Emergence and enforcement of social behavior

Dignum, F. [1] and V. Dignum[1]

[1] *Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands*
Email: dignum@cs.uu.nl

Abstract: Complex social systems require an understanding of individuals, networks, institutions and cultures. Where regulations assure efficient coordination between individuals and groups and their needs and requirements, cultural and personality differences result in different behaviors and interpretations of global aims. Due to the dynamic mechanisms of social interaction, this results in a cyclic pattern with norms enforcing behavior and behavior generating norms. Current modeling tools often only consider one or the other direction of this cycle. I.e. one can either design systems where individual behavior is 'forced' to adhere to institutional norms, or analyze the emergence of global behavior resulting from 'free' individual behaviors.

In this paper, we propose a framework for the design and analysis of social systems that integrates a top-down (prescription of behavior) view with a bottom-up (description of behavior) view. Such models enable the understanding of patterns of influence between individual behavior and social norms and vice-versa. For the social organization, the framework is based on the OperA model that enables the representation of groups and organizations in a way that guarantees independence from individual architectures but enables individual interpretations and the emergence of regulation patterns. For the individual level, the framework uses the BRIDGE agent architecture, a modular cognitive agent architecture that extends BDI models with modules for the specification of personality types, urge hierarchies and different normative and cultural attitudes.

One of the most important premises of the framework is that both emerging norms as well as imposed norms do influence individual behavior but individuals can always violate norms (for personal reasons or in a particular situation). The consequences of violating norms might be an explicit sanction, but can also have only very implicit consequences. E.g. violating a speed limit may lead to a fine from the police, but violating the norm that one should always drive as fast as the speed limit might have as consequence that other drivers don't keep their distance or try to overtake in a dangerous places.

The way these possible consequences play a role in the decision forming process of an individual depends on the personality type of the individual, but also on the culture and even on the urges. An individual that forms part of a culture where conformity to majorities is very important will avoid deviating from any norm. However, in case an individual really needs to find a toilet very quickly he might disobey a speed limit. Finally, some individuals have personalities that put objective ideal situations over subjective judgments of other individuals. They might obey the speed limit if they agree with the motive behind it and will not hesitate to violate it (even if most people disagree) if they don't agree with the norm.

In our framework we want to study the relative importance of these different factors on the effectiveness of norms and regulations on societies under certain conditions. Which type of behavioral patterns will emerge given a certain distribution of individual traits and how will this affect the functioning of the society as a whole?

Keywords: norms, culture, personality, emergence, social behavior

1. INTRODUCTION

Social simulation is a good way to support our understanding of complex social phenomena (Moss, Edmonds, 2005). Especially in cases where many different social aspects interact it will be difficult to capture them all in one theory that can provide answers to all questions. We are interested in social behavior in the context of unusual circumstances such as a crisis situation caused by the flooding of (part of) a city. Simulations of these situations are meant to support authorities to better develop procedures which are difficult to practice in the real world. Important types of questions are, e.g. How will people react to orders for evacuation. Will people use buses instead of their own cars to avoid blocking exit routes? Will people start looting shops when an evacuation is taking place? What circumstances cause looting to happen? Are people willing to help persons that are immobile (sick, elderly, small kids,…)? …

All of these aspects have a crucial influence on the success of managing the evacuation. We do not pretend to answer all these questions at once, but we will concentrate on a particular aspect of social behavior, namely the way it is governed by norms. Norms can be seen as a kind of loose constraints on behavior. They are interesting because they are connecting the value system of a society with the (accepted) behavior in that society. Thus they are also a connection between culture and behavior, where we see culture as an expression of a certain set of values common to a group or society. The way that a person reacts to norms, however, does not only depend on the values and culture, but also on its personality, its goals and present circumstances.

Some norms can be made explicit and enforced through the law. E.g. people are obliged to follow the orders of a police officer. In a crisis situation this norm can be used to let the police order the people to evacuate their houses. Some norms are not so explicit. E.g. a norm to help your neighbors might not be enforced (directly), but relies on a system of reputation and social networking to have effect. Both norms are incentives to behave in a certain way, but rely on different ways to exercise their influence. In order to understand behavior governed by norms one should at least model these different types of processes in a different ways. A crucial point for this model is that norms are not pure constraints on behavior. It is always possible to violate a norm. As a consequence the model should provide a possibility for individuals to violate a norm and preferably couple this violation to a "good" reason.

Besides the fact that norms can be imposed on individuals from the outside, there are also norms that emerge from the individuals. E.g. it might become a norm to share your water and food during an evacuation. Or, in a negative way, it might become a norm to loot abandoned shops. In these cases, a pattern of behavior is started by one or more individuals and copied by more and more individuals in the society. Although unusual circumstances can contribute to provoke this emergence, one does not need unusual circumstances to initiate the emergence of norms. E.g. taking huge financial risks in order to secure short term high profit growth became a norm in banking in the early 2000's without any unusual circumstances. Apparently the behavior was so appealing when it was observed that bankers started to copy it indiscriminately. Often it is the case that such a behavior is difficult to counter by one individual, because that individual will have negative consequences while it is uncertain whether the unwanted behavior stops. A well-known example is that of talking loud in a group of people. If one person starts talking very softly, the only consequence is that no one hears him anymore. The only way of being heard is talking louder than the rest thus provoking the others to talk louder again, etc. Also in this case of norm emergence it seems important to be able to model the decision process that leads to the normative behavior and the reasons and situations when it does not lead to the normative behavior.

Given the aims of the simulation it seems reasonable to use an agent based simulation. In such a simulation we model individuals as software agents that we can ascribe their individual goals, beliefs, etc. This approach enables us to distinguish different types of motivations for different individuals, thus providing a more realistic simulation of the society. However, using a simulation purely based on individual agents would also not be correct. This would model a situation where there are no social structures that connect and limit individual autonomy. We will thus use a modeling framework that combines both a societal and an individual point of view. For the model of the society we use the OperA framework that can be used to model societal roles and their relations, the norms from the society and standard interaction patterns (Dignum, 2004). Within this framework the roles can be played by agents according to their own personality, goals, etc. The culture of a society is (partly) modeled in the social contract that ties an agent to a role. In some cultures this contract is very tight while in others it is very loose. E.g. playing the role of a "father" or "child" can carry very specific rights and duties in one culture while only have a biological meaning in another.

This paper is organized as follows. In the next section we will discuss the notion of norms and what are the necessary concepts that should be modeled in order to create a useful simulation. In section 3 we give an overview of the framework for simulation. In section 4 we expand the society level of this framework, while

in section 5 we describe the architecture of the agents that will populate the simulation. We finish the paper in section 6 with some conclusions and directions for future work.

2. NORMS

Norms in a sociological context can be seen as describing expected or standard ways of behavior (Conte, Castelfranchi, 1998). In philosophy the expected behavior is seen as the ideal behavior, describing how the world ideally should look like. These two viewpoints nicely fit with the idea that norms can be both imposed from outside and emerge in society. The sociological viewpoint talks about expected behavior, but does not indicate whether the expected behavior is also ideal. I.e. it does not presume a predefined ideal behavior against which actual behavior is measured. When we refer to norms from this perspective we will use the term *social norms*. The philosophical point of view does presume some ideal behavior. Norms in this point of view follow from a moral set of values. In this context, norms are described as obligations, prohibitions and permissions, using different types of deontic attitudes (and formalized using deontic logics). We therefore refer to norms from this perspective as *deontic norms*.

Social norms emerge from the bottom up. We therefore assume that they arise from observing and copying behavior of others or through interactions that are beneficial and thus lead to repetition of these behaviors. Deontic norms can be imposed in two main ways. The more abstract way is that they are explicitly made known to all agents in a society (like a law). Thus these norms are presumed to always be there and influence individuals. The second way deontic norms can be imposed is as effect of a request or order of another agent that is in some position of authority. I.e. the order of a policeman or request of a parent leads to a deontic norm to fulfill that request or order.

The main common denominator of both types of norms is that norms are not constraints on behavior that are always guaranteed to be fulfilled (e.g. after dropping an object it will fall down), but rather rules that can be *violated*. The possibility of violation of a norm makes them very subtle (and difficult to model) instruments to influence behavior. The violation of a norm that is imposed from outside is usual connected with some kind of (direct) sanction. In the example of crisis management the norm of using public transport for evacuation can be violated by using one's own car. However, the punishment for this violation might be that your car is impounded. The violation of emergent norms does not have an explicitly (known) direct punishment. E.g. violating the norm of helping your neighbors to evacuate might not lead to any punishment, but can lead to a low social status/reputation later on.

The above examples lead to the question of how and whether violations can be detected, i.e. the issue of norm enforcement. The way a norm is enforced has a large impact on the influence of the norm on behavior of individuals. Norms, and especially those which are imposed from outside (deontic norms), do not automatically lead to the behavior prescribed by the norm. Society should thus provide mechanisms to detect violations and ways to punish these violations. If too many violations go undetected the norm becomes less effective. E.g. if the police does not detect that people use their own car instead of public transport during the evacuation then more people might use their own car anyway. Similarly, if a violation is detected, but one can negotiate the fine with a police officer the effect will be different than when every car that is detected on the road is automatically impounded.

It might be clear that it will not be sufficient to just specify a set of norms as part of the simulation, but rather one has as well to include norm enforcement mechanisms as part of the norms. A second element also comes into focus now. When a norm is violated and this is detected, someone *should* react. That is, there is again a norm that indicates the expected or wanted behavior on violation. But this norm can also be violated. Suppose that if someone uses his own car for the evacuation, is detected by a police officer, but this police officer does not impound the car, but rather asks for a lift. So, there should be norms about enforcing the norms and also a mechanism to enforce these (secondary) norms again.

A last important feature of norms is that the set of norms can be inconsistent. This happens especially when norms have different origins. Suppose that an evacuee wants to follow the norm to help his neighbors and also the norm that he has to use public transport for the evacuation. These norms can conflict if his neighbors are elderly people that cannot ride a bus anymore, but need to be transported by car. The person then needs a mechanism to decide which norm is more important and follow that norm while violating the other. As we have shown in (Dignum, 2006) the problem of these conflicts is that they are not general (logical) conflicts between the norms, but only conflicts in very specific situations or even ways in which norms are fulfilled.

One could argue that social norms, in general should not be conflicting, because they emerge from the population and no conflicting patterns should emerge at the same time. However, many norms that emerge

are restricted to certain groups of people (or roles) in the society. Conflicts might arise when individuals are members of several groups in which conflicting norms emerge due to the different circumstances in those groups. E.g. being a member of my family we want to help our neighbors, but being a young male I might want to get away with my friends.

3. FRAMEWORK FOR SIMULATION

We are interested in developing simulation models to describe and analyze the behavior of individuals in a social normative setting. That is, how do people react in a certain situation and why do they react in that way? Following the discussion on norms in the previous section, we distinguish between social norms (that emerge from collective behavior) and deontic norms (that prescribe desired/required behavior). Following (Dignum et al., 2002) we use the notation $N_a\varphi$ to specify a social norm ('it is normal/it is expected that a will do φ') and $O_a\varphi$ to specify a deontic norm ('a is obliged[1] to do φ'). Furthermore, we use $G_a\varphi$ to represent that agent a has goal φ, and the operator $\geq_a$ to compare the relative priorities or preferences the agent has for different norms and goals.

The simulation framework integrates two modeling paradigms:

- Societal characteristics (goals, requirements, restrictions) are modeled using the OperA framework (cf. section 4.)

- Individuals (goals, decision making characteristics, personal characteristics) are modeled using the BRIDGE architecture (cf. section 5.)

The society model describes a given scenario in terms of goals, roles, interactions and norms. These are prespecified at design time representing the initial situation. Agent models specify the behavior of different individuals (or groups) according to their own desires, intentions and personality characteristics. The result is an organization that enforces a certain set of (deontic) norms, but is populated by autonomous individuals pursuing their own aims. When confronted with a norm on their behavior, each agent will deliberate on its desired course of action. This means that some agents may decide to obey the norm, while others choose for their own plan, accepting the possibility of sanctions (which they can again decide to ignore). Furthermore, agents are endowed with capabilities to observe and reason about each others' external behavior, and may decide to copy or reject it. In this way, social norms may emerge (e.g. an agent may decide to use its own car to evacuate the disaster area, in violation of the deontic norm, if it sees all its neighbors doing the same).

Furthermore, environment conditions will change which may also lead to different behavior of the agents. By separating between the prescribed desired specification and the actual resulting behavior, dynamic situations are simulated, actual behavior can be compared with the desired one. Note that at this stage we assume that this evaluation, and subsequent eventual changes to the prescribed society, are performed manually by the designer and external to the system.

Culture is defined as the set of shared attitudes, values, goals, and practices that characterizes an institution, organization or group. Culture both emerges from and sets the behavior of a group. E.g. in a collectivist culture an individual will look after the interest of the social group it belongs to and thus will very likely follow all social norms pertaining to that group. The other way around, if all agents follow the social norms this will lead to a collectivist culture. We follow the theory of Hofstede on cultural traits from which the following dimensions are relevant for our context (Hofstede, 2002).

1. **UAI** (Uncertainty Avoidance Index). A high index leads to a society where many interactions are regulated by deontic norms.

2. **IDV** (Individualism). A high level means that people put their own goals before that of the society.

3. **PDI** (Power Distance Index). A high index means that individuals are more likely to follow norms coming from authorities.

We equate cultural traits with the role of norms in the society. In the explorative part of our project, we will use the simulation model as a means to generate emergent patterns of behavior. We then analyze those patterns in the light of the characteristics of Hofstede's cultural dimensions in order to identify the culture of the group. We then determine which kind of interventions are most likely to be conductive of a given desired behavior given the identified culture. In the validating phase of the project, we will implement those

---

[1] We use only the deontic operator O (obligation). All other deontic relations can be reduced to an obligation.

interventions in order to check that indeed they support that desired behavior. For instance, it is expected that in individualistic societies, people will more likely choose to evacuate the disaster area using their own can despite the deontic norm prohibiting it, that is, $G_a take\text{-}car \succeq_a O_a use\text{-}bus$. In such society, enforcement of this norm will require the presence of many police or military that have the power to apply sanctions to those using their car. In societies with a high power distance, orders by leaders are obeying without question. That is, the utterance *"take the bus"* by a police officer, will become an high priority obligation which is preferred over own goals, in a collectivist society: $O_a use\text{-}bus \succeq_a G_a take\text{-}car$.

4. OPERA

The OperA model (Dignum, 2004) for agent organizations enables the specification of global requirements and objectives, and at the same time allows participants to have the freedom to act according to their own capabilities and demands. OperA considers agent organization models as having at least two description levels. At the abstract level, which can be seen as a receipt for collective activity, organizations are described in terms of roles, their dependencies and groups, interactions and global norms and communication requirements. The concrete level is a possible instantiation of the abstract organization, by populating it with real agents that play the roles and realize interactions (Vazquez-Salceda et al., 2005). Organizational roles and responsibilities represent general, long-term guidelines while operational control involves specific short-term agreements among agents to perform specific activities for specific time periods (Sims et al., 2004).

OperA provides a layered formal specification language for agent organizations:

– The organizational model (OM): contains the description of the roles, relations and interactions in the organization. It is constructed based on functional requirements of the organization

– The social model (SM): describes the population of the organization, linking agents to roles such that organizational rules are translated to contracts for the agents fulfilling the roles

– The interaction model (IM): describes interactions given organizational model and agents

A generic methodology to analyze a given domain and design its OperA model is described in (Dignum et al., 2004). Our simulation framework uses the OperA OM formalization for the specification of the social structures and interactions. These structures are to be used by the individuals to reason about their role and position in the society, their influences and expectations. Space limitations do not allow for a complete description of the model. In the following, we suffice with a few small examples.

The OM specifies an agent society in terms of four structures: social, interaction, normative and communicative. The social structure specifies objectives of the society, in terms of roles, groups and dependencies that govern coordination. In the disaster evacuation simulation, roles are e.g. *burgers*, *elderly*, *commander, rescuers* and *police*. *Commander* is dependent on *police* to organize the evacuation, *burgers* are dependent on the *police* to know how to leave the area, *elderly* are dependent on the *rescuers* to be evacuated, etc. The interaction structure gives a partial order of the scene scripts that specify the way dependencies between roles are to be achieved. Interaction scene scripts are specified in terms of landmarks, which describe how a result should be achieved, that is, describe the states that must be part of any protocol that implements the interaction scene. E.g. the interaction scene describing the departing of burgers may describe that people are to wait in a pre-specified location, that police gets buses to that location and the order of who gets in which bus, etc. Norms and regulations are specified in the normative structure, expressed in terms of deontic rules for roles and interaction scenes. Finally, the communicative structure specifies the domain concepts and communication illocutions. The way interaction occurs in a society depends on its aims and characteristics, and determines the way roles are related to each other, and how role goals and norms are 'passed' between related roles. For example, in a authoritarian society (PDI), a higher placed role has the power to delegate its goals to its children roles, which then take those as obligations.

5. THE BRIDGE ARCHITECTURE

The agents that are used in the simulation framework are based on the BRIDGE architecture (Dignum et al, 2009). There is not enough space to describe the entire architecture, but for this paper it suffices to know that the agents follow a traditional cycle of sense-reason-act. In the *sense* phase they check the situation of the world and update their beliefs accordingly. These updates can be about facts that became true or are no longer true, but these can indirectly also affect the relevance of a goal the agent is pursuing or a norm that might no longer be valid for him.

In the *reason* phase the agent will deliberate on which action to take next. In our case the decision is based on the following elements:

  – The goals of the agent, G_a
  – The imposed norms, or obligations, O_a
  – The social norms, N_a
  – The beliefs of the agent, B_a

Besides having Goals, Beliefs and Norms the agents also exhibit personality and are influenced by culture. Obviously, there is a connection between norms and culture. At present we assume that the culture reasoning module of the agent architecture determines the way an agent chooses how to handle norms. Whenever the goals and norms are all aligned the agent will just follow his own goals until it achieves them or they are no longer (believed) achievable. The more interesting case arises when the goals of the agent are conflicting with the norms and/or the social norms conflict with the imposed norms. In (Dignum et al., 2002) we discussed a framework to reason about these situations. In the present framework we take a somewhat simpler approach (which will make the simulation more efficient) and assume that culture and personality of the agent together determine whether an agent will follow a norm or its own goal first

We use the MBTI framework to characterize personalities (Beerens, 2006) and make the (very simplified) assumption that personality does not have any impact in the case of collectivist societies (i.e. IDV is very low). In a collectivist society, the agent will follow first the imposed norms, then the social norms and finally its own goals. If the culture is individualist, personality does play a role. The important MBTI dimensions for our context are: Sensing/Intuition and Thinking/Feeling. In this case, if an agent is "thinking" and "intuition" it will first choose to follow the imposed norms, then its own goals and finally the social norms. An agent with opposite personality (sensing, feeling) will first look what other agents are doing and follow their lead. Thus it will follow social norms first, then if most agents follow the imposed norms it will also follow them and its own goals come last. However, if most agents around him follow their own goals before the imposed norms it will also follow its own goals before the imposed norms.

The above mechanism gives a crude way to decide between different sources when they conflict. However, it does not indicate ways to cope with conflicts within the different sources. For this aspect we assume that the norms and goals have preferences that will be adjusted according to experience and situation. The preference to follow an imposed norm depends on two elements: how big is the chance of a violation being detected and how big is the penalty on violation. If the chance of detection is high and the penalty severe than the preference to follow the norm is high. The chance of detection can be set to a predetermined value and adjusted according to experience. The agent might base its experience on its own violations, but can also check whether violations of other agents of the norm are detected.

The preference of social norms is completely determined by how many other agents are observing the norm. This is a gross simplification of the actual process, but we use it as a useful starting point for our simulation framework. So, we say that a social norm is more preferred if it is observed more often. Note that these observations might be very local!

The preference of goals depends on two things. First, how much did the agent already "invest" to achieve the goal and expects it to still need to do before achieving it. Secondly, how much utility does the agent get from the goal. So, a goal that has been pursued for a long time and is almost achieved and also has a high utility will have a high preference, while a goal with low utility for which nothing has been done yet will get low priority. Putting all the above elements together the preferences of a Sensing, Thinking agent from an individualistic culture will look something like figure 1.
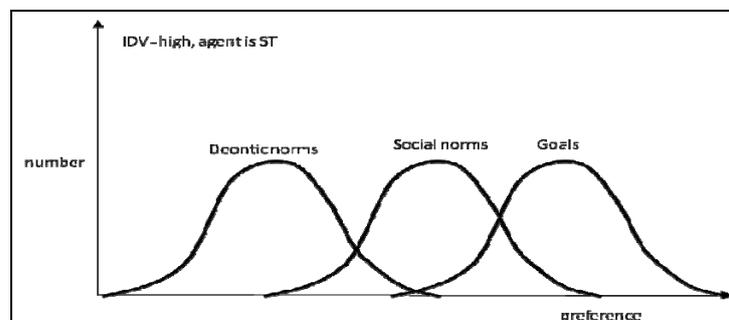

Figure 1: Preference ordering for agent in individualistic culture

Note that not all goals are preferred to all social norms and not all social norms to deontic norms. This is due to the experiences of the agent that shift the preferences of norms and goals upwards and downwards.

Using this framework we are able to simulate the influence of norms, culture and personalities on the behavior of people in crisis situations. This allows us e.g. to test whether in a collectivist culture with high power distance index people will indeed use public transport for the evacuation when there is enough police around to order the people to do so at the start of the evacuation. And that their own goal of taking the car might prevail otherwise.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a framework with which we can simulate the influence of different types of norms on individual behavior. By endowing agents with mechanisms to differentiate their reasoning process, and the ability to decide or not to comply with a given norm, we are able to simulate the evolution of collective behavior. Because agents are able to observe each other behavior, including the violations of norms by others, and the consequences thereof, their normative preferences change, which is reflected in a changing collective behavior, which will determine culture changes. Furthermore, we can use the same framework to analyze the effectiveness of norm enforcement interventions on different cultural settings.

An important direction for future work, in which we are currently engaged, is to fully implement the framework and to model different individual behaviors and normative settings. Besides the crises management scenario within the ALIVE project, we are planning to use the framework to the simulation local economies. The implementation will be based on the ABC-Lab tool, presented in (Tranier et al., 2009) that integrates 2APL agents (Dastani et al., 2007) in a Repast environment[2].

REFERENCES

Berens L. (2006). *Understanding Yourself and Others: An Introduction to the 4 Temperaments 3.0.*, Telos Publications.

Conte, R and Castelfranchi C. (1995). *Cognitive and social action*. London: UCL Press. pp. 74-118.

Dastani, M., Hobo, D., Meyer, J.J. (2007). Practical Extensions in Agent Programming Languages. In *AAMAS'07*, ACM Press.

Dignum, F., Kinny, D., Sonenberg, L. (2002) From Desires, Obligations and Norms to Goals. *Cognitive Science Quarterly*, 2(3-4):407-430.

Dignum, V. (2004). *A Model for Organizational Interaction: Based on Agents, Founded in Logic*. PhD dissertation, Universiteit Utrecht. SIKS dissertation series 2004-1.

Dignum, V., Dignum, F., and Meyer, J.J. (2004). An agent-mediated approach to the support of knowledge sharing in organizations. *Knowledge Engineering Review*, 19(2):147-174.

Dignum, F. (2006). Norms and Electronic Institutions. *Deontic Logic and Artificial Normative Systems.* LNAI 4048, pp. 2-5, Springer.

Dignum, F., Dignum, V., Jonker, C. (2009): Towards Agents for Policy Making. *Multi-Agent-Based Simulation VIII*. LNAI 5269, Springer.

Hofstede, G. (2001). *Culture's Consequences, Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Thousand Oaks CA: Sage Publications.

Moss, S., Edmonds, B. (2005). Towards Good Social Science. *Journal of Artificial Societies and Social Simulation* 8(4).

Sims, M., Corkill, D., and Lesser, V. (2004). Separating domain and coordination knowledge in multi-agent organizational design and instantiation. In *Proc. Agent Organizations: Theory and Practice*, volume WS-04-02. AAAI.

Tranier, J., Dignum, V. and Dignum, F. (2009). A Multi-Agent Simulation Framework for the Study of Intermediation. *Submitted*

Vazquez-Salceda, J., Dignum, V. and Dignum, F. (2005). Organizing Multi-Agent Systems. *Journal of Autonomous Agents and Multi-Agent Systems*, 11(3):307-360.

---

[2] http://repast.sourceforge.net/