1: Ban Bar**

The BanBar was developed in order to give an explicit representation of the implicit motivations of the agent, in this case drives and values, and implement them in a computational model to show how these low-level forces can generate higher-level behavioural patterns.

Values are dispositions to choose one state of the world over another. Drives are a tool to represent an agent’s internal state and can be used to express agents’ motivations. They tell us what an agent needs. When a drive is out of balance, the agent has a need to satisfy. The BanBar model treats all implicit sources of motivation, both individual and social needs, as drives. Motivations derived from one drive can lead an agent to perform actions that are detrimental to another one of his drives.
The conflict between drives is captured by the model connecting the update functions for drives to different features of the state of the world, influenced by the agent’s actions. All trade-offs between connected drives are decided by referring to the agents’ values. Actions promoting values that are relatively more important for the agent will be selected for execution. This shall in turn feedback to the agent’s drives.

In a multi-agent system the state of the world realised by one agent acting to satisfy his needs might, and often will, have side-effects for other agents in the same environment. Side-effects emerge through the links made between the same behavioural or environmental features, and the drives of two (or more) agents. Side-effects can be both positive and negative, capturing the complexity of cooperation and conflict in social interactions.

The behaviour of the individual agents is regulated by the level of four basic drives and the agents, who differ from each other in regard to their attitude towards smoking, act in order to keep the highest number of drives in check. The drives are:

(a) Nicotine, controlling the need to smoke.
(b) Tolerance, which measures the strength of agents' dislike of ETS.
(c) Affiliation, controlling their sense of belonging.
(d) Comfort, a drive affected by the characteristics of the location of the agent.

Output from two groups of agents with differing cultural profiles, defined by the shared ordering of the agents' values, was used to analyse the features of the model.

The agents’ attitudes towards smoking are derived from their values and their strengths are also used to guide the agents' decision-making every time their behaviour impacts on conflicting drives and their related needs. Agents hold the following values:

1. **Health**: the realised benefits of healthy behaviours, together with the awareness of the risks associated with the negative consequences of being exposed to the unhealthy behaviour of others.
2. **Hedonism**: the general attitude to discount future consequences in favour of present rewards.
3. **Individualism**: represents the extent to which agents identify with the group; how much the agents let their own behaviour be affected by shared ways of life.
4. **Equality**: represents the importance agents attribute to differences in power, together with tendency to comply with the rules and obligations associated with the assigned roles.

The relative strength of the values of health and hedonism is responsible for the agents' attitudes toward smoking. The two cultures are defined by the relationship between the values. For each culture separate experiments were conducted.

The variables describing an agent are the four drives (including their activation levels and their thresholds), the four values (and their strengths), and three individual variables:

1. A location: every agent can be either **AtHome** or **ToThePub**,
2. A behaviour: either **Smoking** or **NonSmoking**, he can decide to do so either **Inside**- where other agents can react in a negative way to the presence of smokers or **Outside**- meaning an outdoor venue, a place less comfortable, but where the other agents would not complain about ETS.
3. A response: depending on the location of the agent, its attitude towards ETS and the activation level of the tolerance drive, this variable can be: **Voiced Reproach**, **Was Reproached**, and **Null**.
The system has proved to be responsive to changes in the motivational state of the agent. This feature is important for linking individual preferences to social outcomes. We have shown through our simulation how a structural asymmetry in the motivational state of the agent can produce significant changes in the overall social response of the system.

Our model’s structure permits analysis of the difference between the personal interest of an agent, and the regulatory interest of the agent. The personal interest is represented by the negative feedback between the nicotine and tolerance drives. Representing the regulatory interest of an agent would require the implementation of elements of normative reasoning, which should be grounded in the dynamics of values and drives described. A fuller account of this model shall be published as part of the MABS 2012 workshop.

**EMIL+ LSM**

A second large-scale model has been created to look at the factors favouring the spreading of the no-smoking norm. We suggest that to understand how the no-smoking norm can emerge and spread, it is necessary to model its impact on people’s minds. The simulation model implemented allows us to look at the factors influencing agents’ motivation to comply with the no-smoking norm. In particular, we focus on the role of the norm salience. With salience, we refer to the perceived degree of importance and strength of a norm. Norm salience is a complex function depending on several contextual and social factors, such as: the level of compliance and violation, the amount of non-punished violations, the frequency of punishment, the enforcement typology, and also the credibility and legitimacy of the punishing actor.

In the currently implemented scenario there is one bar where all agents go each evening. They choose independently whether to smoke or not when in the bar. Agents who are not smoking (non-smoking agents, hereafter) may then send them a message to say that they are violating a norm. The enforcement typology is a critical issue here as only dyadic interactions have previously been simulated, but this doesn’t transfer naturally to a bar environment. We are testing parameters currently, in order to observe the emergent effects of agents’ salience updates. These are based on a combination of: reproaches for smoking in the bar, and passively observed behaviour (smoking or not smoking).

Agents in the model have three goals:

A) The goal to smoke
B) The goal to avoid ETS
C) The goal to socialise

They also have a goal to respect the no-smoking norm, named their normative goal.

They update their belief about the salience of the norm, which directly influences the strength of their goal to follow the norm, and to see it followed by others, depending on:

1) Whether they smoked in the bar themselves.
2) The observed smoking behaviour of others.
3) Normative messages observed, received and sent.

The model is currently set up to allow the exploration of a number of more complex social interactions, including friendship groups and multiple types of normative interventions, combining both citing the existence of the norm, and actually punishing other agents.
further planned extension is to include the role of the bar-owner and to model multiple bars. This would allow the current Dutch situation of co-existing smoking and non-smoking bars to be reproduced.

**Simulation of the interaction of Group Dynamics and Smoking Behaviour**

A student undergraduate project investigating questions of interest to the SEMIRA project has been conducted at Wageningen. The aim of this model is to analyse the interaction of group dynamics and smoking behaviour in various cultural configurations. The setting of the model aims to replicate the formation of friendship groups among first year students in a university situated in a city where both smoking and non-smoking bars co-exist.

Agents are either smokers or non-smokers, with varying intensities of addiction or dislike of environmental tobacco smoke, respectively. They all wish to be in friendship groups, being happier the larger the group up to a certain limit. Their major decision in each simulation round is whether to stay in the same group and whether to keep the same smoking behaviour. Groups tend to go to bars that reflect the majority preference of the group with regard to smoking and ETS. Agents’ satisfaction with their current situation is determined by three elements: their number of friends, the extent to which they share their smoking preferences with other group members, and the bar most recently attended with their group. When the agents have low levels of satisfaction they may change either their group or their smoking behaviour. They do this in order either to find a group that conforms better to their preferences, or to better conform themselves to the preferences of their current group. This model allows us to examine social contagion in smoking behaviour and the segregation of groups based on such behaviour. A student thesis report on this model shall be finalised in June 2012. This will provide both a record of the work and a platform on which to build.

The characteristics of agents in the three models presented in this section will be based on both culture theory and data collected in D.3 for the next phases of model exploration, verification and validation.
4. Dissemination activities in the period in question (including list of publications where applicable)

Dissemination 2011

Conference proceedings


Conference presentations


Mascarenhas, S. Paiva, A. Degens, N. Mc Breen, J. Hofstede, G. J. (2011). "How should I say this?" Agents with culturally-appropriate verbal communication styles. *The International Workshop on Culturally Motivated Virtual Characters @ Intelligent Virtual Agent Conference.*

Conte, R. (2011). SEMIRA: Simulating the Emergent Impact of Regulations Across cultures *Complexity-NET session at the European Conference on Complex Systems*

Invited Lectures

Dissemination 2012

Conference Proceedings


Submitted Papers


Invited Lectures


Bibliography of previous work cited in report


