Proceedings of the First International Workshop on Agent-Based Modeling for Policy Engineering (AMPLE 2011)

Francien Dechesne[1], Virginia Dignum[1], Amineh Ghorbani[1], and Julian Padget[4]

[1] Department of Technology, Policy and Management, Delft University of Technology, The Netherlands
[2] Department of Computer Science – University of Bath, United Kingdom

Preface

This volume contains the papers presented at AMPLE2011: the 1st International Workshop on Agent-based Modeling for Policy Engineering@ AAMAS 2011 held on May 2, 2011 in Taipei, Taiwan.

The goal of AMPLE is to discuss the role of multi-agent systems and artificial societies onpolicy making and institutional analysis.

Socio-technical systems are complex adaptive entities that require the engagement of social and technical elements in an environment to reach certain goals. In order to understand, analyze or design such systems, advanced tools are required. One of the major tools for understanding socio-technical systems is agent-based modeling. In recent years, social scientists, including economists and policy makers, have been using agent-based models to tackle their problem domains. Building artificial societies by combining the multi-agent systems view and domain knowledge has become a challenge, because of the complexity involved. AMPLE aims to provide a shared platform for the modeling and the organization design communities in order to discuss mutual effect of models on societies and vice versa. More than mere design or analysis tools, agent-based models and simulations can impact the way groups, organizations, and societies are managed and adaptation is shaped. Computational models are increasingly used for the evaluation of institutional decisions and policies and the what-if analysis of potential changes such as re-engineering of the tasks, structures, innovations and societal effects of policies.

We are grateful to Catholijn Jonker for her invited presentation at the workshop.

These proceedings have been produced using tools made available by EasyChair.

We want to conclude this preface by extending our thanks to the members of the program committee of the AMPLE 2011 workshop that were willing to review the papers in a very short time span and also of course to all the authors for their submissions.

March 10, 2011
Delft

Francien Dechesne
Virginia Dignum
Amineh Ghorbani
Julian Padget

Table of Contents

Program Committee

Frances Brazier	Delft University of Technology. the Netherlands
Rosaria Conte	IRC, Italy
Nuno David	Universidade de Lisboa, Portugal
Frank Dignum	Utrecht University, the Netherlands
Bruce Edmonds	Manchester Metropolitan University Business School, UK
Andreas Ernst	University of Kassel, Germany
Armando Geller	George Mason University, USA
Nigel Gilbert	University of Surrey, UK
Maria Gini	University of Minnesota, USA
Gert Jan Hofstede	Wageningen University, the Netherlands
Catholijn Jonker	Delft University of Technology, the Netherlands
Jens Pfau	University of Melbourne, Australia
Nicole Ronald	Eindhoven University of Technology, the Netherlands
Jaime Sichman	University of Sao Paulo, Brazil
Barry Silverman	University of Pennsylvania, USA
Liz Sonenberg	University of Melbourne, Australia
Yao-Hua Tan	Delft University of Technology, the Netherlands
Jeroen Van Den Hoven	Delft University of Technology, the Netherlands

Additional Reviewers

B

Brandao, Anarosa

C

Casare, Sara

Development and Application of Rich Cognitive Models for Policy Making

Catholijn Jonker[1], Frank Dignum[2], and Virginia Dignum[1]

[1] Delft University of Technology, The Netherlands
[2] Utrecht University, The Netherlands

Policy simulation models are important tools for predicting whether the introduction of a policy will have the desired effects. Policies are incentives for individuals to change their behaviour. However, the way people change their behaviour might not be the intended change in behaviour of people. People interpret the policies in the context of their own situation, according to their own personality and influenced by the social networks they belong to. Modelling the effects of introducing a new public policy is therefore as important as it is complicated. That is, the desired global (or societal) behaviour that motivates the introduction of a policy depends on the individual behaviour as well as on the dynamics of the environment. The pure macro-models that are currently used average over all the individual behaviour and thus can miss important effects that occur due to individuals reacting to each other's behavioural changes.

To remedy the fact that macro-models do not pay sufficient attention to emergent behaviour as a result of individual level changes, agent-based models have been introduced. Agents being autonomous software entities that perceive and act in their environment can be used to model individual behaviour. However, most current Agent-Based Social Simulation (ABSS) models are based on agents with rather simple cognitive capabilities. They typically do not model the real needs and personalities of people or the multi-cultural background which are important factors that shape the effects of a policy, or do only partially so. Models are needed that are able to account for all of these factors and for the related diversity.

The aim of this project is to develop an incrementally complex model to describe the influence of policies on the behaviour of the individuals. This requires formal frameworks to describe policies and rich cognitive agent models, collective behaviour, and to bridge the gap between societal and individual level. The scientific challenges are the following: No integrated models of personality and culture exist, and for a good reason: no linear combination suffices. Our ultimate aim is to develop an integrated model of the combined influence of personality, culture, and social influence on the individual's decision making on behaviour. We will build on our previous work on the influence of culture on trade processes and on our experience on a computational model of the influence of personality on an agent's reasoning. The model will also handle the influence of collectives on individual behaviour by considering the social groups one belongs to as input to the reasoning process.

The Role of MAS as a Decision Support Tool in a Water-Rights Market

Vicente Botti[1], Antonio Garrido[1], Juan A. Gimeno[1], Adriana Giret[1], Pablo Noriega[2]

[1] DSIC, Departamento de Sistemas Informaticos y Computacion,
Universitat Politècnica de València,
[2] IIIA, Artificial Intelligence Research Institute,
CSIC, Spanish Scientific Research Council,
{vbotti,agarridot,jgimeno,agiret}@dsic.upv.es
pablo@iiia.csic.es

Abstract. Water is getting a more and more scarce resource, which motivates the idea of designing a framework where water rights may be exchanged more freely, thus leading to a more efficient use of water. In this paper, we present a water-right market embedded within a decision support tool designed as a multi-agent system. To our knowledge, there are many sophisticated decision support systems for water management from a hydrological perspective, but they lack of a social perspective. Using a multi-agent system allows us to design intelligent agents that mimic humans, thus implementing different factors such as (mis)conducts, trust criteria and users willingness to water-right trading. Within a decision support tool, we can dynamically change norms and regulation at no cost, and explore the impact on the evolution of the market. Mixing all these elements together, we have implemented our *mWater* system as an electronic institution that demonstrates very appealing for decision taking and policy makers to test: i) how regulations and norms may modify the users' behaviour, and ii) how the quality indicators of the market are affected.

Keywords: Applications of multi-agent systems, decision support, simulation tools, electronic institutions

1 Introduction

Water scarcity is becoming a major concern in most countries, not only because it threatens the economic viability of current agricultural practices, but because it is likely to alter an already precarious balance among its different types of use: human consumption, industrial use, energy production, navigation, etc. Underneath this emergent situation, the crude reality of conflicts over water rights and the need of accurate assessment of water needs become more salient than ever.

It has been sufficiently argued that more efficient uses of water may be achieved within an institutional framework where water rights may be exchanged more freely, not only under exceptional conditions but on a day-to-day basis [18], similarly to a traditional goods market. In hydrological terms, a water market can be defined as an institutional, decentralized framework where users with water rights (right holders) are allowed to voluntarily trade them, always fulfilling some pre-established norms, to other users in exchange of some compensation, economic or not [18]. Additionally, when there exist incentives for an efficient use of water allotment, it is time for a straightforward extension to other types of stakeholders that promote trading for non-irrigation uses, such as industrial uses, aquiculture or leisure, thus improving market conditions and efficiency in water use.

This paper concerns the application of a regulated open Multi-Agent System (MAS), *mWater*, that uses intelligent agents to simulate a flexible water-right market. Our simulator focuses on demands and, in particular, on the type of regulatory (in terms of norms selection and agents behaviour), and market mechanisms that foster an efficient use of water while also trying to reduce conflicts among parties. In this scenario, a MAS plays a vital role as it allows us to define different norms, agents behaviour and roles, and assess their impact in the market, which helps enhance the quality and applicability of its results as a decision support tool.

2 Problem Overview

Water-right markets allow rapid changes in allocation in response to changing demands for water and stimulate investment and employment, as users are assured of access to secure supplies of water. Because of water's unique characteristics, such markets do not work everywhere, they are not homogenous as present different organisation schemata, nor do they solve all water-related issues [18]. Therefore, it is essential to design appropriate water laws and regulate, either privately or publicly, the users' actions, interactions and their eventual trade. By doing this, water markets effectively address rising demands for groundwater and for surface water found in rivers, lakes and canals. In that line, international experience in USA (particularly California), Chile, Australia or Mexico has demonstrated that (formal) water markets can improve the economic efficiency of water use and stimulate investment [18].

The willingness of irrigators to buy or sell water highly depends on the difference between the price of water and net revenue each farmer expects to earn by irrigating. Thus, for a given price of irrigation water, a farmer would be willing to purchase water if (s)he expects a unit of water to generate more incomes than it costs. If another farmer expects a unit of water to earn less that (s)he could sell it for, (s)he might want to sell it thus originating the trading process. But it is not always a matter of price expectations, but also of regulation. The emphasis on regulatory aspects is motivated by the fact that the main objective policy makers have in mind is to achieve an adequate behaviour of users to ensure the success of the market. And regulation is the main tool that policy makers have to modify behaviour by means of: i) government laws, ii) basin or local norms, and iii) social norms. However, in practice, users are prone to achieve "order without law" or, at least, to preserve their practices within the established regulation, whereas policy makers adapt regulation to guide users in a constantly changing environmental and political media. But adapting this regulation and taking the best decisions on the design of the norms for the market are difficult and delicate tasks, and cannot be freely applied in the real world. Also, as the result of enforcing norms in a water market is unknown a priori, a MAS-based simulation tool shows very appealing to analyse the impact in the users, the market itself and its success.

3 Limitations in Current Approaches

Literature abounds in examples of sophisticated basin simulation models, particularly decision support systems for water management [1, 12], sustainable planning of water volumes and hydraulic resources [5, 14], and use of shared visions for negotiation and conflict resolution [11, 17]. From a hydrological perspective, these works have successfully bridged the gap between the state of the art in water-resource systems analysis and the usage by practitioners at the real-world level. Clearly, operational management has benefited from the advances in computing and its applications, particularly in modelling, software engineering and simulation techniques, thus improving the operating rules for efficient water allocation. However, the gap can still be considerably narrowed from a social perspective, which is an important limitation nowadays. The underlying idea is not only to consider hydraulic factors, such as river basins, water demands, pumping flows, etc., but also different norms typology, human (mis)conducts, trust criteria and users willingness to agree on water-right trading, which may lead to a win-win situation in a more efficient use of water. This requires the use of intelligent agent technology, including trust, cooperation, argumentation, negotiation and, in general, agreement technologies [16]. Agreement is a crucial concept that helps human agents to cope with their social environment and deal with any type of human interactions. And how to support and promote agreements in water markets is missing in current approaches, which is also an indication of ineffectiveness.

An additional limitation is imposed by current legislation. In many countries, the norms and their regulation are very strict, which do not allow a full and flexible market. For instance, Spanish regulation is too restrictive and does not let final stakeholders participate freely in the modelling and water-right trading process. In particular, the Water Law of the National Hydrological Plan regulates the power of right holders to engage in voluntary water transfers, and of basin authorities to setup water markets, banks and trading centers for the exchange of water rights, but *only* in cases of drought or other severe scarcity problems. This means that the number of water-right transfers is practically non-existent in reality, reduced to few eligible participants and limited to very short periods. Also, in some tentative scenarios aimed at forming water markets the results were unsatisfactory because: i) water-right holders were reluctant to participate in the market, and ii) regulation and legally binding conditions were too tight.

Finally, from a performance standpoint it is unclear which is the best quality indicator of the market because it cannot be measured in terms of just one factor; we need a multiobjective analysis that comprises multiple criteria based on differing objectives, responsibilities and interests among the stakeholders and institutions involved in the market. Factors such as economic development, social welfare, environment preservation, agricultural self-sufficiency and financial feasibility must be considered. All in all, these issues can be achieved at a high global cost which is based on industry structure, population, quality standards, investment for new treatment plants, and policy for water allocation among agriculture, industry and domestic sectors.

4 Why Use a MAS as a Simulation Tool for Decision Support?

Agent technology and multi-agent systems have been successfully applied to problems such as manufacturing, medicine, aero-space, e-commerce, etc. when developing high-quality and industrial-strength products. One of the most promising domain applications of MASs is the simulation of complex real life systems that emulate social behaviour and organizations, where a MAS is used as a powerful tool that mimics real world behaviours of autonomous agents, i.e. individuals and societies [17]. In this way, complex behavioural patterns are observed from simulation tests in which autonomous entities interact, cooperate, and/or compete to achieve a set of goals. This offers several advantages: i) the ability to model and implement complex systems formed by autonomous agents, capable of pro-active and social behaviour; ii) the flexibility of MAS applications to add and/or delete computational entities, in order to achieve new functionalities or behaviours in the system, without altering its overall structure; and iii) the ability to use notions such as organization, norms, negotiation, agreement, trust, etc. to implement computational systems that benefit from these human-like concepts and processes among others [16].

In the specific domain of water-right management there is a need to foster a more rational use of the resource. And it is agreed that this may be addressed by creating an efficient market of water rights that coexist in a complex, social and legal framework [18]. Although most water management models are based on equational descriptions of aggregate supply and demand in a water basin [14], only a few include an agent-based perspective. Under this perspective, we explore an approach in which individual and collective agents are essential components because their behaviour, and effects, may be influenced by regulation and policy-making. The idea is to follow the thread of MAELIA (`http://www.iaai-maelia.eu`) and NEGOWAT projects (`http://www.negowat.org`) that simulate the socio-environmental impact of norms for water and how to support negotiations among stakeholders in areas where water conflicts arise.

From a technical perspective, there are several approaches to implement MAS applications. Some approaches are centered and guided by the agents that will populate the systems, while others are guided by the organizations that the constituent agents may form (for a literature review please refer to [3]). Other approaches rely the development process on the regulation that defines the MAS behavior, which is usually encoded as an Electronic Institution (EI) [7, 9, 13]. We are interested in this latter approach due to the requirements imposed by the environment. In particular, *mWater* —from the perspective of a MAS simulation tool— implements a regulated market environment as an EI, in which different water users (intelligent agents) trade with water rights under different basin regulations. With such a tool, water-policy makers can easily predict and measure the suitability and accuracy of new or modified regulations for the overall water market, i.e. more transfers, fewer conflicts, increased social satisfaction of the water users, etc., before applying them into the real floor. At the same time, it is a tool to manage the water resource in an effective way, both in the short and medium term. All in all, not only is it an aid for a better understanding of the physical and management aspects of the water-resource system in question, but it is also a good tool for data organization and communication among the different teams of the basin administration.

5 Our Approach

mWater uses a multi-tier architecture, as depicted in Fig. 1 [8]. In addition to the three typical tiers of presentation, business and data persistence, we have a module that represents the EI for *mWater*. This way, the construction of *mWater* consists of four stages: i) modelling the system as an EI; ii) designing the information system based on a database of the entire electronic market and basin structure (persistence tier); iii) implementing the agents (business tier); and iv) creating the GUI for simulation tool (presentation tier), which are described next.

5.1 Modelling the system as an EI

Electronic Institutions (EI) are computational counterparts of conventional institutions and represent a set of conventions that articulate agent interactions [7, 10]. In practice, they are identified with the group of agents, standard practices, policies and guidelines, language, documents and other resources —the organization— that make those conventions work. EIs are engineered as regulated open MAS environments in the sense that: i) the EI does not control the agents' decision-making processes, and ii) agents may enter and leave the EI at their own will, which is essential in a market.

An EI is specified through: i) a *dialogical framework* which fixes the context of interaction by defining roles and their relationships, a domain ontology and a communication language; ii) *scenes* that establish interaction protocols of the agents playing a given role in that scene, which illocutions are admissible and under what conditions; iii) *performative structures* that, like the script of a play, express how scenes are interrelated and how agents playing a given role move from one scene to another, and iv) *rules of behaviour* that regulate how

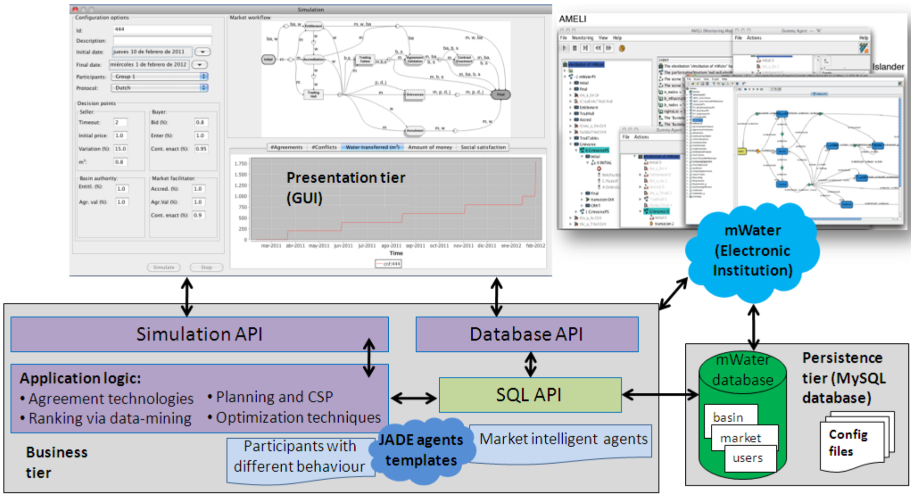


Fig. 1. Multi-tier architecture of the *mWater* decision support tool

commitments are established and satisfied. We have used this specification and modelled *mWater* as an EI. *mWater* uses the notation for the conceptual model introduced in [2], whereas for the actual specification and implementation we use the EIDE platform[1].

The *mWater* institution is specified through a nested performative structure with multiple processes, as depicted in Fig. 2. There are five agents' roles: i) guests, i.e. users before entering the market; ii) water users, i.e. the guests that have valid water rights; iii) buyer/seller, thus representing the particular role the water user currently joins for the market; iv) third parties, i.e. those water users that are direct or indirectly affected by a water transfer —usually conflicting parties; and v) market facilitator and basin authority, thus representing the governing roles of the market. The top structure describes the overall market environment and includes the following elements:

– Entitlement, which represents the bootstrap routine to give access to the market to those water-right holders who prove they are entitled to trade because: i) they have an existing right, or ii) a new right is created by the *mWater* authorities and an eligible holder gets it granted.

[1] EIDE is a development environment for Electronic Institutions, implemented at the IIIA (`http://e-institutor.iiia.csic.es/eide/pub`). It consists of a set of tools that support all the stages of EI engineering, namely: i) ISLANDER, a tool for EI specification; ii) aBUILDER, a tool to support the automatic generation of agent (code) skeletons from ISLANDER specifications; iii) the AMELI middleware that handles the enactment of the institution; and iv) SIMDEI, a testing and monitoring tool.

- Accreditation, which allows legally entitled water-right holders to trade by registering their rights and individual data for management and enforcement purposes.
- TradingHall, which represents a nested performative structure. It basically provides information about the market and, at the same time, allows users and trading staff to initiate trading and ancillary operations. Metaphorically speaking, it represents a place where participants stay to be informed and reconvene after leaving a trading table or grievance process.
- TradingTables, which represent a nested performative structure and the core of our market. It allows a market facilitator to open a new trading table whenever a new auction period starts (i.e. automatically) or whenever a right-holder requests to trade a right (i.e. on demand). Our implementation accommodates different trading mechanisms and negotiation protocols, such as Dutch auction, English auction, standard double auction and blind double auction with mediator negotiation, but new negotiation protocols can be easily included.
- Agreement Validation, which validates agreements on water-right transfers according to the market regulation. More particularly, staff have to check whether the agreement satisfies formal conditions and the hydrological plan normative conventions.
- Contract Enactment, which represents the signature among parties involved in a norm-abiding agreement, thus making the agreement active.
- Grievances, which represent a nested performative structure. It allows external stakeholders to initiate a grievance and conflict resolution procedure that may overturn or modify an active agreement. Even if there are no grievances that modify a contract, parties might not fulfill the contract properly and there might be some contract reparation actions.
- Annulment, which deals with anomalies that deserve a temporary or permanent withdrawal of water rights.

The essence of our market relies on the Trading Tables and Grievances structures. The former implements the trading process itself, which entails the participation of the buyer/seller and staff agents. The latter is necessary to allow normative conflicts to be solved within the *mWater* institution, particularly when the agreement execution turns conflicting with third party agents. In our approach, we include a framework for conflict resolution based on grievance protocols in which alternative dispute resolution (ADR) mechanisms are included in order to settle the conflicts internally in the market [15]. In this framework, any grievance process primarily involves negotiation like in any Trading Table (with or without mediation) and a arbitration procedure, or a combination of both. This way, the result of a conflict resolution can be an agreement among the conflicting parties by which they voluntary settle the conflict, or a decision from the arbitrator (a neutral third party) which is final, and binding to both conflicting parties.

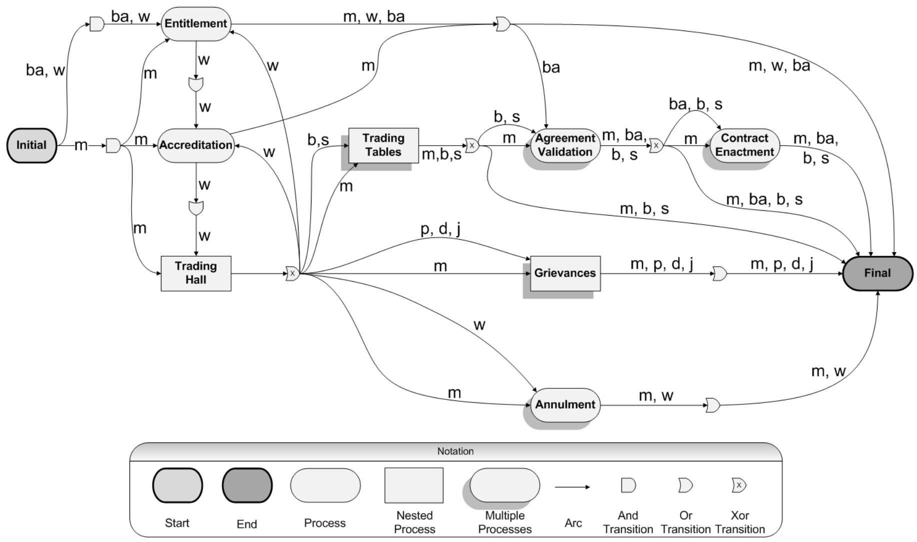


Fig. 2. *mWater* performative structure. Participating roles: g - guest, w - water user, b - buyer, s - seller, p - third party, m - market facilitator, ba - basin authority. See [4] for further details

5.2 Persistence Tier: Database Design

mWater implements the persistence tier by means of a MySQL database with over 60 relational tables in which historical data is stored. In essence, we have three views that comprise the basin, market and grievance structure (see Fig. 3). In the first view we model all the information about the nodes, connections, users, norms and water-right definition. In the second view we model information related to the entire market, including the trading tables and their protocols, the water rights to be traded, participants, agreements and contracts that can be signed. Finally, in the third view we model the information about the legislation and conflicts that may appear after an agreement or contract and the mechanisms for solving such a conflict, that is the negotiation stage or arbitration procedure. This way, policy makers can run the whole market with real and simulated data for drought periods, rainfall, norms and users, and analyse how they affect the final results and the number of grievances. Furthermore, all the changes in the market are registered in the database to provide statistical information and/or distributions to the policy makers, which are essential in a decision-support tool.

5.3 Business Tier: Implementation of Agents

mWater implements a schema of agents that include both the internal and external roles. Broadly speaking, there is a JADE (Java Agent DEvelopment Framework, `http://jade.tilab.com`) definition for each class that represents

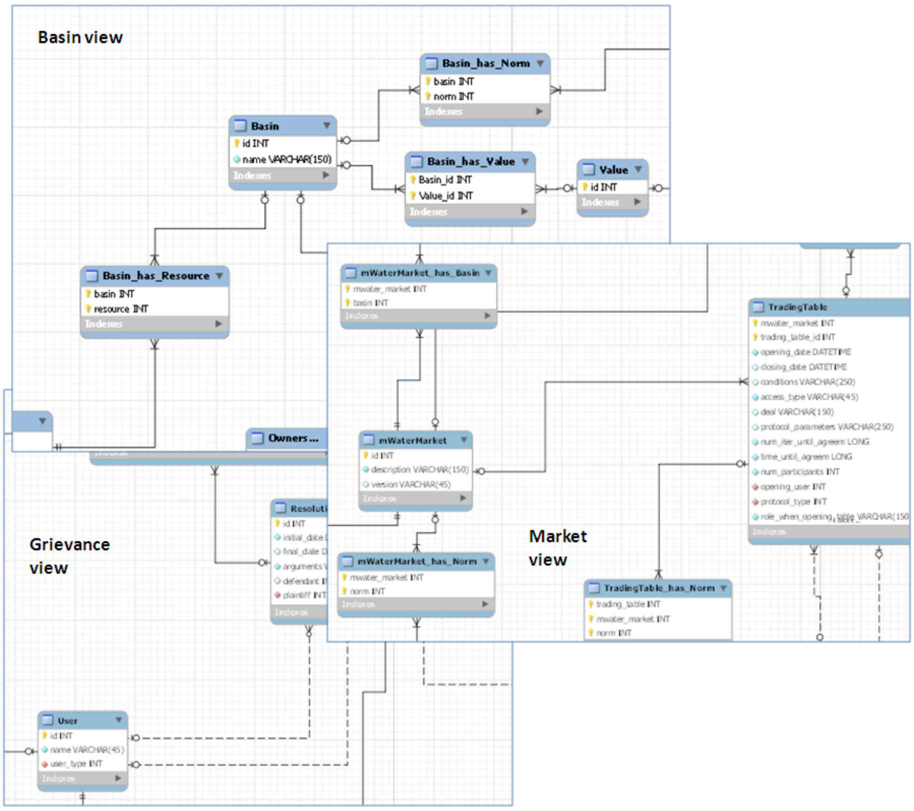


Fig. 3. Fragment of the database: basin, market and grievance views

the roles in the scenes. The generation of the Java classes is done in an automated way, thanks to the tools provided by the EIDE development environment. More particularly, the mapping that is used to generate the agents implementation is shown in Fig. 4. In particular, one Java class is created per valid role (guest, water user, buyer, seller, third party, market facilitator and basin authority) and per scene in which each role can participate. Intuitively, this can be seen as a basic template for an agent participating in a given scene. It is important to note that not all roles participate in all the scenes —recall the definition of the *mWater* EI in Fig. 2—, so there are roles that are translated into more classes than others. The main idea with this is to offer open and flexible templates to implement different agents and norms, which provides more opportunities to the user to evaluate the market indicators under different regulations and types of agents.

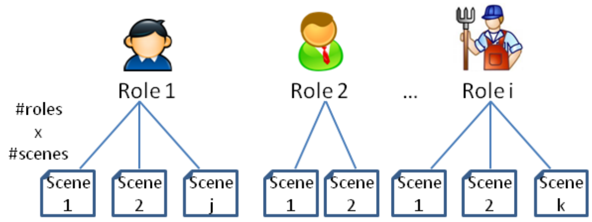


Fig. 4. Schema of the agents implementation. The mapping proceeds by generating one Java class (template) per role in each scene it can be involved

Once the templates have been automatically generated, we can extend them by implementing new classes that represent different behaviours, which is interesting from a simulation perspective. Basically, we override methods to change the original behaviour that allows the agent to move from one state to another, i.e. to execute a transition, or send a message (interact) to other agents. For instance, in the case of the buyer/seller we have implemented a *favourable* and *unfavourable* behaviour. In the former, the agent is always in favour of achieving an agreement to trade and follow the norms of the market, whereas the latter is always against it and does not follow the rules. Additionally, we have placed some decision points that rely on random distributions (inputs of the GUI, see section 5.4) to make the simulation more realistic.

Our implementation introduces an explicit intelligent management into the market in the form of market facilitator. This role has demonstrated very helpful to improve and facilitate the internal behavior of the institution. The market facilitator must be aware of the organizational conventions, the rules of the market and the negotiation structure. But more importantly, (s)he offers intelligent capabilities to help the users under three basic scenarios: i) to decide about opening a new trading table, ii) to decide what user is going to be invited to join that table and why (preliminary process of invitation), and iii) to help within the negotiation (trading) process. First, the facilitator must be aware of the current context of application that may forbid or allow the opening of the most adequate trading table based on the current legislation. Similarly, the market facilitator may offer advice during the grievance procedure, thus making it more efficient. Second, the market facilitator sends invitations to join the table by using data mining rankings that assign a priority to each user for being invited to each table —this involves an intelligent deliberative process based on the user's reputation and trust in previous transactions. Third, the facilitator must obey the particular rules of the protocol to be used within the negotiation, which are usually domain-dependent —different protocols require the application of different sequences of steps—, to make the protocol more agile or to converge more rapidly.

Note that we have also two alternatives for norm enforcement [6]. The former is to implement this reasoning process in the institution side, making it impossible for an agent to violate the norms. Although this provides a trustful and safe environment, it is less flexible and forces the implementation of the agents to be more aware of the legislation of the institution. Moreover, in real life problems, it may be difficult or even impossible to check norm compliance, specially when the violation of the norm cannot be directly observable. And perhaps, it might be preferable to allow agents to violate norms, since they may intend to improve the organization functionality, despite violating or ignoring norms. On the contrary, the second alternative moves the norm reasoning process to the agent side, thus making the system more open and dynamic. In this case, the intelligence of the agent can make it more or less law-abiding in order to obtain a higher personal benefit. If a norm is violated and a third party is affected, the grievance mechanism activates and the conflict resolution stage modelled in the EI is launched.

All in all, and as shown in Fig. 1, this tier includes several techniques to deal with agreement technologies, selection procedures based on data mining processes, intelligent agents that can reason on norms, and planning+CSP methods for navigating through the *mWater* EI, while also trying to find optimal solutions in terms of the amount of water transferred and/or the social satisfaction of the participants.

5.4 Presentation Tier: GUI Simulation Tool

The interface of *mWater* as a simulation tool is simple and intuitive, as shown in Fig. 5. The idea is to offer a straightforward and effective way in which the user configures and runs simulation with the following data: i) the initial and final date for the period to be simulated; ii) the participants, i.e. water users, that will participate in the market (different groups/type of water users lead to different results; e.g. a group in which water users do not trust other members of the group results in a low number of agreements and a high number of conflicts); iii) the protocols to be used during trading, which represent the regulation to be applied in the current simulation; and iv) several decision points to include some random behaviour when users (seller, buyer, basin authority and market facilitator) need to take some decisions. The tool outputs graphical statistical information that indicates how the market reacts to the input data in terms of the number of transfer agreements signed in the market (historical data including information about real or simulated users), number of conflicts generated, volume of water transferred, amount of money, etc. Apart from these straightforward parameters, the tool also shows different quality indicators based on "social" functions in order to asses values such as the trust and reputation levels of the market, or degree of water user satisfaction, among others. This is important to evaluate the quality of the market from the stakeholder's point of view, and not only from a mechanistic standpoint based just on the number of agreements or water transferred, among other.

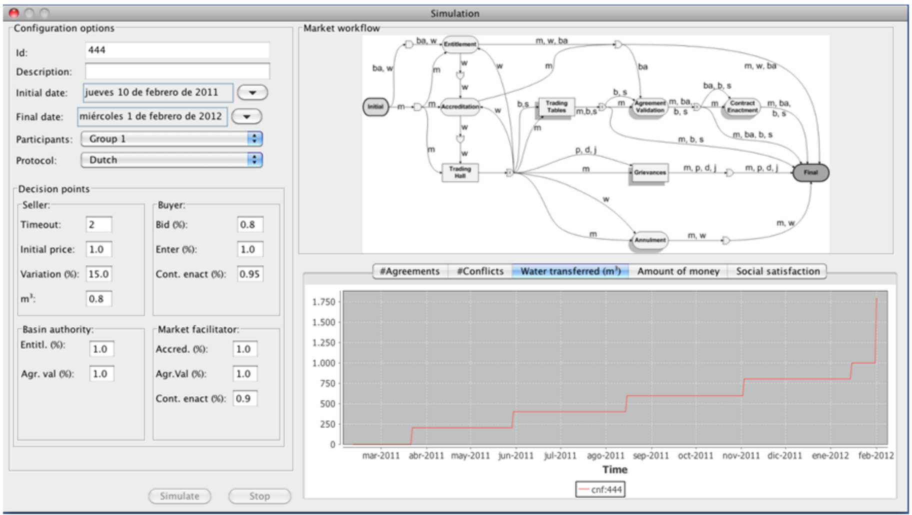


Fig. 5. The *mWater* simulator in action for a given configuration.

5.5 Results: Analysis of the Results

One essential part of a simulation tool to assist in decision taking is to be able to compare the results of different simulations, executed under different configurations. Having this in mind, and aiming at providing as much valuable information as possible, we have also implemented in the GUI a specific decision tier for comparing and analysing simulations. The idea is easy but very effective: the user chooses some simulations from those previously executed and stored in the database, the tool plots them together and extrapolates the best result for each unit of time (day, week, month and so on). For example, if we plot the number of agreements of two simulations, e.g. configurations #337 and #347, and the objective is to maximize this number, a third graphic is added which always shows the highest number of agreements over the timeline (extracted from

#337 and #347), as shown in Fig. 6. This is helpful for the policy makers, as it allows them to find out which *part* of the simulation (and, consequently, which input values for participants, protocols and decision points) leads to the best results in a particular time window, despite the same values are not so good in other windows. In other words, the simulator gives us more precise information on the best result over very particular time units; e.g. the input values for one configuration lead to a higher number of agreements during summer, but the input values for another configuration are better for winter, though none of the configurations in itself is clearly better than the other for a whole year. In particular, in Fig. 6 we can see that configurations #337 and #347 are very similar until May 2011, but afterwards configuration #347 is better —it represents the optimal solution of both configurations. Although the reader may think that this simply puts some sugar on the result simulation form and the user could do this by him/herself, it is important to note that policy makers run dozens (and even hundreds) of simulations for periods that may range from one month to many years. So, doing this analysis by hand and independently for each simulation becomes prohibitive in most scenarios.

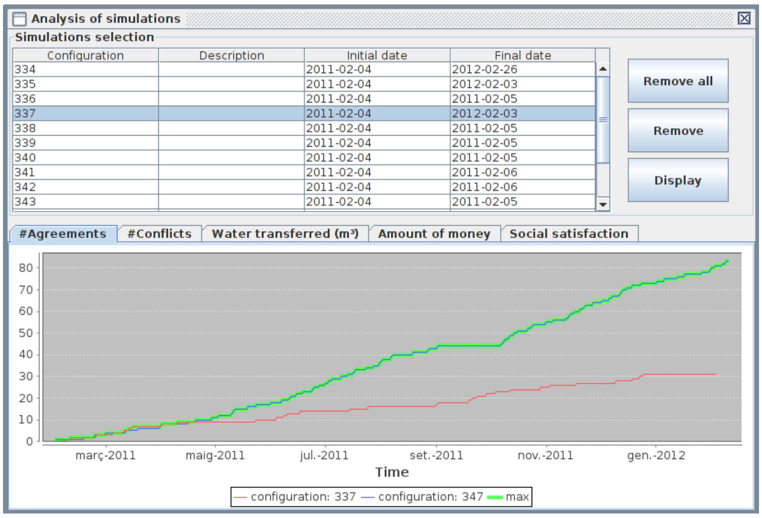


Fig. 6. Analysis of different simulations. Thick line represents the optimal solution, in this case the max number of agreements.

From the experts' point of view and their advice, we can conclude that a model+simulator like this provides nice advantages: i) it successfully incorporates the model for concepts on water regulation, water institutions and individual behavior of water users; ii) it formally represents the multiple interactions between regulations, institutions and individuals; iii) it puts strong emphasis on user participation in decision making; and iv) it finally provides a promising tool to evaluate changes in current legislation, and at no cost, which will surely help to build a more efficient water market with more dynamic norms. Note, however, that the simulation tool is currently mainly policy-maker-oriented rather than stakeholder-oriented. The reason for this is that we have focused on the possibility of changing the norms within the market and evaluate their outcomes —which is the policy makers' labor—, but not in the participation of stakeholders to change the model of the market itself. But clearly, in a social context of water-right management it is important to include tools for letting stakeholders themselves use the system. In other words, the framework should be also able to incorporate the participation of relevant stakeholders, thus helping validate results, which is part of our future work.

6 Conclusions and Future Work

This paper has contributed with *mWater*, a rather sophisticated regulated open MAS-based simulator to assist in decision taking and policy makers; we simulate and test how regulations and norms modify the users' behaviour and how it affects the quality indicators of the market. The core component of *mWater* is an agent-based virtual market for water rights that intends to grasp the components of an electronic market, where rights are traded with flexibility under different price-fixing mechanisms and norms. In addition to trading, *mWater* also simulates those tasks that follow trading, namely, the negotiation process, agreement on a contract, the (mis)use of rights and the grievances and corrective actions taken therein. These ancillary tasks are particularly prone to conflict albeit regulated through legal and social norms and, therefore, they represent a crucial objective in policy-making and a natural environment for the application of agreement technologies. In summary, this type of MAS has a vital importance for decision support as it provides the foundations for the study of that interplay among agents, rule enforcing and performance indicators.

Our current works addresses the following issues. First, to develop a richer normative regulation in order to allow us to simulate more complex types of norms and to observe what are the effects of a given regulation when different types of water users interact in the market. Second, to elaborate more expressive performance measures to evaluate social issues in the market behaviour in order to asses values such as trust, reputation, and users' satisfaction. We believe that this type of measures will provide the policy makers with extra valuable data for decision making about new regulation. Third, although we now consider *mWater* as a simulation tool for decision-support taking, as a long-term research we are also interested in it as an open environment to human users for conducting social and participatory simulations. This would allow us to: i) let stakeholders use directly the system, ii) apply this approach to a specific basin and particular regulation, and iii) see how this is able to reproduce some real data. In such situations, human subjects will take part in the simulation to see the effects of their interaction with virtual agents, applicable norms and their adaptation. Finally, although we focus on a water-right market, the MAS framework is open to other types of (virtual or real) markets, such as energy (electricity) or stock markets. In this line, it would be interesting to compare whether this agent-based water-right market would differ from electricity trading and the systematic effect on the market outcomes.

Acknowledgments. This paper was partially funded by the Consolider AT project CSD2007-0022 INGENIO 2010 of the Spanish Ministry of Science and Innovation; the MICINN project TIN2008-06701-C03-01; and the Valencian Prometeo project 2008/051.

References

1. Andreu, J., Capilla, J., Sanchis, E.: AQUATOOL, a generalized decision-support system for water-resources planning and operational management. Journal of Hydrology 177(3–4), 269–291 (1996)
2. Arcos, J., Esteva, M., Noriega, P., Rodriguez-Aguilar, J., Sierra, C.: Engineering open environments with electronic institutions. Engineering Applications of Artificial Intelligence (18), 191–204 (2005)
3. Argente, E., Giret, A., Valero, S., Julian, V., Botti, V.: Survey of MAS Methods and Platforms focusing on organizational concepts. In: Frontiers in Artificial Intelligence and Applications. vol. 113, pp. 309–316. IOS Press (2004)
4. Botti, V., Garrido, A., Gimeno, J., Giret, A., Igual, F., Noriega, P.: An electronic institution for simulating water-right markets. In: Proceedings of the III Workshop on Agreement Technologies (WAT@IBERAMIA) (2010)
5. Cai, X., Lasdon, L., Michelsen, A.: Group decision making in water resources planning using multiple objective analysis. Journal of Water Resources Planning and Management 130(1), 4–14 (2004)
6. Criado, N., Argente, E., Garrido, A., Gimeno, J.A., Igual, F., Botti, V., Noriega, P., Giret, A.: Norm enforceability in electronic institutions? In: Proceedings of the MALLOW Workshop on Coordination, Organization, Institutions and Norms in Agent Systems in On-Line Communities (COIN@MALLOW). pp. 49–64 (2010)
7. Esteva, M.: Electronic Institutions: from specification to development. IIIA PhD Monography 19 (2003)
8. Giret, A., Garrido, A., Gimeno, J., Botti, V., Noriega, P.: A MAS decision support tool for water-right markets. In: Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (DEMO@AAMAS 2011). p. In press. (2011)
9. Noriega, P.: Agent-mediated auctions: The fishmarket metaphor. IIIA Phd Monography 8 (1997)
10. North, D.: Institutions, institutional change, and economic performance. Cambridge Univ Pr (1990)
11. Palmer, R., Werick, W., MacEwan, A., Woods, A.: Modeling water resources opportunities, challenges and trade-offs: The use of shared vision modeling for negotiation and conflict resolution. In: Proc. of the Water Resources Planning and Management Conference (1999)
12. Rajasekaram, V., Nandalal, K.: Decision support system for reservoir water management conflict resolution. Journal of Water Resources Planning and Management pp. 410–419 (2005)

13. Rodriguez-Aguilar, J.: On the design and construction of agent-mediated electronic institutions. IIIA Phd Monography 14 (2001)
14. Martin de Santa Olalla, F., Dominguez, A., Ortega, F., Artigao, A., Fabeiro, C.: Bayesian networks in planning a large aquifer in Eastern Mancha, Spain. Environmental Modelling and Software 22, 1089–1100 (2007)
15. Schultz, T., Kaufmann-Kohler, G., Langer, D., Bonnet, V.: Online dispute resolution: The state of the art and the issues (2001), available at: http://ssrn.com/abstract=899079 (accessed January 2011)
16. Sierra, C., Botti, V., Ossowski, S.: Agreement computing. KI - Künstliche Intelligenz In press, DOI: 10.1007/s13218-010-0070-y (2011)
17. Smajgl, A., Heckbert, S., Straton, A.: Simulating impacts of water trading in an institutional perspective. Environmental Modelling and Software 24, 191–201 (2009)
18. Thobani, M.: Formal water markets: Why, when and how to introduce tradable water rights. The World Bank Research Observer 12(2), 161–179 (1997)

No Smoking Here: Compliance differences between legal and social norms

Francien Dechesne[1], Virginia Dignum[2], and Yao-Hua Tan[2]

[1] Philosophy – Department of Technology, Policy and Management, Delft University of Technology, The Netherlands
[2] ICT – Department of Technology, Policy and Management, Delft University of Technology, The Netherlands

Abstract. The values shared within a society influence the (social) behaviour of the agents in that society. This connection goes through implicit and explicit norms. Agents act in situations where different, possibly conflicting, norms are applicable. In the case of a norm conflict, an agent will decide to comply with one or more of the applicable norms, while violating others. Our interest is how the *type* of the norms may play a role in such decision, and take the chosen behaviour of an agent to depend on a personal preference order on the norm types.
We distinguish three different types of norms: legal norms, social norms and private norms. We use the introduction of the law prohibiting smoking in cafes as illustration: we present a simulation of this situation involving agents' preferences over different norm types. The results of this simulation are used for an explorative a model for normative reasoning based on norm types. We discuss a possible connection between the composition of a society in terms of these profiles and its culture.
We briefly discuss the relevance of the model with respect to *value sensitive design* of socio-technological systems. In particular, the model will provide insights in how to balance central regulation on the one hand, and self-regulation on the other, in the construction of institutions.

1 Introduction

Values are ideals that are considered to be worth pursuing. Examples of widely shared values are health, safety, security, freedom, joy, beauty, friendship, justice. People will generally exercise, promote and sometimes command behaviour -by themselves and others- that supports such values.

A norm is a more or less general rule of conduct within a group or society, which constitutes a link between the abstract values (goals) and concrete behaviour that is considered to serve one or more of those goals. In this paper we focus on the effect of norms on behaviour, taking into account the different *types* of norms: implicit norms that emerge among the people, norms that are explicitly imposed on the community (by a governing body) on the other, and norms that agents develop privately over their lives (by being part of different communities and having certain experiences). This last type can be seen as a sort of default behaviour of an agent. We will refer to these three types as social, legal and private norms respectively.

Agents may find themselves in situations where different, possibly conflicting, norms are applicable. In the case of a norm conflict, an agent will decide to comply with one or more of the applicable norms, while violating others. One could analyze such deliberation through the preferences agents may have on the *values* underlying the conflicting norms. This would however require a semantic framework to capture the meanings of particular values (e.g. to state what exactly it means that an agent prefers joy over health). Also, it would require a formal specification of the nature of the connection between norms and their underlying values, which seems hard to capture in general. With the development of a generic formal model for norm compliance in mind, we abstract from particular values and their connection to norms: we take the chosen behaviour of an agent to depend on a personal preference order on the *norm types*.

In particular, we study the difference in conforming to social conventions versus complying with explicitly given laws (with penalties). This is partly motivated from an interest in the design of new governance models for socio-technological systems, which aim to include elements of self-regulation. From the perspective of a governing body for such system, it is interesting to know whether introducing a formal rule prescribing or prohibiting certain behaviour will lead to the best support of the general goals of the system, or whether it could be better to let rules of conduct emerge (after convincing the agents of the relevance of the underlying goals or values). In this sense, the current paper can be seen to be a contribution to the program of *Value Sensitive Design*, which studies theories and practices to include (moral) values into the design of technology and organisational systems.

The paper is organized as follows. Section 2 discusses related work in the areas of multi-agent systems, sociology and value-sensitive design theories. In section 3 we define the norm type classification proposed in this paper. The concrete example of the introduction of anti-smoking laws is discussed in section 4 and a simulation of the resulting behaviour is presented in section 5. In section 6 initial development of a formal model for reasoning and analysis of norm types is introduced, with particular focus on the link to the culture of the society being affected by the norm. The consequences of this work to Value Sensitive Design are discussed in section 7. Finally, conclusions and directions for future work are presented in section 8.

2 Related work

The work in this paper is a first step towards a norm preference model which enables to describe and reason about norm compliance with respect to the type of that norm. This extends current work on multi-agent models for norm compliance, focusing on different norm types and study how this aspect of norms may lead to different overall behaviour of the system. In this section we mention existing work this paper builds on.

The norm compliance in multi-agent systems (MAS) has resulted in different approaches [5, 14]. A regimented view of norms [6] in which norms are viewed as constraints, and a regulated view [1, 9] in which norm enforcement is explicit and agent's motivations play an important role in compliance decisions. When regimenting norms, all agents' actions leading to a violation of those norms are made impossible. I.e. the design of the system makes it impossible to perform of forbidden actions (think of gates at the metro station that prevent entering without ticket). From an engineering perspective, these approaches are straightforward but they seriously limit agent autonomy. Regulated approaches require both the establishment of institutions that monitor and enforce norm compliance, and the existence of normative agents that are able to reason about the effect of norms on goal achievement.

In most cases, work on norm compliance assumes norms to be implemented and enforced by an institution. That is, only legal norms in the sense we define above are considered. In sociology, norms are the behavioural expectations within a society or group. As such, norms are the rules that a group uses to determine appropriate and inappropriate values, beliefs, attitudes and behaviours. These rules may be explicit or implicit. Failure to follow the rules can result in several punishments, including exclusion from the group. Norms are the attitudes and behaviours of a group, and the values are what is important to that group.

[10] gives a conceptual analysis of values and represents them formally through the basic notion of *preference*. Norms are action guiding principles that are connected to (shared) preferences. Norms give a sense of shared values, but values can also create their own norms. [3] defines social norms as non-formal, non-sanctioned (in the explicit sense), and relative to (perceived) number of people adhering to the norm, and expectations of others w.r.t. adhering to the norm. More formal and applicable is the work by Castelfranchi, Conte and Dignum on norm compliance [4]. However, the distinction between legal, social and private norms is not explicit in their work. Their work distinguishes between acceptance of norms and complying with them which is an important aspect for future work.

Which enforcement mechanisms are effective and how sanctions are likely to be followed is directly related to the values of an society. Moral values are the standards of good and evil that guide an individual's behaviour and choices [19]. Individuals, groups, and societies develop own value systems used for the purpose of ethical integrity. Value Sensitive Design (VSD) is a methodological design approach that aims at making moral values part of technological design, research, and development [20]. In particular, we are interested in policy design, that is methods to determine which of various alternative policies will most achieve a given set of goals in light of the relations between the policies and the goals. However, value descriptions do not provide enough formality to be usable at the system specification level. Therefore, an important aim of VSD is to provide a formal representation of values, that 'translates' abstract concepts into a formal representation, to enable a system that supports specification and analysis of policies. The analysis presented in this paper is aimed to support the evaluation of policies and their effect on the behaviour within the community.

Research by Hofstede has shown that national cultures differ in particular at the level of, usually unconscious, values held by a majority of the population [11]. Values, in this case, are "broad preferences for one state of affairs over others". The Hofstede dimensions of national cultures are rooted in our unconscious values. Because values are acquired in childhood, national cultures are remarkably stable over time; national values change is a matter of generations. In this paper, we make a first attempt to link norm type preferences to society culture, which will enable policy makers to decide on the best norm type to use as implementation of a policy. For instance, in cultures that prefer private norms over other norm types, a campaign illustrating the negative effects of smoking will be more effective than a formal law prohibiting smoking. The latter will in turn be more effective in cultures preferring legal norms.

3 Norm Types

As indicated above, we will distinguish three types of norms in this paper, which we call legal, social and private norms. In this section we characterize each of them.

We use the term *legal* norms for rules of conduct that are explicitly formulated and imposed on the community by a central entity. The laws of a community are typical examples of legal norms. Legal norms make explicit for the entire community how to behave in order to support some underlying value.

Acceptance of a legal norm may depend on the extent to which the underlying value is supported, and the prescribed behaviour is considered to *count as* support for that value by the agents in the community. In practice, legal norms usually also have an explicit sanction for violation, which may or may not be enforced. In this paper, for the sake of simplicity, we will keep these considerations implicit.

Social norms are more implicit and more flexible: they only cover a subgroup of the community, their boundaries are hardly defined, and an agent can decide to (temporarily) leave a certain subgroup on the basis of lack of support for the social norms in that group. Also, while the change of a legal norm is a momentaneous central event, social norm change is rather a dynamic process of diffusion.

Following [3], an agent will comply with a social norm or rule if:

- a sufficiently large number of others conforms to the rule and;
- a sufficiently large number of others expects her to conform to the rule, and may sanction behaviour.

In this paper, we take 'sufficiently large' to be the numeric majority ($> 50\%$) of the population, but this would be an interesting parameter to differentiate between more types of agents. Also, the aspect of of expectations of others is not yet part of our model, but we consider it to be an important factor in practice in the deliberation whether or not to comply with a certain social norm. Social norms are more dynamic and context dependent, while legal norms draw clear lines in what counts as desired behaviour.

With *private* norms, we indicate the personal normative beliefs a person has developed over his or her life. We abstract from the way they came to be the personal norms of an agent (partly derived from social norms, partly from legal norms, in the different societies an agent has been part of), and assume them to be fixed for each agent by their personal history. They are the standards of behaviour a person holds for him- or herself. We take these norms to be invoked if no other norms are applicable, or if the agent in question is insensitive to other norms.

The scope of a norm depends on its type. We will take the legal norms to hold for the entire population of the agent society. The scope of a social norm is some, possibly dynamic, subgroup of agents in a certain situation, for example the agents currently present in a given public area. We take the scope of a private norm to be only the agent who holds that norm for him- or herself.

Different considerations play a role in the agent's decision to behave according to the norm or not, depending on its type. The first row lists the primary value connected to following norms of such type, the second row how and by whom they come to exist, and the third row indicates the monitoring and sanctions.

Legal norm:	Social norm:	Private norm:
Compliance	Conformity	Integrity/ Being consequent/
imposed by institution explicit	emergent / dynamic implicit	fixed (by history) implicit
enforced sanctioned	power of numbers/ exclusion	conscience lower self-esteem

We characterize agents by their primary preference which norms he considers guiding for his behaviour:

1. *legal* agents: law-abiding, whatever the law prescribes, they do.
2. *social* agents: whatever most of the agents in a certain shared context prefer, they do as well
3. *private* agents: irrespective of law or context, they do what they themselves judge to be right.

4 Motivational Case

We take the example of the introduction of anti-smoking regulation as illustrative example, because it involves a transition from social norms about smoking in cafes (which were probably different for different cafes, depending on their clientele) to one uniform law prohibiting smoking in all cafes. Formal smoking prohibitions for restaurants and cafes have been introduced in several European countries over the past years, with Ireland being one of the first (2004), and The

Netherlands relatively late (2008). While most people in current society support the underlying value of the introduced law, viz. that smoking is unhealthy for the smoker and its environment, the introduction of the prohibiting law provoked considerable resistance and –at least in some countries– vast violation.

In this example, different values are at stake. On the level of behaviour, such values are health, care for others (with respect to their health), but also: joy or pleasure. Also the economic interest of the bar keepers may be at stake (bar keepers claim their clientele dramatically dropped after enforcing the new law). But particularly connected to the cafe setting are certain values regarding the authority that is accepted, like freedom or autonomy. The relative weight of these values in the cafe setting, may explain the lesser degree of acceptance of the law compared to the acceptance in other seemingly similar settings, like cinemas or restaurants.

In the next section, we provide a simulation for the transition from the (informal) social norm ("you should not smoke in public areas, especially not if it bothers others") to the (formal) *legal* norm ("it is forbidden to smoke in cafes"), and we review how this may affect the size of the total clientele, and the size of the clientele that accepts smoking in the bar (either by themselves or by others). This inspires us to reflect on how norm type preferences explain the witnessed behaviour, and how this links to reality.

5 Simulation

We developed a simple simulation to illustrate how different preferences over the three norm types may result in different behaviour changes after the introduction of the anti-smoking laws. Agents in this scenario have a private attitude towards smoking and a preference order on the three types of norms (legal, social and private) discussed in the previous section. For the sake of this simulation, we simplified this into each agent having one preferred norm type (i.e. the top element in his preference order on the norm types).

Agents meet each other in a cafe. We model the effect of the introduction of the smoke prohibition in cafes, both on the average number of cafe clients (depending on their preferences, agents may leave the cafe if they can no longer smoke there, or stay longer if it actually becomes smoke-free), and on the number of those clients that violate the law (agents may still go to the cafe and ignore the law). The legal norms range over the entire society, the social norms are relative to the contingent context of those people present in the cafe. This will give the simulation its particular dynamics.

5.1 Simulation design and implementation

We have developed a simple environment "CafeWilhelmina" that simulates a community of people that frequent a cafe. Each agent is either in the cafe or not, holds a personal preference with respect to smoking, and has a preferred norm type. The social norm whether smoking is accepted, is determined by the majority of the agents (rules **R4-R5**). At any time, all agents can decide either to go to the cafe or to leave the cafe (rules **R1-R2**). Based on demographic information, we assume that smokers are more likely to frequent the cafe than non-smokers (as reflected in parameters PGS and PGN). Agents that are in the cafe can decide to leave either because they don't like the environment (i.e. the prevailing social norm is different from their preference, rule **R3**) or simply because they've stayed long enough (parameter PL).

Formally, a CafeWilhelmina setting is defined by the following parameters:

- $A = \{a : a = ((\textit{private-pref}, \textit{norm-pref}, \textit{in-cafe}), (\textit{go-cafe}, \textit{leave-cafe}))\}$, are the agents. They are characterised by their beliefs and actions, stored in parameters *private-pref* (a personal belief whether smoking in public is OK or not), *norm-pref* (which norm type they would primarily follow), *in-cafe* (whether or not they are in the cafe), and *go-cafe* (whether they decide to go to the cafe if they're not yet there).
- $SN \in \{\textit{smoke}, \textit{not-smoke}\}$ is the social norm prevailing in the cafe at any moment. Here '*smoke*' abbreviates the norm that it is permitted to smoke, not merely the action 'to smoke'. In particular, you don't need to be a smoker to comply with '*smoke*'.
- $SN_0 = \textit{smoke}$, is the initial social norm
- $N = 100$, is the number of agents
- $T = 200$, is the number of ticks per run
- $LawT = 100$, is the tick of the introduction of the anti-smoking law
- $DN \in \{\textit{smoke}, \textit{not-smoke}\}$, is the prevailing legal norm
- $PGN = 0.15$, is the probability for non-smoking agents to go to the cafe
- $PGS = 0.45$, is the probability for smoking agents to go to the cafe
- $PL = 0.25$, is the probability for all agents to leave the cafe
- R are the rules, defined as follows:

R1 $\forall a \in A, do(\textit{go-cafe}(a)) \rightarrow \textit{in-cafe}(a)$
R2 $\forall a \in A, do(\textit{leave-cafe}(a)) \rightarrow \neg \textit{in-cafe(a)}$
R3 $\forall a \in A, \textit{in-cafe}(a) \wedge (SN \neq \textit{private-pref}(a)) \rightarrow do(\textit{leave-cafe}(a))$
R4 $(\textit{smokers-in-cafe} > \textit{others-in-cafe}) \rightarrow (SN = \textit{smoke})$, where *smokers-in-cafe* and *others-in-cafe* indicates the number of clients with a preference for smoking / not smoking, respectively.
R5 $(\textit{smokers-in-cafe} =< \textit{others-in-cafe}) \rightarrow (SN = \textit{not-smoke})$
R6 $\forall t >= LaWT : DN = \textit{not-smoke}$

Private preferences and norm preferences are assigned randomly to each agent, according to a given probability distribution (see the percentages in the bottom line of figure reffigure:simulationresults). During each run of the CafeWilhelmina simulation agents will go to or leave the cafe. Each run consists of a number of 'ticks'. Each agent can use each tick either to act or to reason (not both simultaneously). The objective is to study the behaviour of the community with respect to the introduction of anti-smoking laws. We have implemented CafeWilhelmina using the RePast simulation environment [18].

5.2 Results

Figure 1 shows the results of the simulation for different population compositions. In this scenario, agents have a fixed private preference towards smoking (assigned randomly with 50% chance) and a fixed norm type preference (i.e. they will either follow legal, social or private norms). We varied the percentages of agents with a certain norm preference, as indicated in the x-axis. We run each configuration 10 times, the values in the graph are the average numbers over these runs. Each pair of columns shows the situation before and after the introduction of the law.

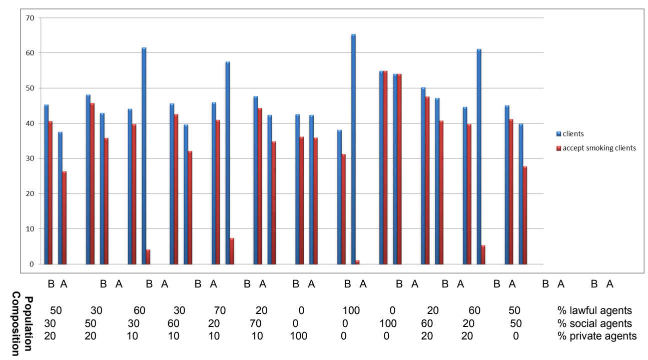


Fig. 1. Results of the simulation for different compositions of the population

As can be expected, highly normative societies (where the percentage of lawful agents is above 50%) react positively to the introduction of the smoking ban[3]. This can be explained by the fact that non-smokers will be more inclined to go to the cafe, as they can be sure that the place will be smoke free. In configurations where social agents are in the majority, the number of clients typically diminishes after the introduction of the law. Non-smokers and lawful agents will not stay in the cafe: they don't feel comfortable, because of the smoke or because the law is not being upheld respectively.

6 Model: Norm type orders

In this section, we make a first attempt to reconstruct the role of different norm types in the agent's deliberation to comply or violate a certain norm.

6.1 Modelling choices

By restricting the model to the effect of norm types on agent normative reasoning (in opposition to reasoning based on the object of the norm), we make an explicit decision on the scope of the model, as follows:

[3] The residual smokers number has to do with the transition moment in which there were still smokers in the cafe.

– We do not explicitly model sanctions, even though the existence of formal sanctions to norm violation constitutes an essential difference between legal and social norms.
– We do not include a mechanism that explains how social norms emerge.
– We do not address what is "a sufficiently large number" in the definition of social norms by [3] (we took it to be a numeric majority in the simulation).
– We do not include the realistic possibility of one agent convincing another to comply to a norm (or not)

Further assumptions:

– All agents (within the society) share all legal norms.
– We take social norms to be context and subgroup-dependent.
– Individual agents each have their own private values, which are fixed (so we do not consider how they grew from personal history).

6.2 Preferences over norm types

The concepts around which we build our model are: values, norms, types of norms, preferences and actions.

Values are abstract goals for the agents in the society. **Norms** guide the **actions** of the agents, to support one or more values (or, one could say, norms *reflect* values in that norms prescribe conduct that promotes them). Norms are action oriented, values are not. Their relations are depicted in Figure 2. We as-

$$\begin{array}{c}\boxed{\text{values}}\\ \Downarrow\\ \boxed{\text{norms}}\\ \downarrow\\ \boxed{\text{actions}}\end{array}$$

Fig. 2. Layering of the concepts

sume all agents to have **preferences**. We follow [10] in taking these to represent values an agent maintains.[4] With norms functioning as links between values and actions, preferences reflecting values can explain why –in particular in case of norm conflict– a certain action is chosen by an agent rather than another. In our model, we take the norm types to represent agents' values concerning following rules of conduct: compliance, conformity, consistency (cf. the table in section 3).

Our model builds upon the assumption that each agent maintains a complete (strict) order on the norm types, i.e. one of the following:

– Legal $\succ$ Social $\succ$ Private

[4] The nature of this representation relation is however not specified in [10].

- Legal $\succ$ Private $\succ$ Social
- Social $\succ$ Legal $\succ$ Private
- Social $\succ$ Private $\succ$ Legal
- Private $\succ$ Legal $\succ$ Social
- Private $\succ$ Social $\succ$ Legal

We take this individual order of the norm types to represent the agent's values with respect to norms. In the next section, we explore this connection.

6.3 Connection to Character and Culture

These six orders of the norm types can be taken to define a part of the agent's "personality". Preferences over norm types can be taken to be a reflection of how agents see themselves as part of the society: whether they accept the authority of the central government (if they have a high preference for legal norms), whether they believe in the self-regulating power of groups (if they have a high preference for social norms), or whether they give most authority to their private judgement.

Note that in the simulation we only used the most preferred norm type for each agent, for simplicity. Taking the complete order over the three norm types, contributes to a richer representation of different attitudes towards rules of conduct. For example, an agent who least prefers private norms, can be described as one who believes norms should be commonly shared (as is the case with legal and social norms), and one who least prefers social norms, could be seen as one who likes norms that are clear and unambiguous to him. An agent who least prefers legal norms, can be seen as less authority-sensitive.

Here we give some tentative characterisations of the six agent types corresponding to the six norm type orders. The structure of the orders gives us some oppositions:

- Legal $\succ$ Social $\succ$ Private: *authoritarian*
- Legal $\succ$ Private $\succ$ Social: *absolutist*
- Social $\succ$ Legal $\succ$ Private: *collectivist*
- Social $\succ$ Private $\succ$ Legal: *relativist* (opposite of absolutist)
- Private $\succ$ Legal $\succ$ Social: *individualist* (opposite of collectivist)
- Private $\succ$ Social $\succ$ Legal: *anarchist* (opposite of authoritarian)

This characterisation of the norm type orders gives us three character dimensions that are not necessarily orthogonal: absolute–relative, authoritarian–anarchist, collectivist–individualist.[5]

Each society is composed of agents with different norm type preferences. The ratio in which each of the agent types is present in a society, reflects its culture with respect to rules of conduct. For example, the highly individualist

[5] Note that we restrict here to the mere order of the norm types, without 'weights' assigned to the norm types. This means that "Legal $\succ$ Social $\succ$ Private" represents both agents who exclusively consider legal norms, and agents who put the three norm types on the same level. Adding weights could refine this.

non-hierarchical character of a society is reflected by it having a large portion of agents of the last type (Private$\succ$ Social $\succ$ Legal). The model in terms of norm types can in that way be used to represent different cultures in their response to the introduction of new (types of) regulation.

A very well-known characterisation of cultures is the one of Hofstede [11]. It distinguishes 4 cultural dimensions: Power Distance Index (PDI), Individualism (IDV), Masculinity Index (MAS), Uncertainty Avoidance Index (UAI).[6] A characterisation of different countries in these dimensions can be found at [12].

A link between cultural dimensions to our norm type orders, would provide a translation from the (known) Hofstede cultural characterisation of societies with their norm type preference profile. This deserves more thorough investigation. For the moment, we postulate some possible correlations between a predominant preference for a norm type and cultural dimensions:

- legal norms: are associated with high PDI (legal norms come from an authority) and/or high UAI (through their explicit formulation, legal norms create clarity and intersubjectivity);
- Social norms: are associated with low PDI (equality), low MAS (caring for others), and low IDV (the importance of belonging to the group)
- Private norms: high IDV (the private context is taken as guiding), high MAS (assertiveness), low PDI.

A further development of such correlations could result in a model predicting for a given society whether the introduction of a new policy will lead to the desired behaviour change or not, or whether it would be better to let it be an emerging process. This would enable policy designers to understand the effect of norm types on different cultures and decide on the most appropriate type to use as medium for the introduction of a policy. In that case, a theory for arguments promoting values and their norms could be of use [2].

6.4 Reflecting on the simulation

In our simulation, we see how the effect of the introduction of the anti-smoking law depended on the composition of the society in terms of norm type preferences (cf. Figure 1). Let us use triples (x, y, z) to represent such composition in terms of percentages lawful (x), social (y) and private (z) agents.

Note that, despite the fact that we only used the most preferred norm type in our simulation, the aggregate level for the society (of 100 agents) does result in a complete (possibly non-strict) order over the norm types (with weights, determined by the composition of the society in terms of the top preferences of the agents).[7]

[6] For simplicity, we ignore for now the dimensions Long-Term Orientation (LTO), and Monumentalism which were added later.

[7] To derive one aggregate preference order on the norm types for a society based on the individual complete norm type orders, would be much more subtle. This is what the Discursive Dilemma [17] illustrates.

Let us first consider the three extreme cases. In a fully law-abiding society (100,0,0), the effect of the introduction of the smoking prohibition is dramatic: after the introduction, everyone complies and cafe attendance increases (because of the consensus created by the law among those lawful agents). In the absence of lawful agents, the introduction of the law will have no effect: see (0,100,0) and (0,0,100). It may be argued that in realistic societies with laws at least some lawful agents exist. If not, the laws of that society would in fact be useless.

It is interesting to note, that with a majority of lawful agents but also a portion of social agents, the cafe attendance will actually increase after the introduction of the smoking prohibition (see (60,30,10), (70,20,10), (60,20,20)), with a vast majority of the attendants complying with the law (an exception being (50,30,20)). The prohibition is noticeably less effective with a majority of social agents and less lawful agents (columns (30,50,20), (30,60,10), (20,70,10)), with less attendants and only a slightly lower percentage of smokers.

We end with an attempt to link our simulation results with the reality of the smoking prohibitions in different countries. Let us look at the results of a Dutch study evaluating the effect of the smoking prohibition [13], despite some important differences with our simulation in what it counts. It compares the effect of the introduction of the smoking prohibition laws for cafes in Ireland (2004) and the Netherlands (2008) over the first months.

It turns out that the number of cafes that became smoke-free has been significantly higher in Ireland (2004) than it has been in the Netherlands (2008). Smokers were asked whether people were smoking in the last cafe they recently attended.[8] In the Netherlands, this was the case in 96% of the cafes before the ban, and 31% after, while these figures were 98% and 5% respectively in Ireland. This suggests that the Irish have a high proportion of lawful agents, with also a significant number of social agents, who follow the changing majority.

Unfortunately, looking at the cultural differences between the Netherlands (NL) and Ireland (IRL) does not give a clear picture as to explaining this difference, and giving a connection between cultures and the effect of the introduction of a new law. The cultural difference is most significant in the masculitinity dimension (NL:14 versus IRL:62). The differences in the other dimensions are less extreme: the Dutch are more uncertainty avoiding that the Irish (UAI: NL:53, IRL:30), both are individualistic, but the Dutch more extremely so (PDI: NL:80, IRL:65), and both are egalitarian, but the Dutch less so (PDI: NL:38, IRL:22). With the proposed link we made between the cultural dimensions and the norm type preferences, one would expect the Dutch to respond better to the introduction of the new law.

However, there is another important reason why it is not so easy to directly link these figures to our simulation or our model. The Irish law differs from the Dutch one, in that it prescribes a complete ban of smoking, while the Dutch law

[8] Recall that in our simulation, we did not count *actual* smoking, but whether agents supported the permission to smoke in cafes

allows cafes to install separate, unserviced, smoking areas. How this is reflected in the figures above is not so clear.

7 Application to Value Sensitive Design

Value Sensitive Design (VSD) [21, 8] is an approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process. The assumption here is that technology is not neutral with respect to values [15]. VSD recognises that the design of technologies bears "directly and systematically on the realisation, or suppression, of particular configurations of social, ethical, and political values" [7]. This is obviously true about the design and implementation of social policies.

According to VSD the process of implementing a (institutional and/or technologic) system should be guided by social values which not only must be made explicit but also must be systematically linked to design choices. Several authors have proposed a derivative process that resembles the relations depicted in Figure 2 [1, 16]. The process requires values to be identified and (formally) described, then translated into concrete norms describing what should (or not) be done, and annotated with concrete, contextualised information that is needed to enforce the norm. That is, besides describing the object of the norm (e.g. smoking in pubs) an explicit description of enforcement and exceptions must be specified (e.g. enforcement will be done by the Health Authorities, and pubs can decide to introduce a smoking space separated from their main area). Lastly, the operational norm is translated into specific system designs expressing the actions to be monitored (e.g. client smoking in main pub area) and the actions that can be taken to enforce the norm (e.g. fine the pub owner, close the pub). Making these choices explicit will enable to trace the design decision back to the values that supported it.

Our work contributes to Value Sensitive Design as it enables to link design choices to value and norm preferences. Given the considerations in section 6.3, the acceptance or not of a certain policy is influenced by the cultural background of the groups affected by that policy. The analyses the norm preference model of that group guides the choices on policy implementation. E.g. a society where social norms are preferred will more likely react positively to a policy that is introduced by word of mouth in social networks, whereas a society that prefers legal norms will react better to an implementation of the policy by legislation means.

8 Conclusions and future work

In this paper, we have explored how distinguishing different types of norms can provide a way to reconstruct agents' acting according to, or in violation of, norms. An advantage of looking at norm types rather than norms, is that we can study the effect of norms in a uniform way, i.e. without having to give formal accounts of the particular values that support given norms, and the precise connection between those values and the norms they support.

We distinguished three types of norms: legal, social and private norms, and proposed the six possible orderings of those types to characterise agents. We can then think of the culture of a society (with respect to agents' attitudes towards norms) in terms of how it is composed of such agent types. We provided a simulation for the introduction of a law, prohibiting smoking in cafes, in an agent society. It showed that different compositions of the society in terms of agents' norm type preferences will respond differently to the introduction of the law.

We see this research as a contribution to the research programme of Value Sensitive Design, as it aims to be a way of making the connections between values and design (e.g. of an institutional system) more explicit, more formal, and more manageable. Taking into account the preference profile of a community with respect to norm types, and thereby aligning with the values of that community, should help to design more effective policies.

The ideas presented in this paper show many directions for further research. We mention a few. First, many interesting refinements can be made with respect to the framing of the norm types. Aspects that will be enrich the analysis are: including sanctions, varying the threshold for 'sufficiently large number of others' and taking into account expectations of other agents (for social norms), possibly making private norms more dynamic. A more refined model would result in more refined agent profiles.

Interesting dynamics would be added by the inclusion of an argumentation mechanism, modelling how agents (of a certain type) could try to convince others to change norm type preferences.

We should further explore the connection between Hofstede's cultural characterisation of societies, and the composition of societies in terms of our norm type characters. We have tentatively indicated a way of expressing our norm types in terms of the Hofstede dimensions, but further investigation should be able give better founded relations between the two frameworks. Especially when more and more refined statistics on the effect of the anti-smoking laws in different European countries become available, this will give material for validation.

References

1. H. Aldewereld. *Autonomy vs. Conformity - an Institutional Perspective on Norms and Protocols.* PhD thesis, Univ. of Utrecht, 2007.
2. K. Atkinson and T. Banch-Capon. Co-ordination and co-operation in agent systems: Social laws and argumentation. *Argumentation in Multi-Agent Systems*, pages 122–140, 2009.
3. C. Bicchieri. *The Grammar of Society.* Cambridge University Press, 2006.
4. R. Conte, C. Castelfranchi, and F. Dignum. Autonomous norm acceptance. In J. P. Müller, M. P. Singh, and A. S. Rao, editors, *ATAL*, volume 1555 of *LNCS*, pages 99–112. Springer, 1998.

5. F. Dignum. Autonomous agents with norms. *Artificial Intelligence and Law*, 7:69–79, 1999.
6. M. Esteva, J.-A. Rodrguez-Aguilar, C. Sierra, P. Garcia, and J. Arcos. On the formal specification of electronic institutions. In F. Dignum and C. Sierra, editors, *Agent Mediated Electronic Commerce*, volume 1991 of *LNCS*, pages 126–147. Springer, 2001.
7. M. Flanagan, D. Howe, and H. Nissenbaum. Values in design: Theory and practice. In M. van den Hoven and J. Weckert, editors, *Information Technology and Moral Philosophy*. Cambridge University Press, 2008.
8. B. Friedman. Value sensitive design. In *Encyclopedia of Human-Computer Interaction*, pages 769–774. Berkshire Publishing Group, 2008.
9. D. Grossi. *Designing Invisible Handcuffs - Formal Investigations in Institutions and Organizations for Multi-Agent Systems*. PhD thesis, Univ. of Utrecht, 2007.
10. S. O. Hansson. *The Structure of Values and Norms*. Cambridge University Press, 2001.
11. G. Hofstede. *Culture's Consequences, Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Sage Publications, 2001.
12. G. Hofstede and G. Hofstede. Cultural dimensions, 2003. Website: `http://www.geert-hofstede.com/`.
13. International Tobacco Control Nederland. ITC policy evaluation project – de effecten van de rookvrije horeca op rookgedrag. eerste nameting, 2009. Available through `www.stivoro.nl` (in Dutch).
14. F. López y López, M. Luck, and M. d'Inverno. Constraining autonomy through norms. In *AAMAS '02*, pages 674–681. ACM, 2002.
15. N. Manders-Huits and J. van den Hoven. The need for a value-sensitive design of communication infrastructures. In P. Sollie and M. Dwell, editors, *Evaluating New Technologies*, volume 3 of *The International Library of Ethics, Law and Technology*, pages 51–60. Springer, 2009.
16. S. Overbeek, M. Dignum, and Y.-H. Tan. Towards a value-sensitive system to support agents in norm fulfillment and enforcement, 2010. Presented at IAT4EB.
17. P. Pettit. Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11(1):268–299, 2001.
18. Repast organization for architecture and development, 2003. Available at `http://repast.sourceforge.net`.
19. E. Scott. Organizational moral values. *Business Ethics Quarterly*, 12(1):33–55, 2002.
20. M. van den Hoven. Design for values and values for design. *Information Age +, Journal of the Australian Computer Society*, 7(2):4–7, 2005.
21. M. van den Hoven. ICT and Value Sensitive Design. In P. Goujon et al., editor, *The Information Society*, volume 233 of *IFIP*, pages 67–72. Springer, 2007.

Towards Qualitative Reasoning for Policy Decision Support in Demonstrations

Natalie Fridman, Gal A. Kaminka and Avishay Zilka*

The MAVERICK Group
Computer Science Department
Bar Ilan University, Israel

Abstract. In this paper we describe a method for modeling social behavior of large groups, and apply it to the problem of predicting potential violence during demonstrations. We use qualitative reasoning techniques which to our knowledge have never been applied to modeling crowd behaviors, nor in particular to demonstrations. Such modeling may not only contribute to the police decision making process, but can also provide a great opportunity to test existing theories in social science. We incrementally present and compare three qualitative models, based on social science theories. The results show that while two of these models fail to predict the outcomes of real-world events reported and analyzed in the literature, one model provide a good results. Moreover, in this paper we examine whether machine learning techniques such as decision trees may provide better predictions than QR models. While the results show that the machine learning techniques provide accurate predictions, a slightly better prediction than our QR model, we claim that QR approach is sensitive to changes in contrast to decision tree, and can account for *what if* scenarios. Thus, using QR approach is better for reasoning regarding the potential violence level to improve the police decision making process.

Keywords: Demonstrations, Social Simulation, Qualitative reasoning

1 Introduction

A violent demonstration, resulting in casualties among its participants, police forces and innocent bystanders, is unfortunately not a rare phenomena. This paper deals with improving the police decision making process, by providing useful predictions as to the potential outcomes of demonstrations, given the specific settings and the ability to account for *what if* scenarios. The hope is to decrease the number of casualties by preventing violence.

In general, there are several model-based technologies that can be used to generate predictions. Agent based simulations [7] require detailed individual cognitive modeling, and furthermore, modeling at the individual participant level is too fine a resolution for useful predictions. Numerical simulation [13] models at an appropriate resolution

* We thank IMOD and ISF Grant 1357/07 for partial support of this research. Thanks to K. Ushi.

(global group behavior), but unfortunately requires complete and precise domain information, which is not available here. There exists significant literature on the factors that impact violence during demonstrations, but it mostly reports on partial, macro-level qualitative descriptions of the influencing factors. However, there are also technologies that can be used to generate predictions which are not requires building a model such as machine learning techniques, for example decision tree. Decision tree takes as an input set of properties and build a model, a set of rules, that allows most accurate classification to the given data.

In this paper we describe a novel application of Qualitative Reasoning (QR) [10, 3] to modeling and reasoning about potential violence level during demonstrations. QR is a sub-area of AI, which enables reasoning with partial or imprecise numeric information. Using QR, it is possible to draw useful conclusions even with only qualitative representation of data and order values (such as little/medium/large). Thus such modeling provides an opportunity to test existing social science theories regarding the influencing factors on the violence level during the demonstrations.

Based on social science research, which provides qualitative information regarding the factors influencing the violence level in demonstrations, we incrementally present and compare three qualitative models of demonstrations. The first two models are based on an extensive research report initiated by Israeli police [2]. The third is our extension of the second model based on sociological consultation. We evaluated the models on twenty four real-life scenarios. The results show that the first two models make incorrect predictions, but the BIU model makes better predictions on the examined test cases.

Moreover, in this paper we examine whether machine learning techniques such as decision trees may provide better predictions than QR models. While the results show that the machine learning techniques provide accurate predictions (better than our QR models) we will claim, in this paper, that QR approach is sensitive to changes, in contrast to decision trees, and allows *what if* reasoning. Thus, using QR approach is better for reasoning regarding the potential violence level to improve the police decision making process.

2 Related Work

Usage of computer simulation is considered to be a leading approach for modeling and reasoning regarding different social phenomena [6]. There are several micro and macro level techniques that enable such modeling, e.g., usage of agent based simulation, cellular automata and system dynamics. However, there are also techniques that do not require building a model to enable predictions, such as machine learning techniques, in particular a decision tree.

Agent-based simulation is a micro-level approach where by social behaviors are simulated by simulating each individual, and their interactions. By applying agents as an "intelligent" entity we have the ability to model complicated social interactions. Such simulations have been successfully used in modeling crowd behaviors [5, 7], economic phenomena [17], and more. However, it is a bottom-up approach in the sense that to receive a macro-level behavior we must model the micro-level interactions which necessitates detailed individual modeling, and when number of agents is scaled up it may provide significant computational barriers. Furthermore, there are domains such as predicting the likelihood of violence that modeling at the individual participant level may be too high a resolution and even unnecessary.

System dynamics approach [6] is a macro level approach in the sense that it models an entire system. It uses defined stocks, flows and feedback loops to model system behavior. The models are basically sets of differential equations that describe changes in the system. In our domain, such accurate and full definitions are not available.

Qualitative Reasoning (QR) is another macro level approach, allowing modeling and reasoning with partial and imprecise information. It has been used to allow for common-sense reasoning in physics [10, 3]. However, it has also been applied to other branches of science: ecology [14], social science [8], politics [4] etc. However, our use of QR to model and predict the violence level during demonstrations is novel.

Fuzzy Cognitive Maps (FCM) [9] is also a macro level approach which enables causal reasoning using fuzzy directed graphs. Similarly to QR, FCM enables imprecise and qualitative representation of the model. However, the output of FCM is a recommendation on a single action or goal, while QR returns the set of all possible behaviors that the model may manifest.

Machine learning techniques such as decision tree [12] enables reasoning regarding social phenomena without providing a model. Decision tree takes as an input set of properties and build a model, which is set of rules, that allows classification of the observed data according to the given properties. It is mostly used in domains that there no significance for the model and only the classification counts. However, as we show in this paper, prediction (classifications) accuracy is not the only requirement for policy decision-support.

3 Qualitative Reasoning and Simulation

Qualitative simulation enables reasoning about possible system behaviors that can emerge from an initial world state. The simulation takes as input the initial state of the world which contains a structural description of the model and produces a state transition graph. A final state graph captures the set of all possible behaviors that the model may manifest from the initial state. It consists of a set of states and the transitions between them (state-transitions). Each state is a possible unique behavior that the model develops, it contains a unique set of values and inequality statements (quantities) which describe the current behavior of the system. State transitions transform one state into another, by specifying the changes in values and in inequality statements. Each state may contain multiple transitions which enables multiple possible developments of the current state. A sequence of states connected by state transitions where each state is the immediate successor of the one before, is called a behavior path.

Each state is composed of a set of quantities. Quantity is the lowest resolution representation for continuous parameters and it is composed of a pair of values: magnitude and derivative. The magnitude represents the amount of quantity and derivative represents the direction of change. The set of possible values is described by Quantity Space (QS) which is a finite and ordered set of landmark values. Changes in the system are defined explicitly by causal relationships. There are two types of casual relationship between quantities, direct $(I+, I-)$ and indirect $(P+, P-)$ influence. Each influence may be positive $(I+, P+)$ or negative $(I-, P-)$ meaning the derivative of the target quantity increases or decreases accordingly.

In each cycle and on each quantity, all influences (direct and indirect) are combined. When positive and negative influences are combined ambiguities may occur. The qualitative simulation considers all the possible combinations thus, when qualitative description is incomplete, it provides a non deterministic prediction.

4 Modeling violence level in demonstration

Knowledge regarding demonstrations is not accurate nor complete. There are many micro-theories in social science regarding the influencing factors on the violence level: Each such theory focuses on a small sub-set of factors. Integrating all of them into a single unified model is real challenge. The Israeli police initiated a comprehensive study to address this challenge, resulting in a report [2] that provides a collection of factors and their influence on the violence level and also on each other. Their goal was to classify and analyze different kinds of demonstrations in order to propose appropriate methods for the police force in dealing with the mass. They studied 102 crowd events (in particular demonstrations) during the years 2000–2003 and interviews with 87 policemen and police officers. They analyzed a variety of factors that may affect violent behavior, as well as relevant literature. This report is a qualitative collection of factors which provide a challenge to the reasoning process. We use this report as a source of knowledge based on which we developed our models and by using qualitative simulation we provide an ability for reasoning regarding potential violence level.

Base Model. The first (*Base*) model was developed based on the report's literature review [2]. It was proposed there as a first attempt at building a baseline, purely based on literature review. According to the Base model the most influential factors on the violence level during the demonstration are (1) the crowd's a-priori hostility towards the police; (2) willingness to pay the personal price (such as willingness to be arrested); (3) low chance for punishment for violent actions (e.g., a belief that police will not respond strongly); (4) group cohesiveness; (5) previous history of violence. All of these increases the level of violence.

We built a qualitative model based on the presented theory. Figure 1 presents the model's graphical structure. We define one entity (called population) with six quantities, five of them are based on the presented theoretical description and another one is the violence level which is the outcome. For each quantity we defined a quantity space (QS) and defined direct influence (I) between them and the violence level. For example, high hostility for police increases the violence level while low hostility for police decreases. There are also quantities with one directional influence such as previous history where existence of previous history of violence increases the violence level while lack of history does not decrease the violence level and it actually has no effect on it.

Police Model. The second model, Police model, is an extension of the Base model. Karmeli and Ravid-Yamin [2] significantly expanded the Base model, based on their

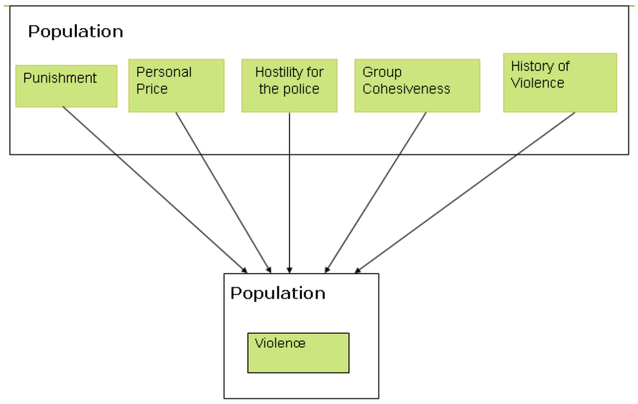


Fig. 1. Basic Model: Structure

interviews with police officers and their investigation into 102 demonstrations. In addition to the factors from the Base model, the Police model adds 12 more variables, roughly divided into several groups. *Environmental factors* include weather, time of day, location sensitivity (e.g., for religious reasons), and time of year sensitivity (e.g., Christmas). *Participant factors* include the number of participants, the existence of violent core among the participants, the existence of group leader, and the cohesiveness of the group (e.g., if they all come from a particular ethnic minority). *Procedural factors* include a request for demonstration license, the purpose of the event (emotional or rational), the timing and strength of police intervention.

The research results showed significant relations between these variables and also their impact on the event outcome (the violence level). For example, political or social demonstrations that express protest or support for leader or cause usually end with low level of violence. However, demonstrations with nationalistic flavor that intend to express emotions (letting off steam) are characterized by much more violent outcomes. The research results also showed a relationship between existence of license and united identity: it was found that some united identities tend to apply for a license before the protest while others do not. It was also found that the time of the day has impact on the violence level; more violent demonstrations occur at night than during the day [2]. A graphic representation of the qualitative model is presented in Figure 2. It shows three entities (Population, Nature and Police) and 18 quantities: 6 are of the Base model, and additional 12 listed above.

Figure 2 presents the model's graphical structure. We defined three entities (Population, Nature and Police) and 18 quantities, where 6 are similar to the Base model and 12 are additional ones. Moreover, based on the research conclusions we defined influence (I) among different variables. For example, emotional event increases the existence of violent core among the participants.

BIU Model. The third model, BIU model, is our own novel extension of the Police model. Based on interviews with social and cognitive scientists, as well as additional literature surveys, we added four additional variables, and updated 19 influences (relations) among the variables. The added factors are: (1) event order (indicates amount of

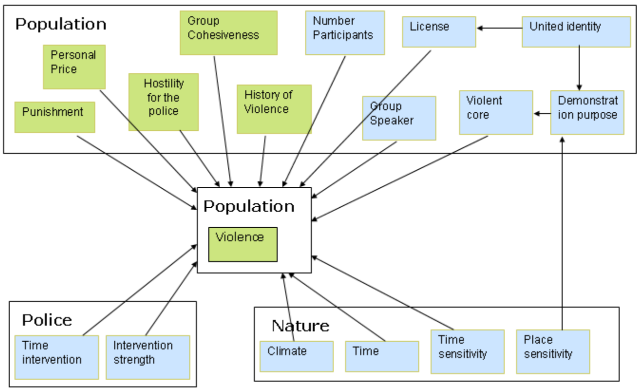


Fig. 2. Police Model: Structure.

preparation that was made following the event, such as delineation, disposition of the police forces etc.) (2) participants anonymity (indicates whether the participants can be recognized), (3) participants' visual cohesiveness (such as similar outfit as among football fans) and (4) light. Moreover, based on sociological consultation, we update the influences among the variables.

Figure 3 presents the model's graphical structure. We used the same entities (Population, Nature and Police) as in Israeli-Police model and added four additional quantities to the existing ones, in addition we updated the quantity space of the police intervention strength and also updated the influences among the variables.

We provide here several examples for updated influences. First, we updated the influence of police's intervention strength, thus instead of direct impact on violence level as in the Police model, it impacts the participants' belief that they may be punished, and their hostility for the police. In BIU model, high intervention strength increases participants' hostility for the police and increases the participants' chance for punishment. However, low intervention strength just decreases the participants' chance for punishment without a change in hostility for the police factor. Another example is that existence of group speaker and existence of license increase the maintenance of order, which decreases the violence level. In contrast, in the Police model, license and group speaker variables had a direct influence on the violence level. Moreover, for the variable *number participants*, we no longer allow direct influence on the violence level as in Police model, but instead have it influence the participants' anonymity level ("the more participants around me the less recognizable I am"). Another example of addition to the BIU model is: participants visual cohesiveness has an impact on group cohesiveness, it actually increases the sense of belonging to the same group.

5 Evaluation

First we wanted to examine whether QR approach may provide successful predictions regarding the potential violence level of demonstrations. We builded the models using QR approach and compare them on twenty four real-life scenarios. Twenty two of them were taken from Hebrew Wikipedia under category demonstrations. The cases were taken both from the main category, and from the subcategories: "demonstrations in

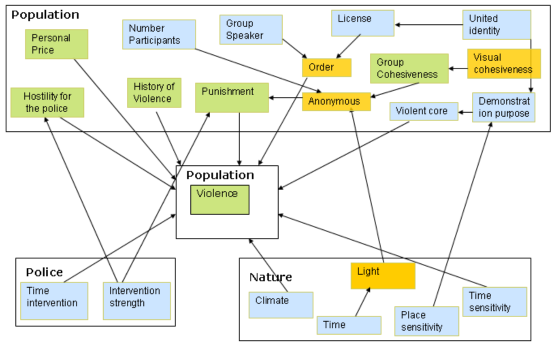


Fig. 3. BIU Model: Structure

Israel" and "massacres in demonstrators". We disqualified general descriptions which are not describe a certain event and also omitted two cases due to lack of information (for a total of twenty cases). Additional three cases are well known events which where extensively analyzed and described [2, 11, 16, 15] by experts. The last event was an event that we participated in and video-taped.

5.1 Models Evaluation

We implemented the models in GARP, a QR engine which enables building and simulating qualitative models and was successfully used in many domains [14, 1]. GARP takes as input an initial setting of the world state (partial state information is acceptable) and produces a simulation state-transition graph. Each sequence of states, following transitions from the initial state and ending with a different outcome state is a possible system trajectory—a possible sequence of qualitative state changes that may occur given the initial state, and the qualitative dynamics specified. The end state in of each such path is where the system dynamics allow no further evolution (i.e., the system is stable). Taking the value of the outcome variables (in our case, violence level) in these final states allow categorical predictions.

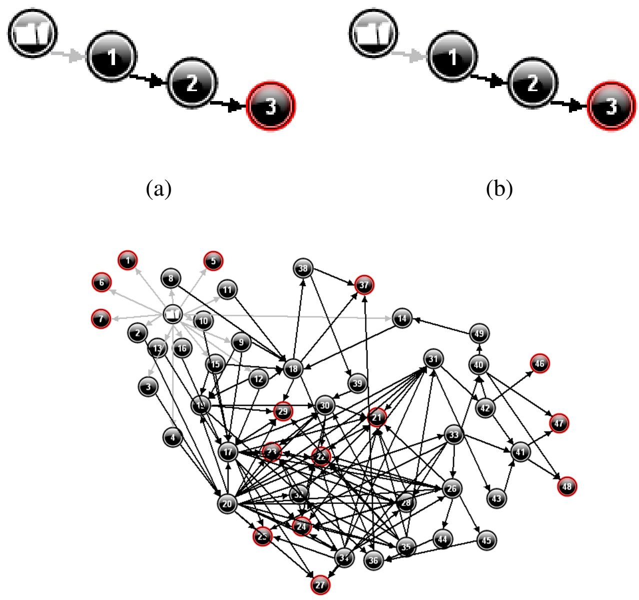
The violence level variable can take three categorical values: zero, low and high. The zero value represents demonstrations that ended without any causalities and also without property damage. The low value represents demonstrations that ended with property damage but without any causalities, and the high value represents all those demonstrations that ended with causalities.

However, it is not enough to know whether a demonstration might be violent; in a sufficiently complex model, all three possible values will have at least one state transition path leading to them. Instead, our goal is also estimate the likelihood of different outcomes. Such knowledge may provide a sufficient addition to the decision making process of the police force. To do this, we use the received state-graph as an input and based on this developed graph we calculate the likelihoods of different outcomes as follows: we count the number of behavior paths that lead to a specific violence level and divide it by the total number of paths.

To initialize the test cases, we utilized the information appearing in their descriptions in the literature and in the Wikipedia. We initialized only the quantities for which we had explicit information; Quantities that we had no information or ambiguous information were removed from the initial set. Qualitative simulation can work with such partial information.

Each model was examined on twenty four test cases as described in Section 5.1. GARP qualitative simulation takes as an input the developed model and its initial state and produces a transitions state-graph. We use this produced state graph for our calculation of the numeric probability as presented in section 5.1.

Figure 4 represents the example of transitions state-graph built by GARP of one of the experiment. Figure 4(a) represents the Base model builded state-graph, Figure 4(b) represents the Police model state-graph in same experiment and Figure 4(c) represents the BIU model state-graph in the same experiment. The circles represents states and the arrows represent state transitions. The end path circles are the final states which contains one of the possible outcomes: no violence, low violence and high violence. For each such builded graph and for each outcome, we calculate the probability as following: we calculate the number of behavior paths that lead to a specific outcome and divide it by the total number of paths.



(a) (b)

(c)

Fig. 4. Transitions state-graph

Table 1 presents the experiment results. The first column corresponds to the examined test case. The second column corresponds to the event outcome as it occurred in real life. Then we present the models predictions, its numeric distribution, for each of the three possible outcome: no violence, low violence and high violence. For each such outcome we present the probability. Below of each experiment, we presented the conclusion regarding models prediction success:

- yes, denoting success
- one level error, corresponding to one ordinal level mistake such as instead of classify to high violence the model classified to low
- two level error, corresponding to two ordinal levels mistake such as instead of classify to high violence the model classified to zero.

Results: The results demonstrate that BIU model made much better predictions than Base and Police models. The Base and Police models made five two level errors in classifications, while the BIU model made just one two level error and four one level error in which instead of classifying to "no violence" it classified "low violence", while the Police model classified three of these cases as high violence.

# exp	Event outcome	Model outcome	**Basic Model**	**Police model**	**BIU model**
Exp1	Very violent event	High violence	100%	66%	70%
		Low violence	0	0	9%
		No violence	0	34%	21%
	Is it correct?		yes	yes	yes
Exp2	Very violent event	High violence	100%	74%	94%
		Low violence	0	5	2%
		No violence	0%	21%	4%
	Is it correct?		yes	yes	yes
Exp3	Very violent event	High violence	100%	75%	61%
		Low violence	0	0	9%
		No violence	0	25%	30%
	Is it correct?		yes	yes	yes
Exp4	Very violent event	High violence	0	66%	83%
		Low violence	0	0	0
		No violence	100%	34%	17%
	Is it correct?		**2 lev. err.**	yes	yes
Exp5	Very violent event	High violence	0	80%	61%
		Low violence	0	0	7
		No violence	100%	20%	32%
	Is it correct?		**2 lev. err.**	yes	yes
Exp6	Very violent event	High violence	66%	66%	87%
		Low violence	0	0	3
		No violence	34%	34%	10%
	Is it correct?		yes	yes	yes
Exp7	Calm event	High violence	66%	80%	8%
		Low violence	0	0	40%
		No violence	34%	20%	52%
	Is it correct?		**2 lev. err.**	**2 lev. err.**	yes
Exp8	Calm event	High violence	100%	78%	59%

		Low violence	0	8	9%
		No violence	0	14%	32%
	Is it correct?		**2 lev. err.**	**2 lev. err.**	**2 lev. err.**
Exp9	Very violent event	High violence	100%	100%	81%
		Low violence	0	0	6%
		No violence	0	0	13%
	Is it correct?		yes	yes	yes
Exp10	Very violent event	High violence	100%	80%	82%
		Low violence	0	0	6%
		No violence	0	20	12%
	Is it correct?		yes	yes	yes
Exp11	Very violent event	High violence	100%	79%	67%
		Low violence	0	7	3%
		No violence	0	14	30%
	Is it correct?		yes	yes	yes
Exp12	Calm event	High violence	0	40%	4%
		Low violence	0	0	64%
		No violence	100%	60%	32%
	Is it correct?		yes	yes	**1 lev. err.**
Exp13	Very violent event	High violence	100%	100%	75%
		Low violence	0	0	7%
		No violence	0	0%	18%
	Is it correct?		yes	yes	yes
Exp14	Very violent event	High violence	66%	66%	63%
		Low violence	0	0	6%
		No violence	34%	34%	31%
	Is it correct?		yes	yes	yes
Exp15	Very violent event	High violence	100%	100%	96%
		Low violence	0	0	3%
		No violence	0	0	1%
	Is it correct?		yes	yes	yes
Exp16	Very violent event	High violence	66%	66%	99%
		Low violence	0	0	1%
		No violence	34%	34%	0%
	Is it correct?		yes	yes	yes
Exp17	Very violent event	High violence	66%	80%	57%
		Low violence	0	0	30%
		No violence	34%	20%	13%
	Is it correct?		yes	yes	yes
Exp18	Calm event	High violence	0%	66%	4%
		Low violence	0	0	64%
		No violence	100%	34%	32%
	Is it correct?		yes	**2 lev. err.**	**1 lev. err.**
Exp19	Very violent event	High violence	100%	66%	73%
		Low violence	0	0	7%

		No violence	0	34%	20%
	Is it correct?		yes	yes	yes
Exp20	Calm event	High violence	0%	33%	23%
		Low violence	0	33%	52%
		No violence	100%	34%	25%
	Is it correct?		yes	**2 lev. err.**	**1 lev. err.**
Exp21	Calm event	High violence	66%	66%	26%
		Low violence	0	0	44%
		No violence	34%	34%	30%
	Is it correct?		**2 lev. err.**	**2 lev. err.**	**1 lev. err.**
Exp22	Very violent event	High violence	66%	79%	83%
		Low violence	0	7	6%
		No violence	34%	14%	11%
	Is it correct?		yes	yes	yes
Exp23	Very violent event	High violence	100%	66%	57%
		Low violence	0	0	8%
		No violence	0	34%	35%
	Is it correct?		yes	yes	yes
Exp24	Very violent event	High violence	100%	66%	56%
		Low violence	0	0	8%
		No violence	0	34%	36%
	Is it correct?		yes	yes	yes

Table 1: **Experiments results**

5.2 Comparison to the Machine Learning techniques

We wanted to examine whether the machine learning techniques such as decision tree may provide a better prediction than our models. We used Weka, an open source software that contains collection of machine learning algorithms and used the J48, decision tree algorithm. We built three files that were used as an input to Weka. Each file contains collection of attributes with their values and was builded based on the quantities initialization set of each QR model (Base model, Police model and BIU model). The target class value of attribute violence in each file was set based on the real-life event outcome. The output of J48 algorithm is the learned decision tree and classification statistics.

Figure 5 present the decision trees that were learned based on the each QR model initial quantity set. Figure 5.2 presents the tree that was learned based on the quantity set of the Base model (Base tree), Figure 5.2 presents the tree that was learned based on the quantity set of the Police model (Police tree) and the same tree was learned based on the quantity set of the BIU model (BIU tree).

The results show that Police tree and BIU tree have 100% of success in classification, while Base tree has 70.8% of success. While the machine learning techniques provide an accurate prediction, a slightly better prediction than the BIU model with QR approach, we will claim, in the next section, that QR approach is much more sensitive

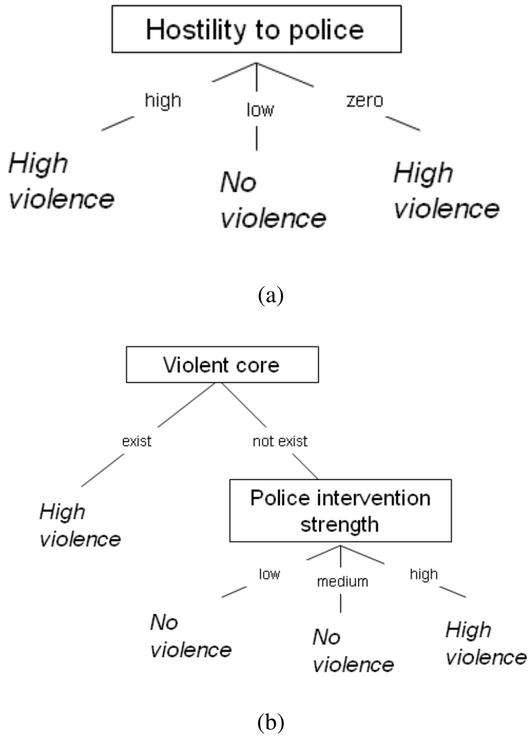


Fig. 5. Decision trees

for changes and can account for *what if* scenarios. Thus, using QR approach is better for reasoning regarding the potential violence level to improve the police decision making process.

5.3 Sensitivity Analysis

In the following experiments we want to demonstrate the use of QR approach and machine learning techniques for variety of hypothetical changes. According to experts [11, 16, 15] in several of the events we modeled (Exp. 15–17), the police intervention strength was found to be one of the important factors for the violence eruption. Thus, in this section, we want to examine the presented QR model's prediction and the machine learning techniques in *what if* scenarios.

First we want to examine whether the presented models with QR approach and machine learning techniques are sensitive enough for the changes in term they can account for *what if* scenarios. Moreover, we want to examine what influence has the police intervention strength on the event outcome, could it be the main factor than can prevent the violence or the event essence to be violent and the police intervention strength has little to do with it? Then we want to examine hypothetical situation of changing the chance for violence in several test cases scenarios by changing different controlled factors and not just the police intervention strength.

Sensitivity Analysis: Experiment 1 In this experiment we want to examine whether the presented models builded with QR approach and the machine learning technique, may account for changes in the police intervention strength. We used the same twenty four test cases as described in Section 5.1 and examined the police intervention strength attribute with it's all possible values. As in Section 5.1 we estimated the likelihood of different event outcomes. The model will consider to be sensitive for the changes if for different values in examined attribute, it will provide different outcome. The change can be one of the following: different distribution values but no change in classification and different distribution values with change in classification.

We compared the BIU and Police models built with QR techniques to decision tree that was built with BUI initialization set. The Base model built with QR techniques is irrelevant for this experiment since the Base model not accounts for the factor of police intervention strength therefore there are no change in the model's predictions.

The results show that Police model changes its distribution in five test cases (from twenty four) and in two of them it also changes its classification. The BIU model changes its distribution in all of the examined test cases and in seven of them it also changes its classification. The decision tree cannot provide distribution for all possible outcomes, it can only provide a final classification, thus unless there was a change in classification the prediction remains the same. From twenty four examined test cases, the decision three change its classification on six of them. Thus, the results show that BIU model is more sensitive to changes than the Police model and the decision tree. However, the question is whereas these models provide a correct changes in the predictions. We answer this below.

We used the three test cases which were explored by experts and for which we have analyzed data ([11, 16, 15]). The first event, test case 15, is the Heysel Stadium Disaster which occurred in 1985 [11]. It was the 1985 European cup final, Liverpool vs. Juventus which was a very tragic and violent event with many casualties. According to Lewis [11] who analyzed this event, one of the reasons for this violent outcome is the police's lack of intervention to prevent the developing violence.

The second event, test case 16, is the Los Angeles Riots which occurred in 1991. This was also a very violent event with many casualties, with 55 killed people and over 2000 injured. Useem [16] who analyzed this event, argued that the police were not properly organized and did not react on time with appropriate force to prevent the eruption, but for the six hours from the beginning of the event, police did little to prevent it which allowed the violent core to grow.

The third event, test case 17, is the London Riot Disaster which occurred in 1990 [15]. As opposed to the previous two events, here the police used enormous force against the protests without distinguishing between anarchists and peaceful marchers. The marchers, with nowhere to go to escape had to fight back. What started as a peaceful protest turned to a very violent event with many casualties.

Table 2 presents the experiment results. The first column corresponds to the examined test case. The second column corresponds to recommended police intervention strength. Then we present the models predictions for each possible outcome: no violence, low violence and high violence. Below of each experiment, we presented whether the recommended reaction changed the model's prediction. "Dist.Change" denotes a change in the distribution, but not in overall classification. "Classif.Change" signifies a change in the classification.

The results demonstrate that the decision tree technique is not sensitive for the examined changes that were claimed by the experts. The Police model performed a slightly better than the decision tree (changed the distribution in test case 15) but failed in two others test cases. However, the BIU model provided a good results which shows that it can account for *what if* scenarios in opposed to the decision tree and the Police model.

Exp.	Recommended Change	Model Outcome	**Police Model**	**BIU Model**	**Decision tree**
Exp15	Increase strength [11]	High violence	66%	83%	
		Low violence	0	6%	
		No violence	34%	10%	
	Classification Dest.Change/Classif.Change/No-Change	High	High **Dist.Change**	High **Dist.Change**	High No-Change
Exp16	Increase strength [16]	High violence	66%	87%	
		Low violence	0	3%	
		No violence	34%	10%	
	Classification Dest.Change/Classif.Change/No-Change		High No-Change	High **Dist.Change**	High No-Change
Exp17	Decrease strength [15]	High violence	80%	19%	
		Low violence	0	45%	
		No violence	20%	36%	
	Classification Dest.Change/Classif.Change/No-Change		High No-Change	High **Classif.Change**	High No-Change

Table 2. Experiments results: changed police intervention strength

Sensitivity Analysis: Experiment 2 In the final experiment we want to examined the hypothetical situation of changing the chance for violence in several test cases scenarios. Specifically, we wanted to examine whether we can decrease even more the violence level in test case 15 (Heysel Stadium disaster) and 16 (LA riots). We used same initializations with several updates as explained below. Some factors such as weather or history of violence cannot be changed, while others can be controlled. For example, police's intervention strength, anonymity, order are examples for features that can be manipulated in the sense that there are actions that can be done to change their values. Police may increase the intervention strength by using more men power or by using different kind of weapon. The existence of projectors and cameras in the gathering zone decrease the perception of anonymity of the participants.

Table 3 presents the experiment results. In this experiment we examined the BIU model and the Decision tree. First column corresponds to the examined test case and the second column corresponds to the changed initial values of the quantities. Then we present the models predictions before the change and after.

# exp	Changed initializations		**BIU Model before change**	**BIU Model after change**	**D.tree before change**	**D. tree after change**
Exp15	Police strength: medium	High v.	96%	80%		
	Punishment: hight	Low v.	3%	6%		
	Anonymity: low	No v.	1%	14%		
	Classification Dest.Change/Classif.Change/No-Change			High **Dist.Change**	High	High No-Change
Exp16	Police strength: medium	High v.	99%	80%		
	Punishment: hight	Low v.	1%	6%		
	Order: hight	No v.	0%	14%		
	Classification Dest.Change/Classif.Change/No-Change		High	High **Dist.Change**	High	High No-Change

Table 3. Experiments results: hypothetic manipulations

Here again the results demonstrate that the decision tree technique is not sensitive for the examined changes which is not surprising since the only components of the learned tree which can change its classification is the existence of violence core and the police intervention strength. However, the BIU model found to sensitive for the changes.

6 Summary and Future Work

In this paper we described a method for modeling and reasoning about social behavior of large groups, and applied it to the problem of predicting potential violence during demonstrations. We used qualitative reasoning (QR) techniques, which to our knowledge have never been applied for modeling crowd behaviors, nor in particular demonstrations. Based on social science research, we incrementally presented and compared three QR models for predicting the level of violence in demonstrations: A Base model, Police model and BIU model. We evaluated these models on twenty four real life test cases scenarios. The results show that BIU model makes better predictions than the compared models and it also was found out to be sensitive for changes. We also compared our performances to the machine learning method, a decision tree. While, the machine learning method made an accurate predictions, it fails in the sensitivity analysis. Thus, the BIU model builded with QR approach can account for *what if* scenarios is opposed to the decision tree and is more preferable for reasoning regarding the potential violence level to improve the police decision making process

In our future work we plan to expand our model to account for bidirectional influences (feedback loops). For example, in the BIU model the ”hostility for the police” quantity increases the violence level. However, increasing the violence level has no impact on hostility. We believe that such expansion is necessary to provide a more accurate prediction. Also, in our future work, we plan to provide a statistical analysis of the developed state-graph and enables reasoning regarding the developed process and not only regarding the final outcome. The third direction is to expand our evaluation process by testing model on additional real-life test cases.

References

1. B. Bredeweg and P. Salles. Mediating conceptual knowledge using qualitative reasoning. In S. Jrgensen, T.-S. Chon, and F. E. Recknagel, editors, *Handbook of Ecological Modelling and Informatics*, pages 351–398. Wit Press, Southampton, UK, 2009.
2. A. Carmeli and I. Ravid-Yamin. Research report on the subject of crowd events and public order. Technical report, Ministry of public security, Bureau of the chief scientist, 2006.
3. K. D. Forbus. Qualitative reasoning. In *CRC Handbook of Computer Science and Eng.,*. CRC Press, 1996.
4. K. D. Forbus and S. E. Kuehne. Towards a qualitative model of everyday political reasoning. In *Proceedings of the Nineteenth International Qualitative Reasoning Workshop*, 2005.
5. N. Fridman and G. A. Kaminka. Towards a cognitive model of crowd behavior based on social comparison theory. In *AAAI-07*, 2007.
6. N. Gilbert and K. G. Troitzsch. *Simulation for the social scientist*. Open University Press, 2005.
7. W. Jager, R. Popping, and H. van de Sande. Clustering and fighting in two-party crowds: simulating the approach-avoidance conflict. *Journal of Artificial Societies and Social Simulation*, 4(3), 2001.
8. J. Kamps and G. Péli. Qualitative reasoning beyond the physics domain: The density dependence theory of organizational ecology. In *Proceedings of QR95*, 1995.
9. B. Kosko. *Fuzzy cognitive maps: advances in theory, methodologies, tools and applications (studies in fuzziness and soft computing)*. Springer, 2010.
10. B. Kuipers. Qualitative reasoning: modeling and simulation with incomplete knowledge. *Automatica*, 25(4), 1989.
11. J. M. Lewis. A value-added analysis of the heysel stadium soccer riot. *Current psychology*, 1989.
12. T. M. Mitchell. *Machine learning*. McGraw-Hill, 1997.
13. S. Patrick, P. M. Dorman, and R. L. Marsh. Simulating correctional disturbances: the application of organization control theory to correctional organizations via computer simulation. *Journal of Artificial Societies and Social Simulation*, 2(1), 1999.
14. P. Salles and B. Bredeweg. Modelling population and community dynamics with qualitative reasoning. *Ecological Modelling*, 195:114–128, 2006.
15. C. Stott and J. Drury. Crowds, context and identity: Dynamic categorization process in the poll tax riot. *Human Relations*, 2000.
16. B. Useem. The state and collective disorders: The los angeles riot/protest of april 1992. *Social Forces*, 76, 1997.
17. J. Wander. *Modelling consumer behavior*. PhD thesis, University of Groningen, 2000.

An Analysis and Design Framework for Agent-based Social Simulation

Amineh Ghorbani, Virginia Dignum, and Gerard Dijkema

Delft University of Technology,
Faculty of Technology, Policy and Management,
Delft, The Netherlands
a.ghorbani@tudelft.nl

Abstract. Agent-based modeling is one of the popular tools for analyzing complex social systems. To model such systems, social attributes such as culture, law and institutions need to implemented as part of the context of a MAS, independently of individual agents.
In this paper, we present MAIA; a methodology for agent-based modeling based on the Institutional Analysis and Development Framework (IAD). The IAD is a well established comprehensive framework which addresses many social attributes. To make this framework applicable to agent-based software implementation, we inspire from some of the detailed definitions in the OperA methodology.
The MAIA framework consists of two layers. A conceptualization layer for understanding the system and a detailed design layer which leads to the implementation of social models. The framework covers the different types of structures affecting agents at the operational level; physical, collective and constitutional. Moreover, this framework includes the conceptualization and design of evaluation.
MAIA allows the balance of global institutional requirements with the autonomy of individual agents thus enabling system evolution and reflecting more of reality in artificial societies.

Key words: Agent-based modeling, methodology, social simulation, IAD

1 Introduction

Socio-technical systems are complex adaptive entities that require the engagement of social and technical elements in an environment to reach their goals [23]. These systems are not easy to analyze and understand due to the complex and unpredictable relationship between their elements. Agent-based modeling is an increasingly popular tool to emulate socio-technical systems and explore phenomena concerning complex relations between entities [3, 12]. Examples of such phenomena include societal response to accidents and disasters [7] or the introduction of new policies and regulations [1].

Even if simulation has been described by some as a third way of doing research as opposed to induction and deduction [2], the process of doing simulation is far from standard and is not without its caveats. Some of the risks include failure to state clear objectives and research questions, taking a too complex or too simplistic approach, and misinterpretation of results [29].

In the particular case of social simulation, the following design guidelines have been identified. First, in order to create a useful simulation, one needs to think through the assumptions clearly [14]. Currently, in many practices not only the assumptions are not clear to the modeler or domain expert, they are also not documented. Second, a simulated model needs to be open to inspection by others, replicable and reusable (parts or whole) which is currently not the case for many ABMs [17]. Third, emergence in social systems is different from emergence in physical systems where there is no response to emergent phenomena from the individual entities. In human societies, emergent phenomena is identified, understood and reacted to. For example, people would respond to culture, institutions or policies which are emergent properties of a society [16]. Individual beings in social systems also vary greatly in many aspects such as thinking, capabilities and needs unlike physical systems. All these aspects regarding emergence need to be taken into consideration for social simulation [14].

The feasibility of a simulation of socio-technical systems increases when proven theory are used to ground factors that need to be considered and are important for the model [14]. This should include the representation of social issues such as norms and institutions which require theoretical background; and theories to link micro and macro behavior that are able to reconcile intentionality, deliberation, and autonomous planning with playing social functions and contributing to the social order [6]. Methodologies for social simulation that provide such theoretical basis are non-existing [17].

In this paper, we introduce MAIA, a conceptually rich and systemic *M*ethodology for *A*gent-based modeling based on *I*nstitutional *A*nalysis. MAIA covers the ABM development cycle, structured into two phases: conceptualization and analysis, and detailed implementation specifications. Through out the process, diagrams and tables are drawn and filled in, to give a clear representation of what the system to be model is about and how it will be modeled.

In response to the three issues raised above, MAIA first provides a structure for conceptualization and design of systems which makes the assumptions made by the modeler clear. Through the diagrams and tables produced using MAIA, one has a documentation of the decomposed system with all the assumptions which can also be communicated to the domain experts for verification before model implementation. This also makes the model open to inspection by others even after implementation.

Second, this methodology is explicitly used for developing models of socio(-technical) systems. Therefore, it addresses emergence in human societies where there is response to emergent phenomena through feedback loops between different structures explicitly mentioned in the methodology (e.g. whether an agent would follow an norm or not). To answer the third issue, we have based MAIA on the institutional analysis and development(IAD) framework. The IAD, developed by Nobel laureate Eleanor Ostrom [25], is a well established framework that is widely used in institutional economics and the social sciences. The purpose of the IAD is to explain all the major components of a social system that affect social structures, in order to understand, analyze and (re)design institutions [24].

One major benefit of using the IAD for agent-based models is that it can be considered as a blueprint which covers a diversity of social concepts and explains how these relate to individual agents. This theoretical base of MAIA also attracts more acceptance of this framework in the social science community. However, institutional frameworks including the IAD have a macro nature in terms of analyzing systems for policy making. Although the IAD covers the social concepts required to analyze a system, many details are required to link to the micro level behavior and develop agent-based models. This is why we inspired from the OperA [10] methodology for agent-based software development to further develop MAIA. OperA aims to cover the social aspects of agent-based software development and thus meets our requirements.

As well as the contributions mentioned above, the MAIA methodology has improvements for the verification and validation of agent-based models as these procedures are considered from the beginning of the modeling practice. Also, this methodology is independent of programming language or software and can be used decompose a system for any kind of modeling environment.

The paper is structured as follows. In section 2 we introduce the IAD framework and explain the OperA methodology. In section 3 we introduce the MAIA framework meta-model and explain its details through a example. In the last section we discuss the potential benefits of this methodological framework for agent society design and give directions for future research.

2 Background

In this section we introduce the IAD framework and the OperA methodology which the MAIA framework has inspired from.

2.1 The Institutional Analysis and Development Framework

The term institution has become widespread in the social sciences in recent years which reflects the growth in institutional economics and the use of the institution concept in several other disciplines, including philosophy, sociology, politics [18]. Ostrom [25] defines an institution as "the set of rules actually used by a set of individuals to organize repetitive activities that produce outcomes affecting those individuals and potentially affecting others". Agreements or rules can be called institutions only if they are accepted by those involved, are used in practice, and have a certain degree of durability [21]. Therefore, with this definition, institutions are emergent properties of social systems that can range from an explicit rule in a company to the implicit culture of the whole society.

Institutions can also be considered as set of rules that influence, guide and limit the behavior of actors. Institutions have two sides: they enable interactions, provide stability, certainty, and form the basis for trust and on the other side, they may also cause power relations. If institutions fail to fulfill stability or bring about mobility of bias, there is ground for institutional (re)design [20].

Institutional redesign refers to deliberate changes in institutional characteristics. In order to (re)design institutions, one should be able to understand and analyze them. The institutional frameworks are developed for this purpose. One of the well-known frameworks is the institutional analysis and development framework (IAD) by the 2009 Nobel laureate Ostrom. The IAD [25] clearly specifies different elements of the system description unlike other institutional framework such as Williamson's four layer model, which has many similarities but offers less detail [21]. The IAD has become one of the most accepted and used frameworks in institutional design and policy making [19, 28]. Also, many social scientists have built agent-based models for common pool resource management based on the IAD [22, 4, 9]. In [13], institutional frameworks are compared and the benefits of applying them for modeling socio-technical system is discussed.

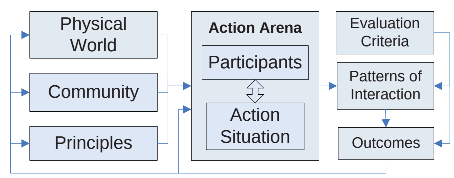


Fig. 1: The IAD framework

Figure 1 illustrates the IAD framework. The central concept is the 'action arena', in which individuals (or organizations) interact, exchange goods and services, solve problems, or fight. The action arena is described by the participants (who have a set of resources, preferences, information, and selection criteria) and the action situation: the actual activity (or 'game') that is to be analyzed. The action situation consists of roles (or positions), actions, information related to the situations, potential outcomes of the situation and the costs and benefits related to it.

What happens in the action arena leads to patterns of interaction and outcomes that can be judged on the basis of evaluative criteria. The action arena itself is influenced by attributes of the physical world (e.g. climate, technological artifacts), the attributes of the community in which the actors/actions are embedded , and the set of principles that the individuals involved use to guide and govern their behavior (e.g. norms, institutions).

Although physical world and community influence the action arena, it is the principles of the game that actually define it. Therefore, in IAD quite some attention is given to these principles. Principles are statements about what actions are required, prohibited, or permitted and the sanctions authorized if they are not followed. Seven distinct types of principles are distinguished:

- Boundary: Specify who is eligible to play a role: who is in and who is out of the game.

- Position: Determine to what extent a distinction is made regarding the position of the different participants. For example, a buyer or seller on a market have a different role than an auctioneer (and thus different access to information, and different choices).
- Choice: Specify what a participant must, must not, or may do at a specific point of the process.
- Payoff: Assign external rewards or sanctions to particular actions that have been taken.
- Information: Describe what information may or may not be shared by participants and whether they have a set of common, shared information.
- Scope: Define what outcome variables should or should not be affected by the actions undertaken.
- Aggregation: Specify who has responsibility for an action: for example, whether a single participant or multiple participants should come to a decision.

These principles can be analyzed within three distinct layers: the operational, the collective choice and the constitutional choice levels. These layers are equivalent to the four layers of Williamson's model [21]. The different levels relate to different time-frames: day-to-day activities fall within the operational level (Williamson's layer 1), the collective choices determine what operational activities take place and these are reviewed over a 5-10 year time frame(Williamson's layer 2), whereas the constitutional level determines how the process of collective choice is organized which is a long-term process (the two top layers of Williamson).

Ostrom claims that the IAD framework with the details briefly described above, covers all the aspects that have influence on giving a system the social and institutional backbone [25]. Thus, if we design an artificial system that implements these components, we are one step closer to shaping social systems in agent-based models.

2.2 The OperA framework

The OperA framework [10] proposes an expressive way for defining open organizations distinguishing explicitly between the organizational aims and the agents who act in it. That is, OperA enables the specification of organizational structures, requirements and objectives, and at the same time allows participants to have the freedom to act according to their own capabilities and demands.

The OperA framework consists of three interrelated models illustrated in figure 2. The **Organizational Model** (OM) is the result of the observation and analysis of the domain and describes the desired behavior of the organization, as determined by the organizational stakeholders in terms of objectives, norms, roles, interactions and ontologies. The **Social Model** (SM) maps organizational roles to agents and describes agreements concerning the role enactment in social contracts. An agent playing a role is called a role-enacting agent or *rea* for short. Finally, the Interaction Model (IM) specifies the interaction agreements between reas as interaction contracts. The IM specification enables variations

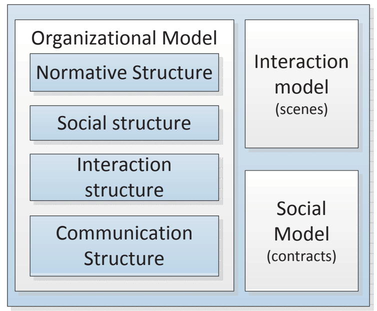


Fig. 2: The OperA Overall Architecture

to the enactment of interactions between reas. The MAIA methodology inspires from the OM and to some extent addresses the SM.

The Organizational Model In OperA, the OM specifies the *means* to achieve such objectives. That is, the OM describes the structure and global characteristics of a domain from an organizational perspective. Organizational objectives are achieved through the action of agents. The organizational model consists of four structures that are briefly described in the following. The design and validation of OperA OMs can be done with the OperettA tool [11] [1].

The *social structure* of an organization describes the roles holding in the organization. Role description should identify the activities and services necessary to achieve society objectives and enable to abstract from the individuals that will eventually perform the role. The definition of a role consists of an identifier, a set of role objectives, possibly a set of sub-objectives per objective, a set of role rights, a set of norms and the type of the role.

The *interaction structure* describes a partial ordering of meaningful scene scripts. A *scene script* describes a scene by its players (roles), its desired results and the norms regulating the interaction. In OperA interaction descriptions are declarative, indicating the global aims of the interaction rather than describing exact activities in details. Interaction objectives can be more or less restrictive, giving the agents enacting the role more or less freedom to decide how to achieve the objectives and interpret its norms.

The *normative structure* defines the norms that regulate roles, and that specify desired behavior that agents should exhibit when playing the role. Norms [2] are specified using a temporal deontic logic. Norms describing obligations, permissions or prohibitions, consist at least of an activation condition, an expiration condition and a maintenance condition.

The aim of the *communicative structure* is to describe the communication primitives. That is, to describe the set of performatives and the domain language

[1] As mentioned in [17], tool support is one of the current shortcomings of ABM.
[2] norms are the same as institutions according to the definition given by Ostrom.

specific to the society, to be used in the communication between role enacting agents. It includes both a knowledge representation language (to describe the domain ontology) and a communication language (to specify the interactions among agents).

The Social Model The Social Model specifies how agents are going to enact roles in an organization. Agent capabilities must be checked against role requirement and roles are assigned on the basis of this. We call agents that want to play roles in the organization applying agents. In particular, it must be checked that applying agents have the required capabilities. The resulting agreement is fixed in a social contract between agent and organization, against which the activity of the agent at runtime can be evaluated. Note that agents are still free to decide on compliance or violation of their social contracts. Different agent's 'personalities' will result in different role enactment behaviors, from social to fully egoistic.

2.3 Theoretical grounding for social simulations: IAD and OperA

We propose a combination of IAD and OperA as an answer to the need for a theoretical grounding of social simulation models identified in the introduction section.
The IAD framework is widely used for analyzing systems that need to be socially restructured by policy makers or economists. Even though there are sufficient details to act as a blueprint and describe a system, the concepts are too general to give formal definitions that can be implemented in a computer model. As an example, the IAD defines a community in general, but the definition of the details within a community such as groups, roles and their dependencies or the way individuals interact with each other is not given.

The OperA framework on the other hand, has a formal model that defines some social concepts in detail including groups and role attributes which are essential for build computer models. However, OperA does not consider all the required concepts of a social system discussed in the IAD such as agents or the necessary physical world [15].

Generally speaking, the development of agent-based social simulations is an iterative and interactive process, that follows the four main phases of software development: Model Conceptualization, Design, Construction and Evaluation. During conceptualization and design, the main components of the system are identified and specified. To develop a comprehensive methodology for conceptualization and design, there are several aspect that the framework (i.e. the combination of IAD and OperA) is required to cover:

- Social concepts such as roles, laws, regulations and culture should be explicitly represented [6, 8, 10].
- A representation of the physical environment is just as important as the social structure for a socio-technical system [15].

– The design of individual agents, their decision making and information processing capabilities is required (e.g. [27]).
– The design of the evaluation domain as well as the actual model itself [17].
– The design of the actions and interactions between the agents of the artificial society (e.g. [10, 30]).

The linkage between the two frameworks is illustrated in table 1. The general concepts in this table are the outcome of a literature survey on the general terms used in system analysis for institutional design. The similarity between the core concepts is the reason OperA was selected for the refinement of the IAD out of the many existing agent-based software methodologies.

Table 1: Differences between the IAD and OperA

Concept	**OperA**	**IAD**
Physical Structure	not mentioned	Physical World
Collective Structure	Communicative structure (ontology, CA, contracts)	Community mentioned (without detail)
Constitutional Structure	Normative/Social Structure (norms)	Principles
Operational Structure	Partly by interaction/social model (land marks, transition rules)	Action Arena
Evaluation Structure	not mentioned	Evaluative Criteria/ Outcome

3 The MAIA Framework

The MAIA framework is an agent-based development methodology consisting of two iterative steps. The first step is aimed at the conceptualization and analysis of the system to be modeled and the second step is a design phase which gives the detailed specifications for the implementation of agent-based models. This paper introduces the conceptualization and analysis phase of the MAIA framework by explaining the meta-model of MAIA. Considering table 1, the MAIA framework has contributed to the physical and evaluation structures by introducing almost all the concepts, the collective structure is mostly a combination of our own specifications and OperA. The constitutional and operation structures are a combination of OperA, own definitions and the IAD.

3.1 Example: Consumer Lighting

The example that will be used for more clarification throughout this paper, is a policy analysis problem, which explores the effects of government policies on the transition to low-electric power consumption consumer lighting [1]. To provide lighting, a variety of illumination technologies are available. With time, each of these technologies improve with respect to cost, energy efficiency, operating

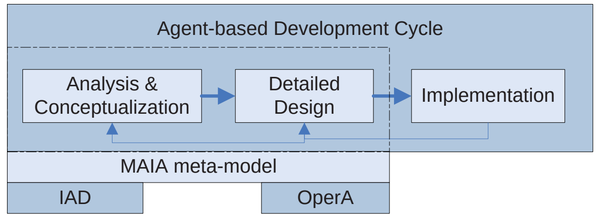


Fig. 3: The MAIA agent-based modeling cycle

hours etc. Consumers buy lamps from retailers based on some selection criteria. Consumers communicate their light bulb experiences within their social network. This creates a word-of-mouth effect that influences consumer buying patterns. Meanwhile, the European Union wants all consumers to replace their lamps with low-energy consumption LEDs. The question on institutional if not policy design then becomes: what kind of strategy works best to effect such transition; subsidy on LEDs, tax on standard bulbs or banning the standards bulbs altogether? Of course, we will not be addressing the whole problem in this paper due to space limitation but merely addressing this example to make the different parts of the framework more clear.

3.2 MAIA Overview

According to the concepts defined in institutional frameworks (including IAD and Williamson) [13, 21] and the ones we see in agent-based methodological frameworks [10], the general structures required to conceptualize and design social system models are presented in figure 4.

The physical structure of a social system is where the physical entities of the system that influence the outcomes(e.g. department, house, money) are defined. The collective structure defines the actual agents in the system, their social network and the common language (ontology) between them. The constitutional structure introduces the social structure of a model including roles, implicit institutions (e.g. culture), explicit institutions (e.g. laws, regulations), groups and role dependencies.

These three structures influence the operational structure of a society where decisions are made and actions are taken. The events in the operational structure result in certain outcomes that are evaluated based on some criteria. The evaluation structure deals with measures that need to be considered when designing an agent-based model [17]. This leads to a more or less ad-hoc approach to evaluation. This setup is not considered in current agent-based modeling practices which mostly consider the operational and physical structures [23, 1]. Likewise, the evaluation structure is not addressed in agent-based software methodologies (e.g. [10, 5, 30]).

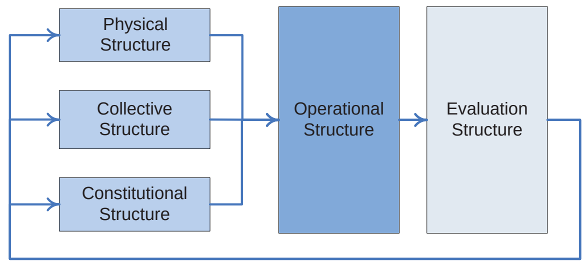


Fig. 4: MAIA Overview

3.3 MAIA: Analysis and Conceptualization Phase

Due to space limitation, we briefly define the meta-model concepts which would be used to conceptualize a system and eventually lead to the design of agent-based models. The concepts in the MAIA framework are presented in tables and graphs (e.g. a role table, an institution table etc...). In the design phase of the MAIA framework, these concepts would be represented in a more programmable form. The process of conceptualization flows according to the sequence of concepts discussed in this section.

Physical Structure The physical structure is made of three distinct concepts namely resource, location and links(illustrated in figure 5). This classification of the physical environment is the result of long term modeling experience (implementation of more than 25 agent-based models). A *resource* has name, property and type (private (e.g. money) or common pool (e.g water, electricity) [25]). A *location* (e.g. house) can have properties such as size, coordinate or capacity. A *link* is a physical connection that can be between resources or locations (e.g. electricity grid linking houses). Each location has a begin node and an end node.

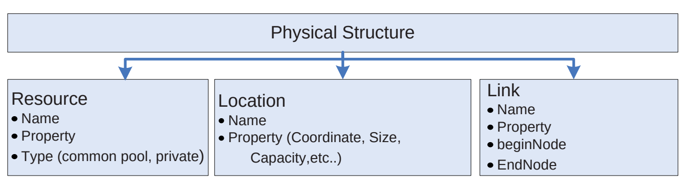


Fig. 5: Meta-model Concept: Physical Structure

Constitutional Structure Generally speaking, the roles the agents take, the groups they form, the dependencies they rely on and the institutions they follow in the social context are defined in the second component called the constitutional structure (see figure 6). The first step is to define the institutions of the system. Institutions (defined in OperA as norms) can be formal(i.e. law, regulation etc) or informal(i.e. culture). A one year warranty on light bulbs or price regulations are formal institutions. Culture(informal institution) makes neighbors talk to each other about the goods they have bought. Fashion is another example of the informal constitutional environment in the society. Institutions can prohibit, oblige or permit roles to do certain tasks.

A role is defined independent of the underlying agent (according to OperA and IAD). Roles have predefined objectives which can be in contradiction with the agent goals. A role can be external to the system which means that agents cannot not take that role, or internal. The boundary rule of a role defines when an agent can be considered as having the role. While the institutions of a role can be ignored by the agent taking that role (e.g. customer not paying for a light bulb) the rights of a role have to be followed by the agent (e.g. customer returns a light bulb with guarantee).

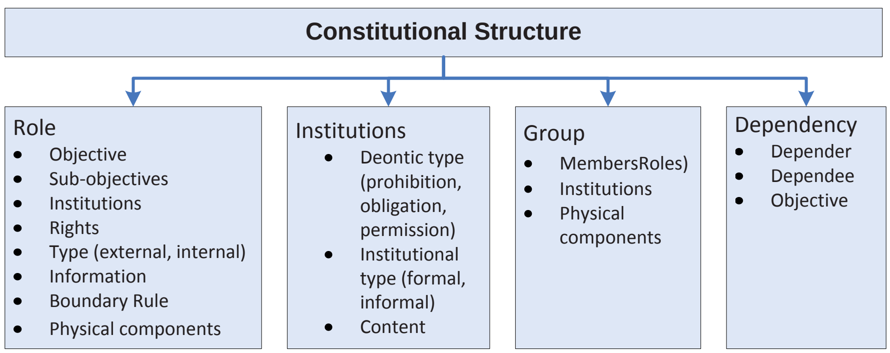


Fig. 6: Meta-Model Concept: Constitutional Structure

To reach their objectives, roles depends on other roles. A role dependency graph for the lighting scenario is depicted in figure 7. The nodes of the graph represent the roles and the edges are the objectives. A link is drawn between two nodes if a role depends on some other role to achieve his objective [10]. Sometimes, an institution may belong to more than one role. In such case, a group is defined. The word of mouth culture cannot happen unless there is a group called friends or neighbors. When a resource belongs to more than one role, it is given to a group (e.g. light bulbs in the corridors of a building complex).

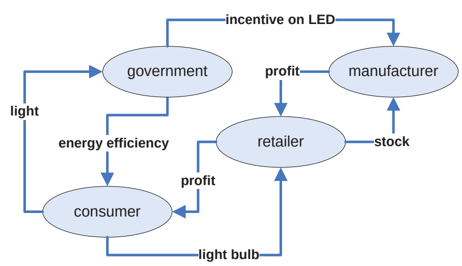


Fig. 7: Dependency graph

Collective Structure This structure specifies the agent, the social network and the domain ontology (shown in figure 8). Agents can take different roles. Each agent has some preferences which may also affect how he would follow the institutions of the role he is taking. For example, an agent taking the role of a consumer can have wealth as his priority instead of quality. So he would buy a cheaper lamp or, he may even steal a lamp (not following the payment institution of a consumer). The preferences of the agent together with his properties such as age or eduction define the personality of an agent which affects the decision making process. The information processing capability of an agent is the amount of information one may have access to. For example, each consumer knows three shops in his surroundings. Similar to roles and groups, agents can also have resources. A car belongs to an agent not a consumer since not every consumer may have one. The number of agents taking a specific role is also defined in the agent table (e.g. 1000 consumerAgents taking the role Consumer).The link between agents and roles is one of the most important aspects of the MAIA framework which will be discussed in the next section.

The social network is a graph where the sender and receiver are agents:roles pairs and the link between two nodes gives a general description of the interaction between two pairs. For example, a consumer *buying lamp* from a shop keeper. Finally, a domain ontology (also present in OperA) defines the common terms and concepts of the system.

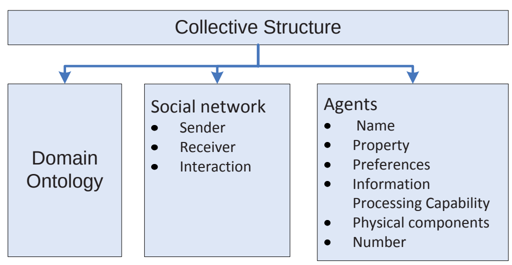


Fig. 8: Meta-model Concept: Collective Structure

Operational Structure The general focus of the operational structure is on the continuous activities of the system such as decision making of the participating agents and the functions that link decisions to outcomes.

This structure is divided into three main components, the action situation, the role enactment agreement and the action sequence. In the action situation, the actual processes of the system take place [25]. Roles take actions which have some costs and may result in benefits. Each action situation may have its own institutions and resources. In a role enactment agreement table, the link between agents and roles is defined. Each agent can take one or more roles and vice versa. The specifications of the role and the agent are all present in this table and each possible combination occupies one row of the table. In the behavior identification column, the decision making process of the agent is defined. This can be done by giving weights to different components or reading external data. The agents can take actions that are forbidden as a part of the role institutions. Since this is done explicitly in the model it is thus traceable.

The action sequence component which is used in the detailed design phase defines the actual interaction between agent:roles. The interaction is the message exchanges and the action is an atomic process taken by only one agent. The sequence diagrams are built according to the social network diagram in the collective structure and the actions defined here.

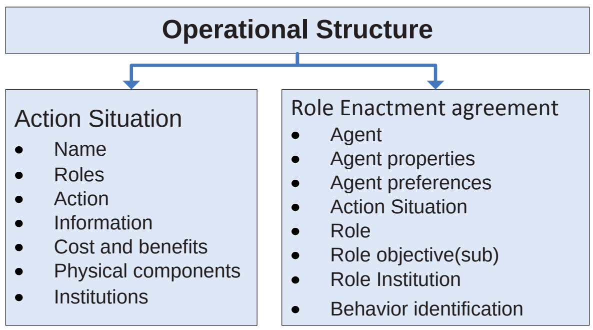


Fig. 9: Analysis and Conceptualization phase: Operational Structure

Evaluation Structure Model evaluation is a step that is often addressed after the software has been implemented and not taken into consideration at the conceptualization or even design phase. However, some agent-oriented methodologies support the development of systems that are required to be robust, flexible and efficient [26]. While for multi-agent software systems, such properties are performance indicators, in the case of agent-based modeling, these factors may not be as critical.

The evaluation indicators of an agent-based model can be divided into two different categories:

1. *Reality closeness parameters*: When modeling a real world system, we first have to make sure that the artificial society is functioning as close to the real system as possible. Choosing the relevant parameters to evaluate the closeness of the model to reality is a function of this category. These parameters are suggested by the domain experts. In the consumer lighting example, light bulb consumption period is chosen as one of the closeness parameters. If the consumers are using the normal light bulbs longer than a certain period that reflects real world light bulbs, then the number is an indicator of some difference to the real system. It is important to realize that the value of these parameters are outcomes of the system and will be calculated after (or during) the model run. The values are not given during the implementation of the model.

2. *Problem domain parameters*: The purpose of the modeling exercise is to answer a set of questions defined by the problem owner. To better understand these questions and make efforts to answer them, certain parameters have to be assigned to each question [16]. These parameter values are again the outcome of the system and only realized during runs or even after. A problem domain parameter for the consumer lighting example is the number of LEDs bought at the end of the game. This parameter is one of the parameters that indicates whether a certain policy has been successful in the transition to LED light bulbs.
   The scope principle of the IAD framework [25] links the actions taking place in the operational structure to the problem domain parameters. This information is shown in a scope matrix. A small matrix is illustrated in table 2 for the shopping situation in the consumer lighting example. The purpose of the scope matrix is to document the relation between the operational events and the global outcomes. Through this matrix one would have a general idea of which local behaviors may lead to certain emergent outcomes. Because of the feedback loops between the structures and also between the different phases of the MAIA development cycle, the matrix would change quite often.

The method of filling in such matrices is left out due to space limitation. The feedback received from these two categories of parameters can be reflected in the physical, collective and constitutional structures of the MAIA framework to further affect agent actions and interactions at the operational structure.

Table 2: Scope Matrix for the shopping situation of consumer lighting

Action/ Domain Parameter	Select lamp	Pay money	Search shops
No. LEDs bought	X		
No. shops (LED)	X	X	X
Income for LED		X	

4 Conclusion

In this paper MAIA, a methodological framework for the conceptualization and design of agent-based society modeling is proposed. The IAD framework for institutional analysis is selected as the theoretical basis for its comprehensiveness and acceptance among the social scientists. For some of the formal details, the OperA methodology was adapted where applicable.

The MAIA framework is especially aimed at building artificial societies for modeling purposes. This framework, which has two different levels of abstraction, can be applied to conceptualize a social system for the design of an artificial society. The major distinctions of MAIA are: first, by following the tables and diagrams, one has a clear documentation of the system decomposition which can be communicated to the domains experts for verification. Second, by building on top of the IAD framework, we aim to cover all the required aspects of a system that forms and evolves social structures. Third, evaluation is considered as a structure that is taken into account from the conceptualization phase [17]. Fourth, there is a clear distinction between the social structure and the individual agents, which helps the exploratory analysis of society formation and evolution while giving enough flexibility for properties to emerge.

In general, the MAIA methodology has been developed to provide an agent-based modeling tool that is independent of software. We provide a common language that extends across domains and is standard so that agent-based models can be documented and reused as required [17]. Even though we have used the IAD concepts which have already been validated many times, formal validation is still necessary. We are also experimenting this framework with different case studies (e.g. e-waste management, woodfuel in forestry) to find out the limitations and drawbacks.

The second level of the MAIA framework is currently in development. This level is for the detailed design of agent-based models which would link conceptualization to programming code. Further specifications of the evaluation structure is one other subject of future research.

References

1. M. R. Afman, E. J. L. Chappin, W. Jager, and G. P. J. Dijkema. Agent-based model of transitions in consumer lighting. In *3rd World Congress on Social Simulation, Kassel, Germany*, 2010.
2. R. Axelrod. Advancing the art of simulation in the social sciences. *Complexity*, 3(2):16–22, 1997.
3. E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280, 2002.
4. F. Bousquet, I. Bakam, H. Proton, and C. Le Page. Cormas: common-pool resources and multi-agent systems. *Tasks and Methods in Applied Artificial Intelligence*, pages 826–837, 1998.
5. P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos. Tropos: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3):203–236, 2004.
6. C. Castelfranchi. The theory of social functions. challenges for multi-agent-based social simlation and multi-agent learning. *Cognitive Systems*, 2001.
7. L. Comfort, K. Ko, and A. Zagorecki. Coordination in rapidly evolving disaster response systems: The role of information. *Agent-Based Simulation: From Modeling Methodologies to Real-World Applications*, pages 208–219, 2005.
8. R. Conte and C. Castelfranchi. Understanding the functions of norms in social groups through simulation. *Artificial Societies*, London: UCL Press.:252–267, 1995.
9. P. Deadman, E. Schlager, and R. Gimblett. Simulating common pool resource management experiments with adaptive agents employing alternate communication routines. *Journal of Artificial Societies and Social Simulation*, 3(2):2, 2000.

10. V. Dignum. *A model for organizational interaction: based on agents, founded in logic*. PhD thesis, 2004.
11. V. Dignum and H. Aldewereld. Operetta: Organization-oriented development environment. In *Proc. ofLADS'2010@Mallow*, 2010.
12. J. Epstein. *Generative social science: Studies in agent-based computational modeling*. Princeton Univ Pr, 2006.
13. A. Ghorbani, A. Ligtvoet, I. Nikolic, and G. Dijkema. Using institutional frameworks to conceptualize agent-based models of socio-technical systems. *Proceeding of the 2010 workshop on complex system modeling and simulation*, 3:33–41, 2010.
14. N. Gilbert. Agent-based social simulation: Dealing with complexity. *http://www.complexityscience.org/*, 2005.
15. N. Gilbert, S. Schuster, M. Den Besten, and L. Yang. Environment design for emerging artificial societies. 2005.
16. N. Gilbert and P. Terna. How to build and use agent-based models in social science. *Mind & Society*, 1(1):57–72, 2000.
17. B. Heath, R. Hill, and F. Ciarallo. A survey of agent-based modeling practices (january 1998 to july 2008). *Journal of Artificial Societies and Social Simulation*, 12(4):9, 2009.
18. G. Hodgson and J. Calatrava. What are institutions. *Journal of Economic Issues*, 40(1):1, 2006.
19. M. Imperial. Institutional analysis and ecosystem-based management: the institutional analysis and development framework. *Environmental management*, 24(4):449–465, 1999.
20. E. Klijn and J. Koppenjan. Institutional design. *Public management review*, 8(1):141–160, 2006.
21. J. Koppenjan and J. Groenewegen. Institutional design for complex technological systems. *International Journal of Technology, Policy and Management*, 5(3):240–257–, 2005.
22. S. Manson. Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico. *Agriculture, Ecosystems & Environment*, 111(1-4):47–62, 2005.
23. I. Nikolic. *Co-Evolutionary Method For Modelling Large Scale Socio-Technical Systems Evolution*. PhD thesis, 2009.
24. E. Ostrom. Institutional rational choice: An assessment of the institutional analysis and development framework. *Theories of the policy process*, pages 35–71, 1999.
25. E. Ostrom, R. Gardner, and J. Walker. *Rules, games, and common-pool resources.* Univ of Michigan Pr, 1994.
26. L. Penserini, F. Dignum, V. Dignum, H. Aldewereld, and D. Grossi. Evaluating organizational configurations. In *Web Intelligence and Intelligent Agent Technologies*, volume 2, pages 153–160, 2009.
27. A. Rao and M. Georgeff. BDI agents: From theory to practice. In *Proceedings of the first international conference on multi-agent systems (ICMAS-95)*, pages 312–319. San Francisco, 1995.
28. M. Rudd. An institutional framework for designing and monitoring ecosystem-based fisheries management policy experiments. *Ecological Economics*, 48(1):109–124, 2004.
29. A. Thesen, L. Travis, and R. Gordon. *Simulation for decision making.* West Publishing Co. St. Paul, MN, USA, 1992.
30. M. Wooldridge, N. Jennings, and D. Kinny. The gaia methodology for agent-oriented analysis and design. *Autonomous Agents and Multi-Agent Systems*, 3(3):285–312, 2000.

A Case Study in Model Selection for Policy Engineering: Simulating Maritime Customs

Hassan Harb[1], F. Jordan Srour[2], and Neil Yorke-Smith[2,3]

[1] Faculty of Engineering, American University of Beirut, Lebanon.
hnh10@aub.edu.lb, srourf@hotmail.com
[2] Olayan School of Business, American University of Beirut, Lebanon.
[3] SRI International, Menlo Park, CA, USA. nysmith@aub.edu.lb

Abstract. The progress of containers through customs is more often an exercise in negotiation rather than a structured queuing process. As soon as a regulatory process involves negotiation, corruption becomes a factor. Studies by the Organization for Economic Co-operation and Development and other organizations reveal that customs corruption is not easily combated by policy changes. We analyze the appropriateness of agent-based modelling techniques for the simulation of potential anti-corruption policies in the maritime customs context. In so doing, we outline an initial simulation design, calibrated on evidence from the Port of Beirut, with the goal of estimating the costs and benefits of various reform policies.

Keywords: model selection, process modelling, maritime customs, corruption, multiagent-based simulation

1 Introduction and Motivation

Container shipments, according to the World Shipping Council, account for 60 percent of international sea-based trade by value. Competitive advantage is gained by properly managing and optimizing container flows through ports. The inspection of container contents and application of regulations and tariffs is a significant part of the import-export process. We study deviations from published maritime customs processes with the goal of using simulation as a tool in policy engineering. The domain is important not only because of the scale of maritime shipments worldwide, but also because of the deleterious impact of corruption, especially on the disenfranchised [47]. Given that corruption can enter the process whenever there is opportunity for human actors to negotiate [46], what simulation paradigm and techniques can we leverage to assess the potential impact of reform policies that might be applied?

This modelling question is important because the choice and application of paradigm impacts the quality of the solution to the domain problem under study, the ease of solvability, and the scope and validity of insights that can be obtained. Given the socio-technical questions in our domain of interest, methodologies from two fields are pertinent: social sciences [19] and logistics [22]. Complex systems theory, agent-based modelling (ABM), and classical operations research (OR) are among the techniques applied to a range of problems in container logistics, port management, and policy analysis in our domain of interest (for ABM, e.g., [23, 31]). We are not aware of research that specifically studies the simulation of maritime customs processes in order to quantify the effect of reform policies.

The literature contains a number of surveys of methodologies (e.g., [33, 45, 13, 14, 18, 27]), which form the lens for this paper. We contribute to the methodological question by highlighting our experience with the selection of a modelling and simulation paradigm; we provide a data point for the discussion of best practices in fitting simulation techniques to the domain problems under study.

The methodological meta-approach we explore, derived from that of Terán [45], can be summarized as: (1) Identify the scenario/system to be simulated, and the goals of the simulation exercise; (2) Make an initial methodological choice; (3) Collect data necessary for model-building; (4) Review the model and language choices in light of the data; (5) Design and build simulation; (6) Run the simulation to examine potential policy decisions; analyze and interpret the results; (7) Collect data on the fit between the selected techniques and the problem under consideration, and validate the model and results; and (8) Apply the conclusions to policy issues in the scenario/system. As can be seen, there is an emphasis on examining earlier steps in light of later steps and reconsidering decisions based on the progress of the process. We expand on these steps through this paper, and report our completion of the first four steps.

In summary, this paper details a work-in-progress case study of simulation technique selection for modelling social complexity in the domain of maritime customs. We give evidence (1) to the applicability of a methodological approach that includes evaluation and reasoned selection of a modelling paradigm, and (2) to the applicability of agent-based simulation.

2 Background and Problem Analysis

Whenever a process has the opportunity or obligation for human actors to negotiate, the possibility of corruption arises. The World Bank defines corruption as "the misuse of public office for private gain" [46]. We distinguish between (1) *routine* corruption (e.g., bribes for normal or expedited completion of processes); (2) *fraudulent* corruption (e.g., tacit or explicit collusion to reduce fiscal obligations); and, the least common but important, (3) *criminal* corruption (e.g., bribes offered to permit a totally illegal, lucrative operation).

While a corrupt act may bring local gain for one actor, the negative repercussions of corruption hang upon institutions, societies, and nations. These include impact upon [25, 38, 47]: poverty, tax evasion, political stability, democracy and rule of law, national competitiveness, and (especially for customs) distortion of trade figures. Further, corruption reinforces disenfranchisement and hinders development, being "one of the most serious barriers to overcoming poverty" with a strong correlation between perceived corruption and national per capita in-

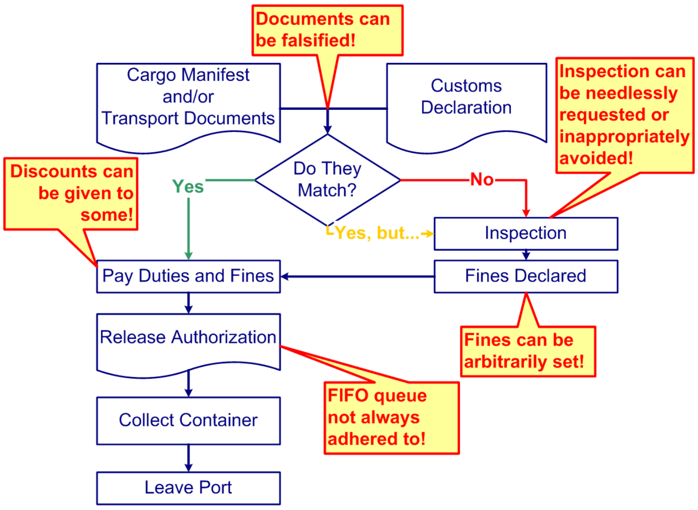


Fig. 1. An archetypal import process. Some opportunities for deviations from the published process are highlighted.

come [47]. However, as Langseth [28] points out, it would be "unrealistic and cost-prohibitive to attempt to eliminate corruption completely. ... Draconian anti-corruption programs, moreover, can have a negative effect on personal freedoms and fundamental human rights if regulations translate into excuses for public officials to become increasingly abusive toward the citizenry."

While eliminating corruption is not reasonable, reducing corruption is a common policy objective. Unfortunately, corruption is elusively difficult to fight. Whereas "strategies based on investigation and sanctions ... can be effective in regulating a situation of low corruption and preventing its further development", localized punitive or incentive-based policies "cannot correct a situation of widespread corruption" writes Hors [25]. The same report notes, based on lessons from three case studies, that "a re-engineering of procedures that leads to an important reduction of the opportunities of corruption should be at the core of the strategy." Complicating matters is the challenge of forecasting and comprehending the potential impact of policy changes. This recognized, even the most careful policy analysis and selection is not sufficient. Studies find that policy reform measures can only be successful if properly set within the national and institutional environment, driven by political will, sensitive to stakeholders at various levels, and part of a continuous improvement process [25, 2].

In order to counter established, widespread corrupt practices, a deeper understanding of the processes in which corruption features is required, together with a deeper understanding of the corrupt practices that occur, within the broad socio-political, socio-economic, governmental and cultural situation [40]. This need for understanding provides the motivation to our study. We posit that simulation can bring the measure of situated understanding sought.

Customs is defined as "the official department that administers and collects the duties levied by a government on imported goods" (Oxford English Dictionary [OED]). The process of moving a container through customs is primarily based on a match between shipping documents (e.g., bill-of-lading) and customs documents (e.g., manifest). If this match is made and the shipper/consignee are considered trustworthy, then the container may proceed following the payment of standard duties. If there is not a match, or should the container be randomly selected, then the container becomes subject to search and may see the leverage of additional duties or fines. The Organization for Economic Co-operation and Development (OECD) notes, particularly for developing countries, that customs revenue is a significant component of public finances, but that customs efficiency is often hampered by widespread corruption, creating "a major disincentive and obstacle to trade expansion" and leading to "disastrous consequences in terms of national security and public finance" [25].

Policy efforts led by the International Monetary Fund (IMF), OECD, World Customs Organization, and World Bank have focused on reducing trade barriers, reforming trade procedures, and building 'cultures of integrity'.

Fig. 1 shows some possible deviations from an archetypal customs import process. These include inaccurate, incomplete, or fictitious documentation; under- or over-inspection; inaccurate value estimation; waiving true fines or imposing additional fines; and delaying or expediting certain containers. In some situations, a whole grey 'parallel customs' system evolves (reported for Bolivia).

Negotiation is the most common entry point for non-standard behaviour within customs processes [26]. The study of negotiation is multi-faceted, including political science, economics, policy research, psychology, and computer science. Turan et al. [48] report on emerging efforts to unify research in behavioural and computational (including agent) communities. While the maritime customs domain holds a rich vein of research in the dynamics of bargaining situations, our objective is not to dwell deeply on the negotiation itself—framed as utilities, internal (affective) states, reasoning, and observable behaviour [48]—but rather to capture inter-actor negotiation within the customs process in order to study policy engineering questions.

3 Case Study: Selecting a Modelling Paradigm

We adopt a methodological meta-approach derived from the work of Terán [45], comprised in full of the ten steps explained below. Our simulation modelling case study stems from work in progress: we have completed the first four steps with some progress on the fifth step. Terán distinguishes four levels of language in Multiagent-Based Simulation (MABS): (1) cultural or natural language, (2) modelling and theoretical paradigm, (3) modelling language, and (4) simulation programming language. We incorporate these four levels of language as we move through the following steps in the selection of a modelling paradigm.

1. **Identify the scenario/system to be simulated, and the goals of the simulation exercise.** *Our aim is to study non-standard behaviours in maritime customs and the impact of policy reforms upon them.*
2. **Make an initial choice of modelling paradigm.** *We chose agent-based modelling for its promise in capturing the practices of actors that interact via negotiation* (modelling fit)*, a perception of* ease of implementation *in considering alternative policies, and insights on emergent behaviours (*explanatory power *for systems/process re-design). Moreover, we judged that MABS offers the potential to predict the impact of individual policy reform measures as well as to explore the effects of process re-engineering.*
3. **Collect data to fuel abstraction and model-building.** *We undertook a series of stakeholder interviews alongside a study of published processes.* Note that the choice of paradigm (Step 2) has some bearing on the type, volume, and quality of data needed to create the model. *Having chosen to model inter-actor negotiations richly and at the micro-level, we sought data sufficient to construct and validate an agent-based model.*
4. **Review data and re-evaluate model and language choices.** *While standard processes could be documented with some confidence, non-standard practices were related only anecdotally and from the literature. We re-considered the option of a lower-fidelity model and traditional Monte Carlo simulation.*
5. **Design and build simulation.** *We began by building an initial two agent system The full MABS will proceed once modelling decisions are validated.*
6. **Run simulation to examine potential policy decisions.**
7. **Analyze and interpret results.**
8. **Collect data on the fit** between the selected technique and the problem under consideration; possibly **revise the model, or even the methodological choice.**
9. Once the results have been validated and considered reliable, **apply the conclusions to policy issues in the scenario/system studied.**
10. **Seek to generalize conclusions to other problems or domains.**

The outcome sought from a rigorous process of selection of modelling paradigm and simulation technique is reliable and generalizable conclusions from the simulation. Two themes help achieve this outcome. The first theme is the ongoing re-examination of earlier steps in light of later steps, and reconsideration of decisions based on the progress of the process. The second them is the principled, multi-level validation of methodological choice, model abstraction, and results [44]. We now turn to a detailed description of our work as related to Steps 1–5.

3.1 Step 1: Identify Target System and Simulation Goals

Sect. 2 explained the domain of customs and the problems situated around corruption in customs. The goal of our simulation is analysis of policies designed to combat corruption. To this end, the areas where new policies may be applied (or old policies enforced) range broadly [46, 28, 26, 25]: (1) computerized data systems, (2) auditing, (3) sanctions, (4) role separation, (5) Customs Officer wages, (6) declaration and monitoring of assets for Customs Officers, (7) training, (8) culture of integrity (e.g., Code of Ethics), (9) legislative reforms, (10) legal reforms, (11) tax and tariff reform, (12) simplification of administrative procedures[4], (13) increased accountability and transparency (e.g., process documentation), (14) public awareness, (15) regular stakeholder meetings, (16) independent complaints authority, and (17) media freedom.

Principled means are required to evaluate and compare policy measures. From the literature (e.g., [29]), from reflection upon published measured data available, and from what interviewees said, we formulated the following metrics to assess evaluation of policy measures for import-export processes: (1) end-to-end time for an item to clear customs, (2) time deviation from desired date of receipt (usually, delay), (3) average tariff rate, (4) cost for an item to clear customs, including any corruption costs, e.g., bribes, that can be quantified, (5) percentage of items receiving electronic approval, (6) number of steps in the published process (a measure of transparency), (7) number of deviations from published process, (8) cost per deviation (are a lot of little deviations as bad as one big deviation?), (9) percentage of customs revenue diverted, i.e., lost to the government, (10) cost of enforcement, and (11) amount of change in a re-engineered process compared to a current process.

3.2 Step 2: Choose Initial Modelling Paradigm

In the second step of our methodological meta-approach, we selected an initial modelling paradigm. To assess the 'quality of fit' between a selected modelling or simulation technique and the problem domain (e.g., customs), problem instance (e.g., corruption), and study questions (e.g., competence of a variety of policies in mitigating corruption), we designated the following set of criteria, which provide a checklist for the choice of methodology. They are: (1) **Modelling fit**: how well does the modelling paradigm suit the (abstracted) system to be simulated? (2) **Cognitive fit**: how well does the modelling/theoretical paradigm suit the thinking of the modeller? (3) **Explanatory power**: how well can the simulation developed answer the study questions? (4) **Ease of implementation**: how well does the implementation language suit the model to be implemented and the questions to be asked? (5) **Computational tractability**: how readily can the simulation be performed?

Our initial methodological choice was influenced by the idea that corruption is a phenomenon that emerges. This influence comes from the recognition that regulations or policies are rarely established with the intent of *encouraging* corruption. On the contrary, published customs regulations, in Lebanon and elsewhere, are designed to regulate the flow of legal goods while capturing government revenue from duties. Nevertheless, corruption patently exists and is endemic in many locations. As such, we sought a modelling paradigm that could exhibit emergent behaviour, driving us quickly to agent-based models (ABMs).

[4] Significant since "systems and procedures [evolve] to maximise the number of steps and approvals—to create as many opportunities as possible for negotiation" [25].

The advantages of agent-based models are argued to include [3, 19]: (1) "descriptive realism ... natural system boundaries" [16] (*modelling fit* between the system studied and modelling paradigm and modelling language); (2) flexibility, ease of modelling (*cognitive fit* between natural language, and modelling paradigm and modelling language); (3) heterogeneity and adaptive behaviour at the micro level); (4) emergent behaviour at the macro level; (5) scalable/parallel computation; (6) some accessible tools (but see [20]), i.e., ease of implementation; (7) explanatory insights, especially into non-equilibra behaviour, social or spatial networks, and analytically intractable systems [3]; and (8) visual and intuitive nature for interpretation and public dissemination.

On the other hand, the disadvantages of ABMs are commonly recognized as [5, 30]: (1) interpretation of the simulation dynamics (ABM are opaque: in some ways explanatory power is limited), (2) replication of results, (3) generalization of the results (including robustness of results), (4) validation of the implementation from bugs, and (5) extraction of an analytic model, if relevant.

Leombruni [30] is among those who argue that these disadvantages can be overcome. Agreeing, Hamill [20], an experienced policy adviser, nonetheless adds "To persuade policy advisers to adopt [ABM] there needs to be a clear benefit in terms of the output." Although the case for the value of agent-based simulation in policy analysis is made well by Dignum et al. [14, 15], Hamill finds "The policy areas and questions that would benefit from ABM need to be identified."

A methodology for multiagent-based simulation consists of seven steps [6, 16, 12, 13] (for policy analysis, see also [27, 15]), which might be argued to hold beyond MABS: Abstraction, Design, Inference, Analysis, Interpretation, Application, Conclusion. The meta-approach we follow, outlined earlier, thus wraps very similar steps with methodological selection and re-evaluation.

The problem studied, while embedded into a social context and highly influenced by organizational, cultural, and social factors, does not fit exactly into any of the paradigmatic models for agent-based social simulation identified by Marietto et al. [33]; the closest match is *socio-concrete models*. Rather, simulating negotiation in maritime customs may be better characterized not as a social simulation per se, but as simulating social complexity—the structure and norms of what is and is not considered acceptable in the realm of customs processes, and the micro-macro link between (emergent) actor behaviours and policies applied to the system.[5] Our objective is not so much forecasting (as in economics) or optimization (as in traditional OR) but understanding of collective behaviours. In purpose, our simulation sits between *Mediative* and *Generative* [1].

Our reasons were as follows, for selecting ABM and MABS as a methodological choice to investigate customs processes with negotiation between possibly corrupt actors. Alternatives considered were simulation based on dynamic sys-

[5] In the taxonomy of Davidsson et al. [13], the domain is 'social system and organizations'; the end-user is 'scientists' and, perhaps, 'policy makers'; the purpose is 'prediction' and 'analysis'; the simulated entity is 'living'; the number of agent types is a small finite number; the structure is peer-to-peer, hierarchal; agents communicate; the input data is mostly artificial; the present maturity is 'conceptual proposal'.

tems, classical OR techniques, and statistical aggregate analysis. In the interest of forthrightness, however, we acknowledge our predisposition to MABS.

First, the naturalness of modelling inter-agent communication. We are modelling human actors negotiating, usually at a peer-to-peer level, for which ABM is well-suited. Second, a perception of ease of implementation: modelling and implementation environments are readily available. Third, sought explanatory power for systems/process re-design. The documented success stories of MABS speak of its efficacy [19]. Fourth, the established track-record of MABS in diverse domains coupled with the weakness of alternative methodologies, especially when studying complex, value-driven, human socio-technical problems [14, 27].

An addition factor in our decision was that agent-based models have been successful in port management and container shipment (e.g., [31, 22]), and agent-based simulation has been successful in port stakeholder analysis (e.g., [23]) and policy analysis in transport (e.g., [9]).

Agents have also been used to study corruption, as we now briefly survey. Hammond [21] develops an agent-based population model in an effort to explain shifts in corruption levels. Corruption is modelled as a simple, game-theoretic repeated interaction on the micro level. In a tax-evasion domain, endogenous shifts in global corruption levels are observed as emerging from the micro-behaviour.

Situngkir [42] is interested in the link between corrupt behaviours in individual agents and the societal/cultural environment in which they interact. He builds a MABS inspired by corrupt bureaucrats in Indonesia and obtains system-wide results. However, these results require validation of the assumptions made.

Deviations from customs processes may be seen as governed by structural and normative aspects of the society. Savarimuthu et al. [39] examine how an agent may infer the norms of a society without the norm being explicitly given. Looking more generally at MABS, Norling et al. [35] seek to add more ‘human-like’ decision making strategies, drawing on studies in naturalistic decision making. Dignum et al. [15] emphasize models that include culture, to capture societal aspects such as social norms; these are relevant to studying corruption. Gal et al. [17] demonstrate empirically that people in the US and Lebanon behave differently in negotiation with automated agents in a repeated game; they attribute the differences to cultural factors such as collectivism.

3.3 Step 3: Data Collection

We chose to study the Port of Beirut, Lebanon, due to its regional prominence and its proximity to our institution. The port handles some 900,000 TEU container units per year (out of 500M worldwide), with annual revenues in excess of $150M. It processes 80% of imports into Lebanon, as well being a significant transshipment point for Syria and beyond [7]. Further, anecdotal and published reports indicate a number of exotic practices [36, 32, 11, 37], which may likely be considered to be a superset of practices elsewhere. As a country, Lebanon is in the bottom third of the Transparency Index [47], with a score of 2.5/10.

The objectives of the initial data gathering phase were to characterize the domain and the processes of interest, and to elicit structural, environmental, institutional, and behaviourial knowledge necessary to build a MABS. Further, the data and its interpretation informs the next step in the methodological process, namely the re-evaluation of the selected models and techniques.

Three sources of information provide the basis for abstraction and modelling in MABS [16]: (1) observation and data collection from the target system, (2) bibliographical review (i.e., theories), and (3) domain experts.

Target observation. Regarding the first source, the only data obtained directly from the target system is published statistics that are available for import-export figures in various jurisdictions. It comes as no surprise that we have, to date, been unable to collect empirical data by observation (e.g., sampling containers and following their progresses through the process); in view of the sensitivity of questions in the domain, this kind of empirical study is unlikely.

Bibliographical review. The bibliographical review yielded more data. While jurisdictions differ in their regulations and procedures, nearly all ports have similar import-export processes [34]. Fundamentally, the processes depend on a match of paperwork between manifest and declaration. Nearly all ports have an IT system of some sort, which includes a maritime container standard [4]. The widest differences between systems are seen in taxation schemes.

We examined documented processes at the Port of Beirut, the Port of New York/New Jersey, and the Port of Rotterdam. For reasons of space, and due to the broad similarities, we describe only the Port of Beirut; for the others, see [4].

Fig. 2 summarizes the administrative hierarchy at the Port of Beirut, and Fig. 3 depicts the import process. Initiated in 1993 with its first release in 1998, *NAJM* is the Lebanese automated customs clearance IT system. *NOOR* is the online portal to NAJM. Over half of all declarations are electronic, and automatically verified containers, known as 'green line declarations', account for the majority of declarations (in the reported statistics). We note that one result of these kind of IT systems, as in ports worldwide, is a separation of roles.

A primary role in the container clearance process is that of *freight forwarder*—a company that manages and organizes shipments for others, sometimes consolidating smaller shipments. Data from freight forwarder companies (e.g., [10]), provided valuable context and insight into port procedures.

To give context to customs practices, we also examined published research on the broader socio-economic and cultural environment [24, 41, 32, 40]. Lebanon's culture is shaped by both Arab and European influences. Lebanese like Arab culture is collective (IDV 38) and has high power distance (PDI 80) and moderately-high uncertainty avoidance (UAI 68), but the Lebanese are famously entrepreneurial. Likely because of the 1975–90 civil war, the Lebanese distrust government and place a low value on obeying rules [36]; the economic and societal context is laissez-faire coupled with strong family and, for some, religious values.

Domain experts. In order to form a more robust picture of standard and non-standard practices at the port, we conducted semi-structured, exploratory interviews with customs brokers, freight forwarders, longshoremen (i.e., the workers

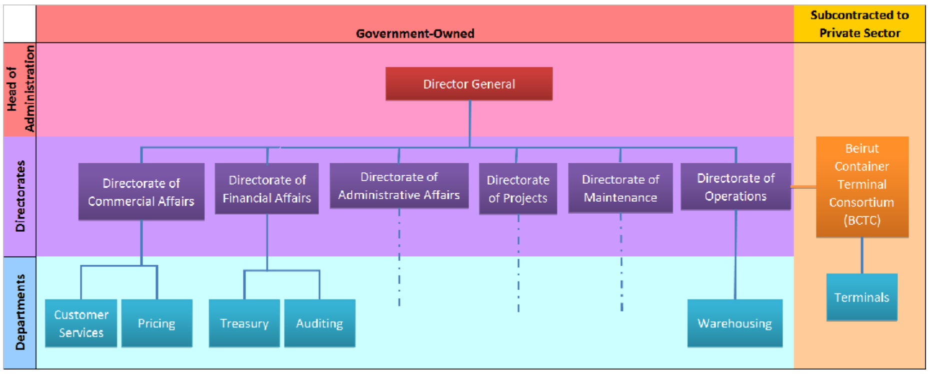


Fig. 2. Administrative hierarchy at the Port of Beirut.

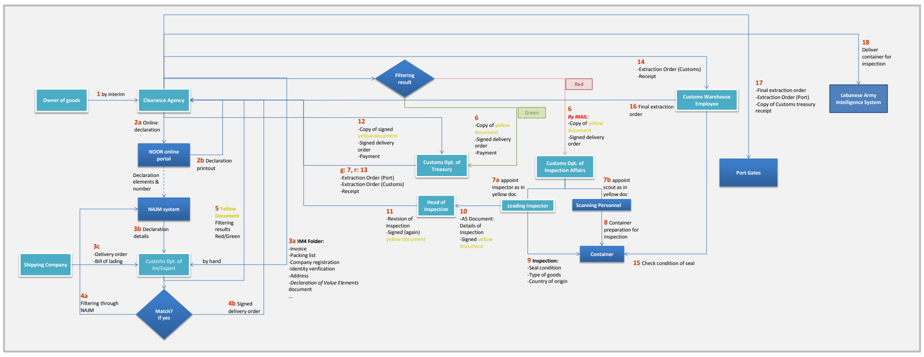


Fig. 3. Simplified customs import process at the Port of Beirut.

who physically move containers), and Customs Officers at the Port of Beirut. For reasons of privacy and given the sensitivity of the topic, we refrain from identifying the interviewees with the statements given.

The actors in import-export processes are listed in Table 1. Our interviewees, when feeling able to speak with some freedom, reported a systemic norm of deviations in import-export process. Customs Officers in practice have greater discretion than their job description states. It was considered routine to engage in 'wasta' [32]—exploitation of influence, political (or other) power, connections—or to offer a bribe or a 'baksheesh'—a widespread practice "(in parts of Asia) [of] a small sum of money given as alms, a tip, or a bribe" (OED).

Non-standard practices reported fall into three categories. First, deviations based on the relationship between actors, where these is no obvious monetary or physical bribe. Relationship levers in negotiation can arise from family connection (nepotism), political tie (patronage), or favour owed. Second, deviations be based on monetary considerations, where there is a tangible bribe, whether

Owner	Excise officer
Owner's agent	Head of Excise
Freight forwarder	Customs broker
Shipping company	Longshoremen
Vessel captain	Customs warehouse employees
Clearance Agency officer	Port security staff
Customs Agency officer	Recipient (consignee)
Inspection officer	Police officer
Head of Inspection	Customs Investigation and Audit officer

Table 1. Actors identified in maritime imports process at the Port of Beirut.

cash or gift, or a debt forgiven. Third, negotiation levers based on threats or extortion, whether physical, financial, or reputation-based.

Our interviewees noted that these deviations from stated practice stem from three sources: discretionary interface between actors, networks of accomplices, and lack of efficient controls. One story is illustrative: a container of household goods was expedited on the basis of an iPad (new and difficult to obtain in Lebanon at the time) that was gifted to a freight forwarder, who in turn used his personal relationships to garner the "good will" of the Customs Officers. Neither party saw this behaviour as non-standard because it fell outside the domain of the IT system and within the domain of standard social/cultural behaviour.

At the instigation of the IMF and other international organizations, Lebanon restructured its customs law and tariffs in 2000 [7, 29]. Despite the IT systems, the legal restructuring, and political will (or, at least, words) [37], our fieldwork correlates with reported statistics that corruption is endemic in Lebanon (e.g., [36, 32, 47]), at least in maritime imports. This only underscores the challenge of determining effective policy measures to fight corruption.

3.4 Step 4: Re-evaluate Model and Language Choices

We reconsidered the suitability of ABM by returning to the quality of fit metrics introduced in Sect. 3.2. To mitigate anchoring bias from our initial choices, we sought the input of experts outside the MABS community. Specifically, we presented the preliminary conceptual design of our MABS at the 2010 Annual Meeting of the Institute for Operations Research and the Management Sciences (INFORMS). The feedback obtained provided us with a new perspectives on alternative modelling paradigms for the domain problem under study.

Among traditional OR techniques is Monte Carlo simulation. In such an approach, the agent's reasoning is modelled as a form of stochastic process. The probability distribution over possible actions the agent could take is estimated. For example, at a given opportunity to offer a bribe, does the freight forwarder offer money or not? Indeed, as Axtell [3] points out, when the system being simulated is stochastic, with the behavioural equations known, then MABS can be seen as a type of Monte Carlo simulation. Although accurate empirical data is unlikely, given the problem being studied, our fieldwork encourages us that

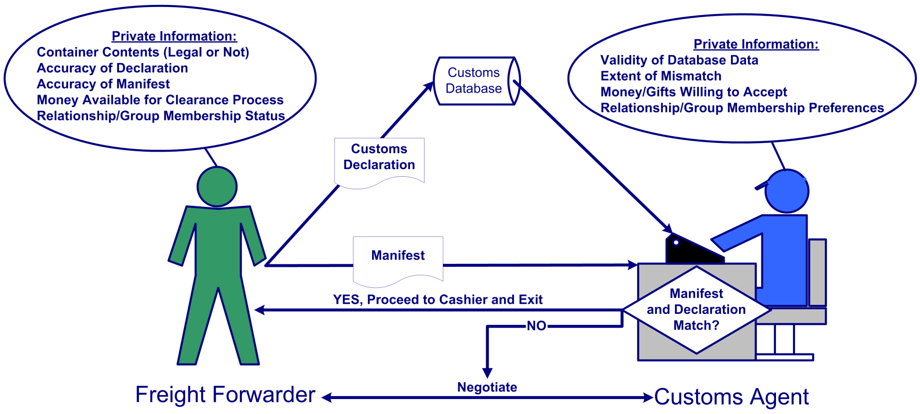


Fig. 4. Private knowledge and opportunity for negotiation.

sufficient, reliable information can be obtained to design and validate a MABS [1]. Further, we believe that ABM allows high-fidelity modelling of inter-actor negotiations, and allows agent behaviours that are heterogenous and adaptive. Hence, only if insufficient data is available to construct and validate an ABM, then a lower-fidelity model and traditional Monte Carlo simulation may fit better. Second, we had concern that a methodology based on stochastic processes risks being too simplistic, since we are trying to capture complex processes and, ultimately, adaptive human negotiations [14]. Third, we hold that the behaviours exposed in our field studies are naturally modelled as entity-level interactions. Fourth, traditional OR techniques are found to be inadequate for complex, value-driven, socio-technical problems [27]. Finally, we hold that MABS can bring insight into processes as well as providing values for global metrics (e.g., mean end-to-end clearance time). Specifically, it allows us to see the change in behaviour(s) at the micro, agent level.

This analysis convinced us to continue with MABS given its capabilities for micro- (agent behaviour) and macro- (emergent phenomenon) level analysis.

3.5 Step 5: Design and Build Simulation

The heart of our MABS is the actors' 'walk' through the documented processes for each shipment, the points of possible deviation, the decisions whether to engage in (or how to respond to) non-standard practices, and the negotiation that may ensue. Fig. 4 depicts how negotiation can enter the procedural interaction between Freight Forwarder and Customs Agent actors.

A proof-of-concept implementation of simulation was undertaken using the Java-based JADE agent toolkit [8]. We chose JADE rather than a dedicated MABS environment (e.g., NetLogo, RePast) because, first, as remarked earlier, ours is not a social simulation per se, but a model of negotiating human agents within simulated processes. Second, JADE comes with a mature set of development tools with which we possessed familiarity.

The scope of the system simulated is deliberately restricted to the customs process. To begin to answer the questions of interest, we do not need to simulate a larger cross-section of society, although it is important to adequately reflect the influence of the broad socio-cultural environment upon the actor's micro-behaviour in negotiation [15]. We also take the approach that we do not need to model in detail each actor's internal cognition and mental state (e.g., beliefs, desires, norms, goals)—just sufficiently to capture negotiation practices.

We encoded negotiation scripts that follow the interactions in the published process (of which Fig. 3 is a simplification), supplemented with branching points that match points where deviations from the published process may occur.

We take each shipment as analogous to a round in a sequential bargaining game. Each shipment brings the opportunity for actors to negotiate, or not. As a means to test our our selected modelling paradigm—agent-based simulation—we initially focused on the actors whose behaviours are at the heart of the process: customs agents and freight forwarders. This preliminary test of interactions allows us to evaluate our selected model and subsequently to scale it to include container owners and longshoremen.

Our initial model of negotiation is between pairs of agents, using a game-theoretic approach with negotiation options described by truth tables. Behaviour is not adaptive, i.e., the agents do not learn from previous 'rounds' nor adapt to others' behaviours. As we confirm the validity of an agent-based model, develop the negotiation behaviours, and add additional port actors to the model, it is likely that we need more robust agent reasoning and negotiation strategies. To this end, Sycara and Dai survey methods for agent reasoning in negotiation [43]. These include, for instance, analytic, defeasible logic-based, argumentation-based, and production rule systems; see also the references in Gal et al. [17] for a human–agent perspective.

Although the full MABS can be built—extending the current approach and making the behaviours adaptive—we prefer to gather more data in order to more realistically capture the parameters of negotiation. For example, supposing a Customs Officer is offered a bribe, we can encode the estimated probability that the Officer accepts the bribe. We rather develop the reasoning over factors that the Officer weighs up, such as the value of container, monetary equivalent value of bribe, relative value of bribe compared to salary, perceived risk of detection, relationship factors, and so forth.

Discussion points in MABS design. We conclude this section by mentioning some points of discussion. First, in capturing agent negotiation patterns: which actors negotiate with which others, especially outside of documented process interactions? Which decision points (corruption opportunities) to model? Second, what is the stopping criteria in negotiation: should it be based on time? number of iterations? or some value? Third, tracking and quantifying non-monetary exchanges: how does one compare "helping somebody's child enter an elite private school" with "handing over €100" or "gifting an iPad"? Fourth, how can we succinctly model relational capital (wasta)? Fifth, how can we model and/or quantify threats? Sixth, how rich a model of emotive state is appropriate for the agents, and how rich a model of the broader environment (i.e., outwith the customs process and the actors associated with it) is pertinent to our measures of policy effectiveness [40, 15, 17]?

4 Conclusion

This paper reported a work-in-progress case study of simulating social complexity in the domain of maritime customs. In this domain, understanding which reform policies are effective against non-standard practices is challenging. We gave evidence to the applicability of a methodological approach that includes evaluation and selection of modelling paradigm, and to the applicability of agent-based simulation. We reported data gathering and initial model building, to lay the foundation to understanding in a quantitative way the costs and benefits of various reform policies aimed at customs processes.

We concur with earlier authors on the value of MABS in public policy, agreeing with Hamill [20] on the need for tools, documentation of best practice, and an outcome-based argument for agent-based modelling in policy contexts. Our exploration of ABM for the maritime customs domain lends support for the ABM methodology. We also agree with Davidsson [13] and Arroyo et al. [1] on the need for validation of MABS that simulate complex human behaviours.

Our immediate next steps are to implement the remainder of the designed agent model, increasing the number of agent roles, and to begin analyzing reform policies in terms of the global metrics identified. We require further fieldwork at the Port of Beirut in order to calibrate modelling parameters, add fidelity to the ABM, and, importantly, to understand the space of more substantive process re-engineering options. Further ahead, we would like to include export processes within our study, since containers imported into one port have been exported from another, and to contrast behaviours within the Lebanese context with other contexts [36]. In particular, we ask whether MABS can give insight into the relative effectiveness of policies in different socio-political, socio-economic, governmental, and cultural contexts.

Acknowledgments. We thank the participants of the session on Agent-based Decision Support in Auctions and Negotiations at the INFORMS Annual Meeting 2010 for discussions. We thank the AMPLE reviewers for their suggestions. We thank J. Hess and C. M. Karam, and the anonymous participants of our field studies. The authors were partially supported by University Research Board grant A88813 from the American University of Beirut.

References

1. Arroyo, J., Hassan, S., Gutiérrez, C., Pavón, J.: Re-thinking simulation: a methodological approach for the application of data mining in agent-based modelling. Computational and Mathematical Organization Theory 16(4), 416–435 (2010)

2. Gonzalez de Asis, M.: Reducing corruption at the local level. World Bank (2000)
3. Axtell, R.: Why agents? On the varied motivations for agent computing in the social sciences. Working Paper 17, Brookings Institution (2000)
4. van Baalen, P., Zuidwijk, R., van Nunen, J.: Port inter-organizational information systems: Capabilities to service global supply chains. Foundations and Trends in Technology, Information and Operations Management 2(2–3), 81–241 (2008)
5. Bankes, S.C.: Agent-based modeling: A revolution? Proc. National Academy of Sciences 99(Suppl 3), 7199–7200 (2002)
6. Barber, K.S., Liu, T.H., Han, D.C.: Agent-oriented design. In: Multi-Agent System Engineering. vol. LNCS 1647, pp. 28–40. Springer (1999)
7. Beirut Container Terminal Consortium: Port of Beirut. `portdebeyrouth.com`
8. Bellifemine, F.L., Caire, G., Greenwood, D.: Developing Multi-Agent Systems with JADE. Wiley, Chichester, UK (2007)
9. Bergkvist, M., Davidsson, P., Persson, J.A., Ramstedt, L.: A hybrid micro-simulator for determining the effects of governmental control policies on transport chains. In: Proc. MABS'04. pp. 236–247 (2004)
10. Cargomaster Group of Companies: Company profile. `groupage.com` (2006)
11. Second food scandal in a week as fake products found. Daily Star (25 Sep 2010)
12. Davidsson, P.: Agent based social simulation: A computer science view. J. Artificial Societies and Social Simulation 5(1:7) (2002)
13. Davidsson, P., Holmgren, J., Kyhlbäck, H., Mengistu, D., Persson, M.: Applications of agent based simulation. In: Proc. MABS'06. pp. 15–27 (2006)
14. Dignum, F., Dignum, V., Jonker, C.M.: Towards agents for policy making. In: Proc. MABS'08. pp. 141–153 (2008)
15. Dignum, V., Dignum, F., Osinga, S.A., Hofstede, G.J.: Normative, cultural and cognitive aspects of modelling policies. In: Proc. 2010 Winter Simulation Conference. pp. 720–732 (2010)
16. Edmonds, B.: The use of models: Making MABS more informative. In: Proc. MABS'00. pp. 269–282 (2001)
17. Gal, Y., Kraus, S., Gelfand, M., Khashan, H., Salmon, E.: An adaptive agent for negotiating with people in different cultures. ACM Transactions on Intelligent Systems and Technology (to appear)
18. Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., del Olmo, R., López-Paredes, A., Edmonds, B.: Errors and artefacts in agent-based modelling. J. Artificial Societies and Social Simulation 12(1:1) (2009)
19. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist. Open University Press, Buckingham, UK, 2nd edn. (2005)
20. Hamill, L.: Agent-based modelling: The next 15 years. J. Artificial Societies and Social Simulation 13(4:7) (2010)
21. Hammond, R.: Endogenous transition dynamics in corruption: An agent-based computer model. Working Paper 19, Brookings Institution (2000)
22. Henesey, L., Davidsson, P., Persson, J.A.: Agent based simulation architecture for evaluating operational policies in transshipping containers. J. Autonomous Agents and Multi-Agent Systems 18(2), 220–238 (2009)
23. Henesey, L.E., Notteboom, T.E., Davidsson, P.: Agent-based simulation of stakeholders relations: An approach to sustainable port and terminal management. In: Proc. Intl. Association of Maritime Economists Annual Conf. pp. 314–331 (2003)
24. Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organizations: Software of the Mind. McGraw-Hill, New York, NY, 3rd edn. (2010)
25. Hors, I.: Fighting corruption in customs administration: What can we learn from recent experiences? OECD Development Centre Working Paper 175 (2001)

26. Klitgaard, R., MacLean-Abaroa, R., Parris, H.L.: Corrupt Cities: A Practical Guide to Cure and Prevention. ICS Press, Oakland, CA (2000)
27. Kunsch, P., Kavathatzopoulos, I., Rauschmayer, F.: Modelling complex ethical decision problems with operations research. Omega 37(6), 1100–1108 (2009)
28. Langseth, P.: Prevention: An effective tool to reduce corruption. In: Proc. ISPAC Conf. on Responding to the Challenge of Corruption (1999)
29. Lebanese Customs Authority: Lebanese customs. `www.customs.gov.lb/customs`
30. Leombruni, R., Richiardi, M.: Why are economists sceptical about agent-based simulations? Physica A 355, 103–109 (2005)
31. Lokuge, P., Alahakoon, D.: Improving the adaptability in automated vessel scheduling in container ports using intelligent software agents. European J. Operational Research 177(3), 1985–2015 (2007)
32. Makhoul, J., Harrison, L.: Intercessory wasta and village development in Lebanon. Arab Studies Quarterly 26(3), 25–41 (2004)
33. Marietto, M.B., David, N., Sichman, J.S., Coelho, H.: A classification of paradigmatic models for agent-based social simulation. In: Proc. MABS'03. pp. 193–208 (2003)
34. Muller, G.: Intermodal Freight Transportation. Eno Transportation Foundation and Intermodal Association of North America, Washington, DC, 4th edn. (1999)
35. Norling, E., Sonenberg, L., Rönnquist, R.: Enhancing multi-agent based simulation with human-like decision making strategies. In: Proc. MABS'00. pp. 349–357 (2001)
36. Rawwas, M.Y.A., Patzer, G.L., Vitell, S.J.: A cross-cultural investigation of the ethical values of consumers: The potential effect of war and civil disruption. J. Business Ethics 17, 435–448 (1998)
37. Rizk, C.: Hariri launches anti-corruption, public services initiative: Premier calls for modernization of bureaucratic processes. The Daily Star (30 Jul 2010)
38. Rose-Ackerman, S.: Corruption and government. International Peacekeeping 15(3), 328–343 (2008)
39. Savarimuthu, B.T.R., Cranefield, S., Purvis, M.A., Purvis, M.K.: Obligation norm identification in agent societies. J. Artificial Societies and Social Simulation 13(4:3) (2010)
40. Seleim, A., Bontis, N.: The relationship between culture and corruption: A cross-national study. J. Intellectual Capital 10(1), 165–184 (2009)
41. Sidani, Y., Gardner, W.L.: Work values in the Arab culture: The case of Lebanese workers. J. Social Psychology 140(5), 597–607 (2000)
42. Situngkir, H.: Money-Scape: A generic agent-based model of corruption. Computational Economics Archive 0405008, EconWP (2004)
43. Sycara, K., Dai, T.: Agent reasoning in negotiation. In: The Handbook of Group Decision and Negotiation. Springer, New York, NY (to appear)
44. Takadama, K., Kawai, T., Koyama, Y.: Micro- and macro-level validation in agent-based simulation. J. Artificial Societies and Social Simulation 11(2:9) (2008)
45. Terán, O.: Understanding MABS and social simulation: Switching between languages in a hierarchy of levels. J. Artificial Societies and Social Simulation 7(4:5) (2004)
46. The World Bank: Helping countries combat corruption: The role of the World Bank, poverty reduction and economic management network. Tech. rep. (1997)
47. Transparency International: The global corruption report 2009. Tech. rep. (2009), `www.transparency.org/publications/gcr`
48. Turan, N., Dai, T., Sycara, K., Weingart, L.: Toward a unified negotiation framework: Leveraging strengths in behavioral and computational communities. In: Proc. IJCAI'99 Workshop on Modeling Intercultural Collab. and Negotiation (2009)

Modelling Culture in Multi-agent Organizations

Alexis Morris, William Ross, Mihaela Ulieru

Adaptive Risk Management Lab
Faculty of Computer Science
The University of New Brunswick
{alexis.morris, william.ross, ulieru}@unb.ca,

Abstract. Culture is a key determinant of relationships and organization formation; however, its role, key properties, and mechanisms are not yet fully understood. This work explores culture and cultural modelling from a complex systems and multi-agency standpoint that takes into account the multi-dimensionality of culture. The need for performing such modelling and simulation is evident since in-vivo organizational experiments are costly, not easily generalizable, and may not be feasable in critical situations. This work contributes to agent modelling of organizations by i) developing a unique approach to culture modelling from a holistic and systems-theoretic perspective according to seven dimensions, and ii) simulating cultural interactions as a multi-agent system of high functioning agents that achieve an equilibrium of beliefs. Experiments present an early model of an agent organization, having distinct roles and influences. As new individuals are added to the system emergent culture develops, with resilient properties.

1 Introduction: Modelling Organizational Cultures

"There must be mechanisms in societies which permit the maintenance of stability in culture patterns across many generations." — Hofsteder, 2001

Cultures develop through complex interactions between parts of an organization, its actors, environment, technologies, etc, [2], (ch. 6). These diversify organizations from each other in important and unique ways that can be seen as compatible, complementary, or even conflicting. This is seen when different cultures are present in a single institution, or when personal values or standards of behaviour are in conflict with those of the organization to which they belong. In such cases there are competing cultures influencing decisions and actions of individuals which cause cognitive dissonance and stress, [13], over which behaviour is appropriate, and hence which belief (and culture) is stronger. As a concept, culture is difficult to classify and model due to inherent imprecision in defining and isolating the components of culture as it is a fuzzy concept with many possible realizations, i.e., in individual and group beliefs, in the physical reality of actions and environment, in the established conceptual ideas about it, and also in long-held traditions. Culture is challenging to understand but plays a key role, as a determinant of relationships among individuals in organizations and as a macro-level driver of individual actions (see [9], (ch. 8), for more on culture as it relates to organizations). Cultural modelling allows for studying the effect and influence of culture, and predicting how the type of culture at work will affect the ability of the organization to function and progress. This modelling is relevant in policy-making, among other domains, as it gives stakeholders a way to visualize and discuss cultural effects in different organizational scenarios.

This paper targets organizational culture modelling, and further presents our recent work, [12], on clarifying cultural relationships, and how "collective programming of individuals," [9], takes place. Culture is defined, and an early exploration of the emergence and evolution of culture in organizational contexts is shown. This is an early step towards future studies about the interplay and eventual integration of two or more different cultures in a shared system environment. The perspective is that culture is not only an intangible social construct, but also an emergent property, and the primary theme is that in order to understand, discuss, and measure culture it must be recognized as a complex, multi-dimensional, and multi-agent system.

Contributions of this work are two-fold: i) it adds to the literature of culture as a complex system by presenting a new seven-dimensional model to describe and discuss culture, and ii) it models cultural interactions as a multi-agent system, of high functioning agents, that achieves equilibrium in beliefs. Section 2 highlights some related work in the area of culture modelling. Section 3 presents a working definition of culture. Section 4 describes the notions behind a complex system and makes the case for culture as such a system. Section 5 discusses a new model for culture in seven dimensions. Section 6 describes the approach to measure culture with high-functioning agents. Section 7 describes three experiments to show the emergence and evolution of culture. Section 8 concludes the paper.

2 Literature Overview

Literature pertaining to culture modelling is vast and interdisciplinary, however this work relates to four perspectives, i) those that use agent-based interaction models, ii) those that use norm-governed models, iii) those that use mathematical models, and iv) those that use a multi-dimensional approach to describe culture in organizations. In terms of agent-based interaction models, MASQ, [17], frames the culture problem with a framework based on quadrants determined from two overlapping dimensions, the individual-collective and the internal-external spectrums. The I-I quadrant refers to the individual, the I-C to a group, the E-I for the physical reality of an individual, and the E-C for the physical reality in a group. In terms of culture the authors define it as strictly internal knowledge, patterns, and rules in the I-C quadrant; this is a similar vision as it advocates culture as shared beliefs, but does not target the emergence and influence of culture. MOISE+/Brahms, [15], is another approach using agents to model organizations based on the structure, work processes (roles), and normative aspects. The aim is toward organization-aware simulation, and although culture is mentioned briefly as tradition, an emergent property of norms, it is not addressed specifically.

In terms of norm-governed models, PreSAGE, [6], presents a rule-based mechanism to develop agent systems based on peer-pressure through reputation, reinforcement learning, and voting strategies. This approach is similar to the vision of understanding cultural influence, but does not discuss culture, or use belief frameworks. Also, in [1], ad-hoc networks are used for resource sharing based on event calculus, rules, and graphs. It is similar in studying permission, obligation, and institutional power of certain agents. Lastly, mathematical techniques of wavelet transform have been used, [8], to model ethnic violence due to poorly defined boundaries and being well-mixed or well-separated. This approach highlights the impact of physical factors and emergence, but does not present a detailed model of culture.

The multi-dimensional perspective of culture is not new, and in organizational sciences there are many models identified (see [2]). Payne, (ch. 10), presents a model of culture based on three dimensions. Ashkanasy et al., (ch. 8), promotes a model with ten dimensions. Dickson et al., (ch. 28), presents a nine dimensional model. Moreover, in [9], Hofstede, (ch. 25), promotes a five and six dimensional model for nations and governments, respectively. See [12], for more on these organizational science dimensions.

3 A Working Definition of Culture

Culture is a "set of shared attitudes, values, goals, and practices that [both] characterizes an institution, organization, or group", as well as emerges from and sets the behaviour of a group, [10]. Culture is also considered as a system, "an entity standing in a state of equilibrium within a specific environment," [19]. These perspectives underscore a holistic view of culture as both a bottom-up/emergent property that achieves a steady state (stable behavioural pattern) and is a top-down influencer of behaviour. The bottom-up view results from individual behavioural interactions, shared beliefs, and learning-by-observation from actors in an organization. The top-down view of culture as an influencer highlights its feedback effect on individuals within the system, as established collective beliefs in the past effect personal behavioural interactions of the present.

Along this line a working definition is the holistic interaction among n agents across seven distinct dimensions that cause stabilization of beliefs within these agents over time. This definition is useful as it targets interaction at the level of individuals, captures the notion of shared beliefs over time, and highlights the need for a multidimensional perspective of culture (in this work the physical, individual, functional, social, structural, normative, and information dimensions are defined). The focus on shared beliefs as a determinant of action is a central concept since beliefs provide an understanding of motivations for behaviour and can be traced to internal and/or external sources, as messages passed between individuals. In this way the influence at both the individual and collective levels can be understood as beliefs. The definition can also be extended from a mono-cultural context to a multi-cultural one and shapes the fuzzy sociological notion of "culture" into a more approachable problem. Using multi-agent systems modelling and simulation it is possible to describe the individual, and also the different interaction configurations that can take place, to analyze the system at both a small or large scale.

4 Culture as a Complex System

Culture can be understood from the perspective of complex systems, that is, having both a micro and macro scope with unique micro-level interaction, and emerging macro-level patterns; these are further situated in a dynamic environment. For organizations this micro-level is the level of individuals and the macro-level is the level of the whole organization as a unit. As a result work on culture requires a holistic method that encompasses system behaviours and structures at both levels of granularity. In addition to these two levels of detail it is important to highlight the openness factor of organizations, since individuals may be continually added or removed from the environment domain. This macro and microscopic focus, combined with the allowance of an open systems perspective, presents a culture as "emerging" from interactions of individuals (bottom-up emergence) yet having reinforcing feedback influence, [16], on these same individuals (adaptation to top-down forces). Additionally perturbations are considered as new elements are encountered from outside the system, resulting in further emergence and evolution over time whenever newer, more dominating, beliefs are accepted and a new steady state "equilibrium" of culture is achieved and maintained.

Emergence, evolution, and equilibrium are common complex systems concepts. Emergence is the notion that "the whole is more than the sum of parts...that constitutive characteristics are not explainable from the characteristics of isolated parts...[but] appear as 'new' or 'emergent'," [19], (ch. 3). Evolution is the accumulation and advancement of macro-level changes in a system over a period of time, across any significant property of the system, in any direction. Equilibrium is the balance, or "centeredness" within a system, [19]; a net effect that stems from all micro-level interactions within the system. It may be considered as "the system in an unchanging state," [3], which in the case of macro-level culture takes place when a shared belief is accepted and no longer challenged by individuals at the micro level, reducing variation in emergent culture.

Hence the properties of emergence, evolution, and equilibrium as they relate to culture are important in the modelling process. They describe complex systems phenomena, i.e., organizing forces that promote growth, and disorganizing (chaotic) forces that promote decay. This delicate balance, from the open systems viewpoint, is fundamental to understanding culture as a system; an organic, stabilized construct that both emerges as well as evolves. Unravelling this complex system of culture will require a better understanding of its component structures across levels, as complexity is understood via "the amount of information necessary to describe a system," [3].

5 Seven Dimensions for Cultural Modelling

The key components of culture are diverse and represent both physical and sociological factors that determine the kinds of culture that emerge in a system. Knowing both the components and their properties will provide useful parameters for changing and exploring culture from the bottom-up. This work advocates an approach to model culture in seven dimensions, each based on a primary question: "Does component, or property, X affect the emergence or evolution of culture?" This builds on our previous five-dimensional modelling framework for joint emergency-response operations, [5], [4], which considers the physical, human (individual), functional, and normative dimensions. The seven dimension approach also incorporates a social and information dimension. These relate to the culture definition holistically, with the physical, individual, and functional dimensions referring primarily to elements at the micro-level; the structural, social, and information dimensions referring to elements equally relevant from both viewpoints; and the normative level relating more highly to the macro-level.

The *Physical* dimension relates to components in the actual world, ranging from tools and technology used to common assets such as buildings, cars, clothing, etc. The *Individual* dimension represents actors in the culture, whether they be simple ants, complex machines, smart sensors, or sophisticated cognitive actors like humans. The *Functional* dimension associates a particular role to the individuals within the system, and rests on the notion that the culture preserves itself through what actions are taken by individuals in accordance with their role. The *Structural* dimension characterizes the organizational hierarchy and involves understanding the chain of command of supervisors, subordinates, and colleagues. The *Normative* dimension characterizes policies and rules that govern the behaviour of individuals within the culture. This highlights not only what ought to be done by whom, but also when it needs to be done. This dimension is highly important, as it dictates what the system looks like and how it ultimately behaves and adapts.

The *Social* dimension is used to classify the type of interaction that takes place between actors, as the nature and speed of social communication are often essential to the whole system (e.g., internet-based cultures develop and evolve quickly). The social also refers to how individuals interrelate, including factors such as trust and reputation ("willingness to take risk", [11]), and information sharing (willingness to share sensitive information). The *Information* dimension represents elements that the system both consumes and produces as it performs its function. This level characterizes information and who the producers and consumers of information are at a given time. The properties of information available (like classification or sensitivity) influence the culture in organizations that depend on this information.

These dimensions are further discussed in, [12], and are useful in defining cultural parameters (or components), depending on the model domain. They are mapped to a particular dimension, and eventually used as a factor in an individual's internal belief system. For instance casual dress code culture depends on physical parameters such as location; individual parameters like the degree of comfort with formal dress of a person; functional parameters like having a back-office role with low visibility/interaction with the public; structural parameters like degree of communication with superiors; social parameters like whether communication is always formal, or implicit based on observation of neighbours; information parameters like whether the dress code was communicated at all; and the normative parameters like the policy of being casual for a particular day-of-the-week. These elements together would describe a single culture system based on dress code. A more detailed example is seen in the following sections (see Table 1).

6 Exploring Emergence and Evolution of Culture with Multi-agent Simulation

In order to test these notions of culture, we model a basic organization having roles, norms, and structure using multi-agent systems simulation. We use the notion of a *belief set equilibrium* to display culture, which represents the balance and change in beliefs over all individuals in the system at a given time. When multiple agents begin interacting, forces cause some beliefs to be accepted by the community and become part of the culture (i.e., social memory). Such a force may be a new manager, for example, who has authority over (a) particular agent(s). As more agents join the organization, the culture that has stabilized becomes more resilient to change. However, if a major destabilizing force occurs (e.g., a key agent such as a manager in an organization is replaced), then a cultural shift may occur, eventually resulting in a new belief equilibrium. As such shared beliefs are maintained as a central concept.

A theoretical motivation for the approach is found in, [2]. Social actors engage in social processes called events, (ch. 3), which result in the notion of meaningfulness and is created by powerful organizational actors, such as managers, who are able to construct and maintain organizational rules. Anyone participating in an organization does so by interpreting events and influencing the meanings that others give to these events, (ch. 6). Rules develop and change through the actions of numerous actors as they establish, enact, enforce, misunderstand, resist, and/or break the rules. It is precisely the configuration of these rules and actors involved that constitute a specific culture.

In order to show emerging culture, we demonstrate how the belief set equilibrium of our basic organization is affected under three conditions: i) the effect of adding the most influential agents in the organization at the beginning of the time interval, ii) the effect of adding the most influential agents in the organization in the middle, and iii) the effect of adding the most influential agents at the end. These agents are described below, with an influencing factor dependent on role occupied, personality, and existing social connections within the organization.

6.1 Cultural Belief Set

A cultural belief set (CBS) contains beliefs that exist in the organization's cultural landscape. These may be about particular attitudes, values, goals, or practices. For this work, each belief in the CBS can assume one of three values, based on deontic logic: prohibited, permitted, or obligated. As an example, a belief that "punctuality = prohibited" means that it is culturally unacceptable to be punctual; "punctuality = permitted" means that it is culturally neutral whether or not someone is punctual; and "punctuality = obliged" means that it is culturally required to be punctual.

Since the belief value in the CBS has been restricted to three possibilities, the current culture's value of a particular cultural belief, x, in the CBS can be ascertained by determining which of the three possible values has the greatest consensus among the various individuals in the organization. If there is a tie and one of the tied values matches the previous value, then the previous value will be used. This can be likened to the effect of tradition. Otherwise, permitted will always be used if it is part of the tie, and obliged if permitted is not in the tied set.

6.2 Influence Calculation

The influence of one agent over another agent is used as the mechanism for changing culture. It is based on the notion described previously that key individuals in the organization have a greater influence on its culture. This influence can be computed using factors from each of the seven dimensions. In this chapter, the factors in Table 1 have been incorporated into the influence calculation and are part of the influence factor set (IFS).

The influence calculation, ι_1, of agent b on agent a is seen in Equation 1 below.

$$\iota_1 = \sum_{j=1}^{n} (IFS_a(j) - IFS_b(j)) * \alpha_a(j), \tag{1}$$

where n is the number of items in the influence factor set (IFS) involving agent a's beliefs about agent b (items 1-7 in Table 1); j is an index to a row in the IFS table and the corresponding impact factor, α, for that row; IFS_a is the influence factor set for agent a; IFS_b is the influence factor set for agent b. Equation 2 represents a similar calculation, but for influences that do not involve agent b directly.

$$\iota_2 = \sum_{j=1}^{n} IFS_a(j) * \alpha_a(j), \tag{2}$$

where n is the number of items in the influence factor set involving agent a's personal values (items 8 -13 in Table 1); j is an index to a row in the IFS table and the corresponding impact factor, α, for that row. Agent a's first-hand experiences are considered in items 8, 12, and 13 in Table 1, where the value

Table 1. Factors incorporated into the influence calculation and influence factor set (IFS)

Cultural Influence Factors		
Structural	1	How does agent A relate structurally (within the context of an organization) to agent B? {supervisor, subordinate, colleague}
Physical	2	How close is agent A's workstation from agent B's workstation? {proximity_Threshold} (agent A has a greater chance of being influenced by agents within its proximity threshold)
Functional	3	How similar is agent A's role to agent B's role? [0-1]
Individual	4	Do agent A and B share the same gender? {true, false} (agent A has a greater chance of being influenced by an agent with the same gender)
	5	Are agent A's and B's personalities congruent? [0-1] (agent A has a greater chance of being influenced by an agent with a congruent personality)
	6	How does agent A's experience in the organization compare with agent B's experience? (agent A has a greater chance of being influenced by an agent with more experience)
	7	How does agent A's leadership ability compare with agent B's leadership ability? (agent A has a greater chance of being influenced by an agent with more leadership ability)
Normative	8	Is the particular belief from the CBS formally or informally specified? (an agent has a greater chance of quickly shifting its cultural belief if it relates to a norm that is formally specified)
Social	9	Does agent A seek peer validation from agent B? [0-1] (this may be due to several factors)
	10	Does agent A trust agent B? [0-1]
	11	Through what medium does agent B principally communicate to agent A? {face-to-face > Web 2.0 > phone > email}
Information	12	Does agent A experience the cultural feedback first-hand or second-hand from agent B? (this speaks to the strength of the confidence interval)
	13	If directly, does agent A receive feedback via verbal or non-verbal cues? (this speaks to the strength of the confidence interval; besides verbal cues may be misinterpreted)

in the IFS is adjusted depending on whether it is known to be true or false. Finally, the total influence calculation for agent a is $\iota_1 + \iota_2$.

Table 2. Influence and impact factors used in the CBS

	Influence Factors	**Impact Ratios**
1	Structural Relation	Structural Impact Ratio
2	Workstation Proximity	Distance Impact Ratio
3	Role Similarity	Role Impact Ratio
4	Gender	Gender Impact Ratio
5	Personality Similarity	Personality Impact Ratio
6	Experience Similarity	Experience Impact Ratio
7	Leadership Similarity	Leadership Impact Ratio
8	Formally Specified	Formality Impact Ratio
9	Seek Validation	Validation Impact Ratio
10	Trust	Trust Impact Ratio
11	Communication Medium	Communication Impact Ratio
12	First-hand Feedback	First-Hand Impact Ratio
13	Verbal Feedback	Verbal Impact Ratio

6.3 Updating the Cultural Belief Set

In the simulation, agents share cultural beliefs with other agents whenever a cultural event takes place. These events occur whenever an agent tests a cultural belief in its CBS'. (CBS' is used to distinguish the agent's personal belief set from the organizational belief set CBS which represents the current culture.) These events take the form of a fact in the world, e.g., $agent_a culturalbelief = value$. The current agent, $agent_a$, is enacting a specific belief in its CBS'. This agent will receive direct feedback—praise or chastisement—from the other agents in the organization. This feedback is in the form of $agent_b culturalbelief = value$. If the value from $agent_b$ matches $agent_a$'s value, the behaviour or belief is being positively reinforced; otherwise, it is being negatively reinforced. An agent's cultural beliefs are reconsidered everytime the agent experiences an event. The other agents also experience the event, but their feedback is received second-hand, or indirectly. Events that are experienced first-hand by the agent will have a greater impact on the value of a cultural belief than events that are experienced second-hand. This is accomplished via $IFS(12)$ in Table 1.

For each belief, x, in an agent's CBS', a confidence value is associated with each of the three possible values—i.e., prohibited, permitted, or obliged. In order for the value of x to change, the confidence related to one of the other possible values must become the new maximum. These confidence values are based on the beliefs expressed by other agents, following a cultural event, combined with the influence of other agents' based on previous calculations in Equations 1 and 2 (see Table 2 and 3). For instance, dressing up casually may start as a prohibited belief for $agent_a$, but as more and more interactions take place with different belief values, eventually the permitted or obligated value may become the new maximum, meaning that $agent_a$'s belief value will change. Equation 3 shows the confidence calculation associated with the three possible values of belief x inside $agent_a$'s CBS'.

$$\Phi_{\mu}(x) = \sum_{i=1}^{k} \frac{\beta(x, i, \mu) * \iota_i}{k} \tag{3}$$

,
where the value of μ is one of the three possible values of x: prohibited, permitted, obligated; x is the belief under consideration in the CBS'; k is the number of agents in the system; ι_i is the influence of $agent_i$ on the current agent (in Equation 1 and 2); β is the function below which produces a 1 if $agent_i$'s value for belief x matches the value currently under consideration, i.e., μ.

$$\beta(x, i, \mu) = \begin{cases} 1 \text{ if } CBS'_i(x) = \mu \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

After each cultural event, the agents recompute confidence for all three possible values for each belief in their CBS'. As it relates to the CBS', if there is a tie between the confidence values for belief x and one of the tied values matches the agent's current belief value, then the agent's current belief value will be used. Otherwise, permitted will always be used if it is part of the tie, and obliged if permitted is not in the tied set.

Ultimately, the belief value with the greatest confidence will be selected by the agent for cultural belief x. However, if an agent's confidence is below a certain threshold (unique to the agent), then the agent will feel free to "test" this cultural belief by performing counter-cultural behaviours, i.e., the agent may perform an action that is counter to the belief value in the CBS. These "agents-of-change" have high confidence and can shift an institution's cultural belief set, [18], leading to evolution, and eventually, as new confidence values for all agents climbs above their individual thresholds, the CBS will stabilize to a new equilibrium.

7 Simulation Experiments

We present three experiments involving a model of a small, generic organization over a fixed time period, from initial inception of the organization (i.e., from a single agent) to its achievement of a full population and a stable culture (i.e., all agents are added to the organization for the period and no more culture testing is done by agents). The objective is to show emergence through interaction in the CBS, emergent evolution, and emergent equilibrium. We use the Brahms multi-agent development environment, [7], that builds on the Beliefs-Desires-Intentions (BDI) paradigm, [14], for ease of implementing belief-based agents. Goals and intentions are not considered, although they may add to future work.

Table 3. Initial values for each agent's self-influence (α_i) and cultural beliefs are shown below.

Agent	α_i	Overtime	Formal Attire	Punctuality
$agent_1$	60	permitted	prohibited	obligated
$agent_2$	55	obligated	prohibited	obligated
$agent_3$	67	obligated	prohibited	permitted
$agent_4$	77	prohibited	obligated	permitted
$agent_5$	44	prohibited	obligated	obligated
$agent_6$	64	prohibited	obligated	permitted
$agent_7$	74	obligated	obligated	prohibited
$agent_8$	64	prohibited	obligated	permitted
$agent_9$	34	obligated	prohibited	obligated

The organization, an IT startup, consists of the following nine agents: an owner, receptionist, payroll manager, IT manager, and five generic workers. These agents are fully connected to each other in terms of communication, but with "subordinate-to" and "colleague-of" relationships based on role. This means that a worker agent that is influential can still be able to communicate with the owner of the organization, for instance, and can represent informal networking of potentially influential agents who may not hold powerful positions in an organization. In the model shown in Figure 1, Agent_1 is the Owner, Agent_2 is the IT manager, Agent_3 is the Receptionist, Agent_4 is the Payroll manager, and the remainder are generic worker agents under the IT manager. Agents 1, 2, and 4 are given the most influential property values from Table 2 and, thus, have the highest influence value across all agents based on Equations 1 and 2.

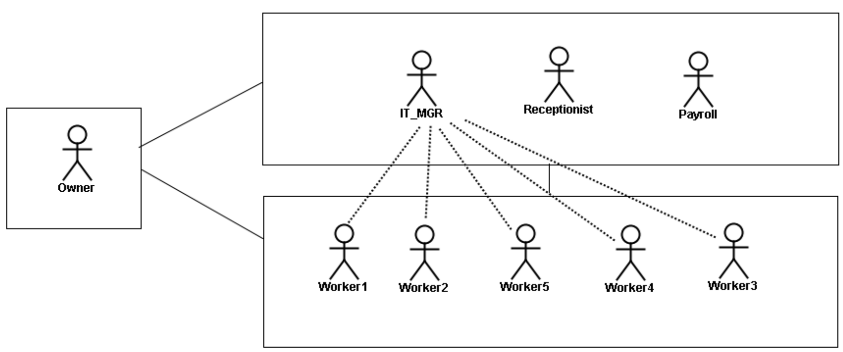


Fig. 1. A simple agent organization consisting of nine agents. The most influential agents are the Owner, IT manager, and Payroll manager. Each agent is fully connected with all other agents, as evidenced by the connection between rectangles. The dotted lines indicate supervisor-subordinate relationship between the IT Manager and worker agents.

Each agent begins with an initial set of beliefs pertaining to both the CBS and the influence factors and impact ratios which were described previously (see Table 3). The CBS in the following experiments is comprised of the following three beliefs that are heavily determined by the culture of the agent organization: i) working after hours (overtime), ii) appropriate business attire, and iii) punctuality. The agents' confidence in whether these are prohibited, permitted, or obligated at any time during the simulation shows the cultural pattern of the organization. As a result, three separate runs of the simulation are conducted, with different ordering for when the most influential agents (Owner (agent_1), IT manager (agent_2), and Payroll manager (agent_4)) are added to the organization. In the first experiment the simulation is run with the three most influential agents added to the system at the beginning of the simulation period. The second experiment adds these agents at the middle of the simulation period. The third simulation adds these agents near the end of the simulation period. It is expected that the culture should evolve differently based on when these agents are added.

7.1 Visualizing the Cultural Belief Set

In presenting culture visually, radar plots are used to show a) the cultural belief values in the CBS that ultimately become the dominant culture (axis labels), b) the number of agents present in the system when a cultural sampling is taken (edge numbers), and c) the *shape* of the resulting cultural system (which will be a triangle, since the CBS used in the experiments contains three beliefs). When the triangle is an equilateral one, it means there is complete cultural consensus among the agents; that is, the emerging culture has reached a state of equilibrium.

It bears highlighting that different orderings of agents result in different cultures emerging (the belief values in the axes are different across the experiment plots). Trends in the shapes, or orientation, of cultures over time show resilience and stability according to the variation of shape. Changes in the size of the plot represent variation in culture maturity, but also indicate the number of agents in the organization. These plots capture the system in a unique way that can be extended to CBS's of different sizes. Points near the origin show number of agents holding a belief to be true at the end of a CBS calculation. Note for each different culture, three different cultures emerge (shown on axes) depending on what interactions take place.

7.2 Experiment 1: Adding Most Influential Agents at the Beginning

In this experiment, the organization begins with the three most influential agents: the owner and the two managers. These agents then have one simulated month to perform cultural interactions before a new agent is added. During this time, two of the agents agree that employees must work after hours and be punctual, and all three agree that business attire is not that important (see Figure 2). After the one month period, another agent is added to the organization. Once again, the agents take one simulated month to perform cultural interactions before the next agent is added.

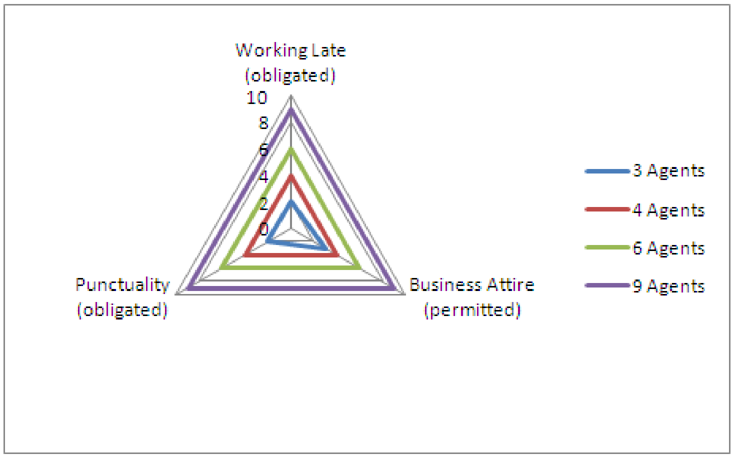


Fig. 2. Experiment 1: Adding Most Influential Agents at the Beginning. Cultural beliefs stabilize after the fourth agent is added.

As can be seen in Figure 2, once four agents are added to the organization, the cultural belief set stabilizes and other agents added to the system adopt the organization's culture. This is because the existing agents are sufficiently influential and eventually convince all existing agents within the organization to conform to their culture. So it can be said that the culture is resilient to change.

7.3 Experiment 2: Adding Most Influential Agents in the Middle

In this experiment, the organization's three most influential agents are added to the organization after three other less-influential agents have performed cultural interactions for a month. The owner and two managers are added separately in successive months, before the remaining three agents are added in the same manner.

As can be seen in Figure 3, complete stabilization of the culture does not occur until six agents have been added to the organization. This suggests that the influence of the most powerful agents impacted the initial culture of the organization, which existed during the first month when the three initial agents were present. This likely occurred because none of the first three agents were sufficiently influential to convince the other agents to adopt their cultural position.

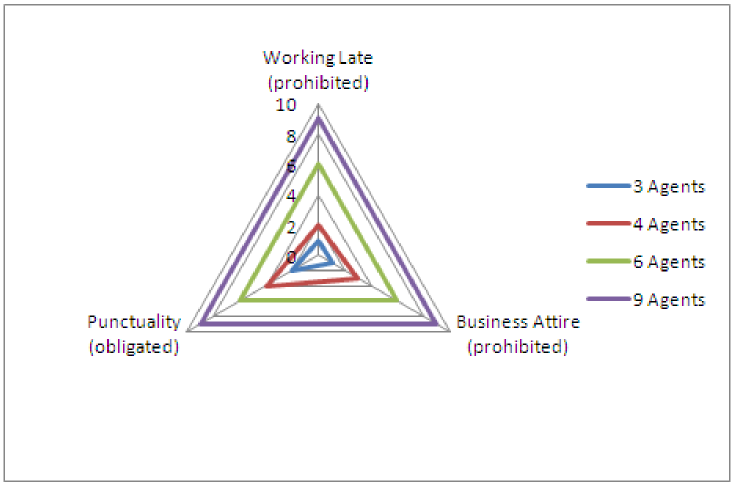


Fig. 3. Experiment 2: Adding Most Influential Agents in the Middle. Cultural beliefs stabilize after the sixth agent is added.

7.4 Experiment 3: Adding Most Influential Agents at the End

In this experiment, the organization's three most influential agents are added to the organization late in the simulation, in incremental time steps, following the initial three agents and the three other less influential agents.

As can be seen in Figure 4, complete stabilization of the culture occurs once six agents have been added to the organization. This suggests that even though the most influential agents are not added until the end, the first six agents are able to create enough "pull" together to compensate for the greater influence of these other three agents. Because these influential agents are added individually, neither one alone is able to overcome the cultural stability (or resilience) already existent within the organization.

8 Conclusion

In this paper, culture has been defined and presented as a complex, multi-dimensional, and multi-agent construct. The complex systems viewpoint is valuable as it allows for considering culture holistically, from both a top-down (emergence) and bottom-up (based on influence and local rules) perspective. The multi-dimensional viewpoint adds to existing literature on modelling of culture's component dimensions with the addition of a seven-dimensional approach. The multi-agent modelling and simulation of culture puts the complex systems and seven-dimensional model into perspective with the notion of achieving belief-based equilibrium of agents over time, according to relationships, communication, and influence idiosyncracies of each agent as individuals in an organizational

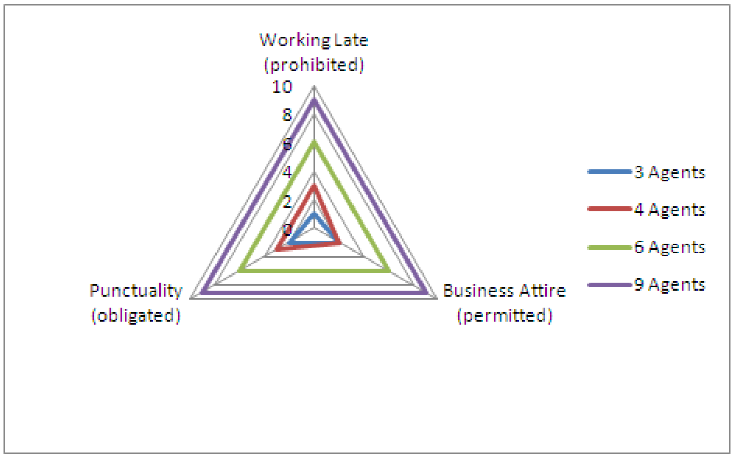


Fig. 4. Experiment 3: Adding Most Influential Agents at the End. Cultural beliefs stabilize after the sixth agent is added.

system. This simulation has been developed for a small test organization with high functioning BDI-based agents.

Three initial simulation experiments have been conducted, showing how culture emerges for different configurations of the same agent organization, depending on when agents of change having high influence levels are added to the system. The developed multi-agent simulation shows that culture can be modelled and visualized in a new way. Future work will involve further testing of the simulation with organizations of different configurations in order to better understand the resilience of culture, and what conditions are needed to allow for an agent of change to sweep through an organization. Furthermore, studying the integration of different culturally-oriented organizations is also of interest.

References

1. Artikis, A., Kamara, L., Pitt, J., Sergot, M.: A protocol for resource sharing in norm-governed ad hoc networks. Declarative agent languages and technologies II pp. 221–238 (2005)
2. Ashkanasy, N., Wilderom, C., Peterson, M.: Handbook of organizational culture & climate. Sage Publications, Inc (2000)
3. Bar-Yam, Y.: Dynamics of complex systems. Perseus Books Cambridge, MA, USA p. 848 (1997)
4. Bicocchi, N., Ross, W., Ulieru, M.: A Cognitive Agent Framework enabling Safety and Security Operations. Downloaded on March 6, 2011, from http://www.cs.unb.ca/ ulieru/Publications/MASS2010.pdf.
5. Bicocchi, N., Ross, W., Ulieru, M.: A simulation modelling approach enabling joint emergency response operations. In: SMC. pp. 1832–1837 (2010)
6. Carr, H., Pitt, J., Artikis, A.: Peer pressure as a driver of adaptation in agent societies. Engineering Societies in the Agents World IX pp. 191–207 (2009)

7. Clancey, W., Sachs, P., Sierhuis, M., Van Hoof, R.: Brahms: Simulating practice for work systems design. International Journal of Human Computer Studies 49(6), 831–866 (1998)
8. Harmon, D., Lim, M., Bar-Yam, Y.: Advanced Mathematical Science of Ethnic Violence. Conflict Management and Peace Science 27(2), 177 (2010)
9. Hofstede, G.: Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. Sage Publications, Inc (2001)
10. Kroeber, A., Kluckhohn, C., Untereiner, W., Meyer, A.: Culture: A critical review of concepts and definitions. Vintage Books New York (1952)
11. Mayer, R., Davis, J., Schoorman, F.: An integrative model of organizational trust. Academy of management review pp. 709–734 (1995)
12. Morris, A., Ross, W., Hosseini, H., Ulieru, M.: Modelling Culture with Complex, Multi-dimensional, Multi-agent Systems. In: Dignum, V. (ed.) Integrating Cultures: Formal Models and Agent-Based Simulations (In Submission). Springer (2011)
13. Morris, A., Ross, W., Ulieru, M.: A system dynamics view of stress: Towards human-factor modeling with computer agents. In: SMC. pp. 4369–4374 (2010)
14. Rao, A., Georgeff, M.: BDI agents: From theory to practice. In: Proceedings of the first international conference on multi-agent systems (ICMAS-95). pp. 312–319. San Francisco, CA (1995)
15. Sierhuis, M., Jonker, C., van Riemsdijk, B., Hindriks, K.: Towards organization aware agent-based simulation. practice 16, 17
16. Sterman, J., Sterman, J.: Business dynamics: Systems thinking and modeling for a complex world with CD-ROM. Irwin/McGraw-Hill (2000)
17. Stratulat, T., Ferber, J., Tranier, J.: MASQ: towards an integral approach to interaction. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2. pp. 813–820. International Foundation for Autonomous Agents and Multiagent Systems (2009)
18. Ulieru, M., Verdon, J.: Organizational transformation in the digital economy. In: Industrial Informatics, 2009. INDIN 2009. 7th IEEE International Conference on. pp. 17–24. IEEE (2009)
19. Von Bertalanffy, L.: General system theory: Foundations, development, applications. G. Braziller New York (1968)

Substantiating agent-based quality goals for understanding socio-technical systems

Sonja Pedell[1], Tim Miller[2], Leon Sterling[3], Frank Vetere[1], and Steve Howard[1]

[1] University of Melbourne, Department of Information Systems, Victoria, 3010, Australia, {pedells,fv,showard}@unimelb.edu.au

[2] University of Melbourne, Department of Computer Science and Software Engineering, Victoria, 3010, Australia, tmiller@unimelb.edu.au

[3] Swinburne University, Faculty of ICT, Victoria, 3022, Australia, lsterling@swin.edu.au

Abstract. In this paper we propose a method for using ethnographic field data to substantiate agent-based models for socially-oriented systems. We use the agent paradigm because the ability to represent organisations, individuals, and interactions is ideal for modelling socio-technical systems. We present the results of in-situ use of a domestic application created to encourage engagement between grandparents and grandchildren separated by distance. In such domains, it is essential to consider abstract and complex quality requirements such as *showing presence* and *sharing fun*. The success of such domestic technologies is based on the meaningful realization of these difficult-to-define quality goals. Our method addresses the need to adequately inform these quality goals with field data.

We substantiate the quality goals with field data collected by introducing an application into the home of three families. The field data adds an understanding of what *sharing fun* means when "filled" with concrete activities. The quality goals served as a template to explore and represent the rich field data, while the field data helped to formulate the requirements for a more complex and refined technology. This paper's contribution is twofold. First, we extend the understanding of agent-oriented concepts by applying them to household interactions. Second, we make a methodological contribution by establishing a new method for informing quality goals with field data.

Keywords: Socially-oriented requirements, ethnography, quality goals

1 Introduction

Despite best efforts, contemporary technologies often fail to meet basic human needs and desires. Recent developments have ensured technologies are generally accurate, reliable, and usable. However, meetings these measurable requirements and qualities constitute only part of what it means to design technology for people. As social beings we have complex and hard-to-measure needs, such as to experience social connection and empathy, to care for others and be cared for, and to share pleasure. These particular types of social requirements cannot be easily reduced to functional specifications. The functionality of a socially-oriented system is often unclear: how does one measure whether a system is able to facilitate a complex goal such as "being fun?"

Pavel et al. [15] argue that agent-based models are suitable for understanding the complex topics inherent to socio-technical systems because the concepts used in these models are suitable for expressing the organisational and behaviour aspects of individuals and their interactions. Our reason for using agent modelling is because they allow us represent human behaviour as well as representing the software system behaviour as a software agent. We define a method to learn more about quality goals in these systems and illustrate it via a case study for exploring intergenerational relationships. We believe that our approach can be applied to other areas where complex social goals have to be considered that need to be written down and implemented in a policy.

Good policies guide decisions and achieve rational outcomes containing the 'what' and 'why' something needs to be done [2]. The 'what' and the 'why' we capture in a motivational goal model. With the associated quality attributes of the goals, we aim to understand the concrete activities for each role to achieve these goals in reflecting on 'how' these goals are fulfilled best. Our tools and techniques are relevant for any social topic involving technology use, but we also argue that they can guide policy decisions in the same way they guide technology design decisions. The difference comes at the production stage, where policies are explicitly written down and have a more formal character in guiding social behaviour. We suggest this paper is relevant for policy making for the following reasons: (1) for giving clear guidelines that can be followed; (2) for understanding of quality goals that are relevant for policy making; and (3) as a basis for communication when defining non-instrumental goals.

In our method, developers first define a high-level goal model that includes the high-level quality goals, such as having fun. Ethnographic techniques are then used to obtain data about the particular domain, and the goal models are used as a template through which the data is analysed. From the data, themes are extracted, and each theme is attributed to a high-level quality goal. If a theme does not match a quality goal, this triggers a discussion as to whether a new quality goal is required. The result is an agent model with concrete themes for achieving quality goals.

Our particular case study focuses on technology for supporting the relationship between grandparents and grandchildren that are separated via distance. This case study presents many interesting and challenging problems for defining innovative technologies with hard-to-define quality goals.

There are several broad aims within our larger research project, including:

1. To increase the modelling capability of social domains using agent concepts.
2. To understand the goals and their associated qualities better in the light of technology use over a distance.

3. To provide a method for designing and implementing quality requirements within complex social settings, such as the domestic space.
4. To build domestic technologies that are better suited to the needs of grandparents and their grandchildren.

2 Socially-oriented requirements engineering

When information and communication technologies began diffusing into the home they did so originally mostly as extensions of our places of work, but this is changing [13]. Domestic technology is generally successful if it satisfies both functional and non-functional needs and if every member of the family from the very young to the very old is capable of operating and enjoying it. But there are characteristics of the home that make designing domestic technologies unique and challenging. Domestic needs are often unspoken; relationships are not straightforwardly hierarchical; lived life is idiosyncratic and even exotic [8]. Technologies for strengthening bonds within separated families must fulfil hard-to-define goals such as *being playful* and *engaging over distance.* Such social goals — which are ambiguous, non-instrumental, subtle and long-term [13] — are difficult to describe and account for in ways that are appropriate for technology development. Development tools typically deal best with clearly defined, hierarchical goals that endure over a specified time frame. Domestic and social goals do not fit well with traditional software engineering methods and processes.

Getting from domestic lives and routines to useful and suitable technologies for the home presents many challenges. One of the big challenges for domestic design is that there is no such thing as a "typical home" [19]. Leonardi et al.[12] describe the home as "*a 'territory of meaning', a place where pleasure, affect and aesthetics are deeply interwoven with the functional and utilitarian dimensions.*" There is still a gap on how to design these technologies as inhabitants have needs that are not easy to articulate, they represent a diverse population, and needs are non-functional and often even ambiguous [8].

Ethnographic data can be used to understand social activity as it happens [18]. In order to create domestic technologies and to inform software development, we need tools that are able to carry the complex, abstract and often ambiguous insights of field data collections into the development process. However, to do this we need a way to represent the insights from fieldwork with artefacts that are shared by field researchers and software engineers, and still carry the voice of the user. Software engineers have their focus on future technologies and social needs are often neglected in existing software development processes. The researchers' focus is on the current lives of people. Consequently there is a gap that both groups have to bridge in the design process.

This work is not about the development itself, but the way discussions and thinking take place when talking about the goals and values of socio-technical systems. First, we gain an understanding of the social part of the system as a basis of what is needed to implement a system that is fit for purpose. The behaviour of the software cannot be defined without understanding the social aspects of the social-technical system in which the software operates.

2.1 Modelling field data

The value of matching socially-oriented studies of human interaction with user requirements has been acknowledged (e.g. Viller and Sommerville [18]). Other researchers describe bridging the gap between the output of field studies and the required input to system designs through meta-modelling [10]. This mapping is based on plans and procedures that need to be clearly specified. However this is not straightforward for socially-oriented requirements. Eliciting socially-oriented requirements from field data involves working in a milieu in which it is essential to capture concepts accurately but flexibly at a high level, without losing the liveliness and vitality of those concepts through over specification. The rich information and knowledge gathered in the field needs to be reshaped to accommodate the more formalized and rigorous models of software requirements elicitation and design when identifying goals for the system, and how they should be operationalised. We want to maintain the richness of data while generating models that can be implemented into technologies. To this end, we suggest that quality goals are a necessary part of the abstraction process because they permit a level of ambiguity that is necessary to represent the complex social concepts found in field data.

2.2 Intergenerational fun

We are particularly interested in how domestic technologies mediate shared experiences and emotions, such as having fun and joy, between grandparents and grandchildren. The grandparent-grandchild relationship is an example of a set of complex social interactions and roles and it is not obvious what kind of technology supports a strong intergenerational relationship. This is complicated further when the intergenerational relationship is nurtured over a distance.

We must look at a family's life more closely to understand emerging interactions in technology use. We analyse these interactions in the light of these interactions and their qualities in order to draw conclusions about the affordances of domestic technologies. Existing technologies are not adequate to bridge the distance between grandparents and their grandchildren [4]. The phone is still the most commonly used technology for children to get in touch with remote family members [16]. However, it is problematic and not suitable for intergenerational interactions as much of the communicated contextual information is lost [1].

2.3 Motivation models

The work in this paper builds mainly on the work of Sterling and Taveter [17]. Their work has focused on how to make high-level agent-oriented models palatable to design discussions. This is achieved using goal models with a straightforward and easy syntax and semantics. Goal models are useful at early stages of requirements analysis to arrive at a shared understanding and ontology [7, 11]; and the agent metaphor is useful as it is able to represent the concepts that we want to capture for socially-oriented systems, such as agents taking on roles associated with goals. These goals include quality attributes that are represented in a high-level pictorial view used to inform and gather input from stakeholders. In Sterling and Taveter's notation, goals are represented as parallelograms, quality goals are clouds, and roles are stick figures; see Figure 1. These constructs can be connected using arcs, which indicate relationships between them.

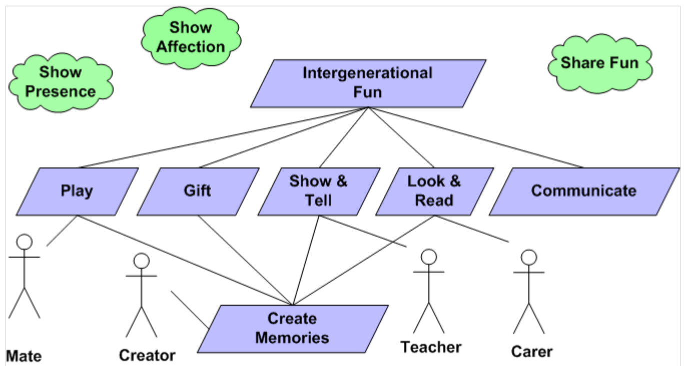


Fig. 1. Motivation model representing *intergenerational fun*.

Quality requirements at the early stages of elicitation tend to be imprecise, subjective, idealistic and context-specific, as discussed by Jureta and Faulkner [11]. Garcia and Medinilla [5] describe high-level quality goals as a specific form of uncertainty that can be used as a descriptive complexity reduction mechanism and to model and discuss uncertainties in the environment. In our requirements elicitation process, we seek complexity reduction without losing the richness of the concepts themselves. Instead of eliminating uncertainty early in the process, we embrace it and withhold design commitment, at least until there is clarity and understanding [6]. High-level goals associated with activities can act as a point of reference for discussing the usefulness of design alternatives to achieve these goals instead of a decomposition into single requirements. The multi-agent paradigm offers benefits over other paradigms because the concepts used in modelling, such as roles, goals, and interactions, are part of everyday language.

From a software engineering point of view the models enable us to take the outputs from a field study and use them to inform socio-technical software design. This is achieved by taking account of the richness of human social interaction provided by the field data, encapsulating quality attributes of that interaction into quality goals in the models and using these models as artifacts for designing technologies that really support and enhance domestic social interaction.

2.4 Modelling with quality goals

Focusing on quality is well established within software and systems engineering. Software engineers are aware of the need to express quality attributes of software as well as functional capabilities of software. These quality attributes are referred to using a variety of terms including: non-functional requirements, constraints, quality attributes, quality goals, or quality of service requirements.

We use the construct of quality goals attached to functional goals to represent the quality attributes of social interactions. Social quality goals are essentially non-functional and are designed to encapsulate social aspects of the context into the software requirements model, thus providing a mechanism to carry subtle nuances of those social aspects through to the implementation phase. These social quality goals remain interpretively flexible, even until the final product, opening up a variety of possible interpretations both in the design and use of the system. We maintain that there is benefit in articulating quality goals without the need to resolve them into measurable goals. Sterling and Taveter's agent-oriented models allow the expression of non-functional requirements by attaching quality goals to goal models.

In our approach there is a direct pairing between system goals and quality goals, whereas non-functional goals do not necessarily have a direct relationship with functional goals [3]. This makes it more difficult to carry them through the process in an unresolved state. Relating an abstract and unresolved quality attribute to a system goal enables a focus on social goals within the design process.

Our starting point is the simple model of motivations of the socio-technical system shown in Figure 1. By capturing and representing quality goals in agent-oriented models we make a commitment to important aspects of social interactions that can remain unresolved, giving interaction designers and software engineers alike a focal concept for analysing and designing around complex social concepts. By externalizing them in a simple format the models become shared artifacts that are able to sustain multiple interpretations across disciplines [13]. Quality goals allow a focus on understanding the reasons why people do things, or the essence of a relationship rather than describing a physical action. In doing so, quality goals capture something that is more dynamic and fluid than other elicitation mechanisms found in usual software engineering practices.

3 Method

Then how can these social goals and attached qualities be fulfilled when using technology and what tools are used best to explore the use of domestic technologies? The success of a design in achieving its goals can really only be investigated after implementation. Therefore we started with building lightweight technologies that focus on certain goals of the goal model. On the goal model level we do not prescribe how to use specific technologies. We purposely keep them on a high level that they are representative and comprehensive to a satisfactory degree, but are independent of one concrete implementation. This way we were able to learn more about qualities that are arising as a consequence from technology use in tying back concrete activities of technology use to the motivational model. With the insights gained from using simple technologies we hope to predict more accurately what will work when building more complex technologies that cover the complete goal model. The main components of our method are the following. We purposely speak about the components of a method and not a process as the activities of these components are taking place iteratively depending on the available knowledge of the user domain.

The main features of our approach are:

- Use of agent-oriented models with a focus on quality goals.
- The implementation of lean, but focused technologies.
- Iterative exploration and discussion of social requirements.
- Lightweight evaluation of quality goals in ethnographic studies.
- Analysis of quality goals and elicitation of social requirements.
- Refining of user needs.

3.1 Electronic Magic Box

We built an application called *electronic Magic Box*, which was inspired by the motivational model in Figure 1. The electronic Magic Box uses synchronous touch screens for displaying and mobile camera phones for sending photographs and messages that were shared among the grandchildren and grandparents households. Each family unit was allocated one mobile phone and one touch screen — that is, one for the grandparents to share — and one for shared use by the children (and parents). The mobile phones were important as we wanted the sharing of everyday experiences that could operate at a distance. Grandparents and grandchildren could carry the phones with them and share photographs of events and ideas with the others sending it to the system. The system was easy to use and tried to constrain the user as little as possible, thereby facilitating flexible interactions without strict assumptions about how technology was mean to be used. The screens were placed in high traffic areas in the family homes such as in the lounge room or the kitchen counter — easily accessible and surveyed by the family members. While our focus was on the grandparent-grandchildren relationship, the parents took on an important role in facilitating interactions and observing them without being directly involved in the use of the system.

We wanted minimal ongoing intervention from the researchers themselves while allowing us to observe the transactions between the participants. The system has logging capabilities to monitor and record the use of the application such as technology probes [9]. Probes are specifically suitable for collecting data in the domestic domain through their ability to capture the nuanced aspects of everyday life. In this study we regard probes as informational, designed to inform about daily life rather than to inspire design. Information and story generation are two important benefits that we see in the use of probes.

The electronic Magic Box allowed the sending of a treasure box that could be filled with photographs and messages. Figure 2 shows the layout of the homepage. On the left side area of the homepage, seven picture based links (home, Magic Box, scroll, collection book, settings, admin, and logout) can be found that guide to a number of destinations within the application. The box is placed in a forest of fern trees and appears either closed (a new box has arrived) or open (no new box has arrived). A scroll either sealed or with a broken seal indicates if the box in the other household has been opened and the content been looked at.

Fig. 2. Homepage of electronic Magic Box.

In order to be able to access the content the receiver has to play a maze game to ‘find’ and open the box. An opened message can be saved in a collection book. Emphasis in this application was put on the goal *gift*, but the concept certainly carried elements of and was inspired by the other high-level goals or motivations like *playing*, *show & tell*, *look & listen*, *communicating* and *creating memories*.

4 Study

4.1 Study design and participants

We introduced the electronic Magic Box to three families. The application was installed in the family homes between three and six weeks over a period of four months. Family one consisted of one eight year old girl living with her mother, with her grandmother living 12 kilometres away (parent of mother). Family two consisted of three grandchildren of 18 months (girl), six years (girl) and eight years (boy) living with their parents 8 kilometres from their grandparent (parents of mother). Family three consisted of two girls in the ages of five and six and their parents with a distance of about 16 kilometres from their grandparents (parents of mother). All grandparents had regular contact with their grandchildren (at least once a week) and all of them described having a strong and loving relationship.

4.2 Data collected

We conducted three to four interviews per household about the probe use (usually grandparent household and grandchildren household separately). The parents were present in the grandchildren interviews. This was an important source of information as the parents were observing the ongoing interactions without being directly involved, and were able to make comparisons on the basis of how the interactions occurred before the introduction of the system.

During the interviews, we did not ask for goals such as *play* or *gift*. These goals are implicit in the system. We were more interested in the actual interactions and how the qualities were judged by the participants. For example, we would ask: 'What kind of interactions did the system support?' and 'what activities did you particularly enjoy?'. The technology probe data collected with the electronic Magic Box application included 102 boxes (electronic letters and photographs) and time stamps for all messages.

4.3 Analysis

The transcribed interviews together with the photographs and messages (text messages and electronic letters) were analysed using content analysis according to Patton [14]. The data was analysed focusing on the quality goals as overarching themes. We investigated and evaluated the activities and interactions and not the technology per se. This procedure enabled us to find sub-themes for all of the quality goals and therefore to learn more about each goal in the light of typical activities between grandparents and grandchildren. Each sub-theme was briefly described and substantiated by compelling examples and instances of these goals in the specific context of intergenerational fun using the applications.

We analysed the interview data according to what we could learn about the quality goals. The photographs and messages were downloaded from the servers and analysed and discussed biweekly. The essence of the quality goals is based on experiences and judgement of the participants regarding their interactions and not easy accessible by the field researchers. Therefore, the interview data played a major role in this analysis as we wanted expand from the activities and original goals to inform the quality goals. The photographs and messages were mostly to back up and illustrate the results with use episodes and participant stories.

This analysis procedure helped us to keep the focus on the human needs with the technology as mediator. We avoid the risk of focusing on the technology as our aim is not to create a perfectly running technology, but implementations that support us in further investigating the social requirements themselves. Even further this approach validated our existing understanding. We were looking for describing social requirements, looking for examples for "this was fun" or "this was not fun". If we had never a comment that said "this was fun" or "this made me feel connected", we would feel invalidated in our original motivational model.

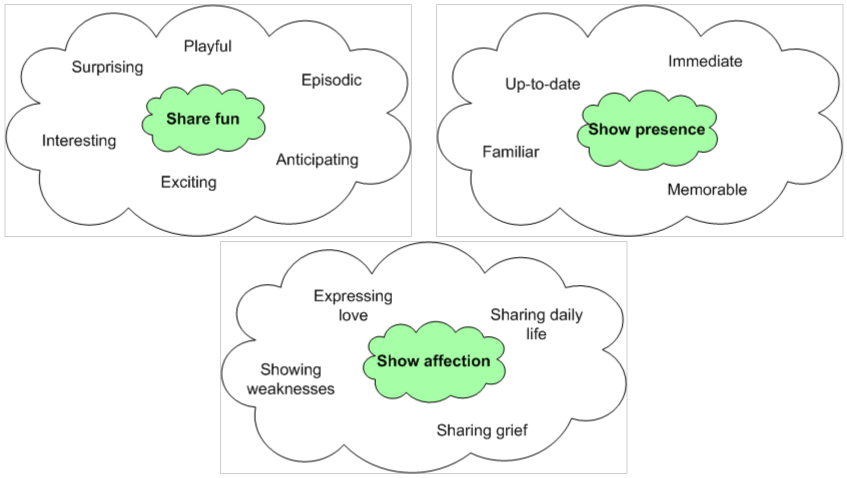


Fig. 3. Quality clouds for the quality goals *share fun*, *show presence* and *show affection*.

5 Results

Many of the insights concerned how to expand or better the system itself such as including more games, providing music options or the possibility to send several boxes in a row. Our focus here is on what we learnt about the quality goals and therefore about the family interactions facilitated by the system and not about the functionality of the electronic Magic Box itself. Different to other approaches, our aim is not to improve the existing system, but to learn more about the quality goals attached to the functional goals, to represent our leanings, and make changes if necessary to our high-level goal model.

5.1 Substantiating the quality goals

The sub-themes that emerged from our data analysis were organised as characteristics to the quality goals into *quality clouds*, shown in Figure 3. The quality clouds consist of one quality goal linked to a functional goal, with associated qualities factored around. The quality clouds can be seen as an abstract representation of field data into which we are able to zoom into the associated quality goal more closely. In this process the sub-qualities or quality attributes were formulated into adjectives to re-connect the qualities in discussions more easily to the functional goals they are attached to. Each sub-quality of a main quality goal is briefly described and directly linked to the respective quotations in the interview data. Here, we only show some of the quotations that led to the quality cloud *show affection* to demonstrate our procedure. Some of the sub-qualities from the clouds that brought us unexpected insights are exemplary described in more detail.

5.2 Quality goal *share fun*

Anticipation: A feeling of looking forward to hear from the other family member.

Grandparents as well as grandchildren were excited when they saw a closed box on the screen and eager to view the box content as soon as possible. All participants involved kept checking the status of the application regularly when they had sent off a box. One grandmother told us:

"... when I opened it up and there was a message — when the box was closed on the screen — that was fun."

Another grandparent expressed it this way:

"I turned it on in the morning. If the red seal was still on I thought 'darn — nothing new' "

Surprise: Something unexpected that happened that was caused by the sender of the box.

When an empty box was sent a kangaroo would jump out of the box (see Figure 4). The families described they had a lot of fun when this function was discovered. One child explained:

"I never knew what was in the box — every box was a surprise — in particular the kangaroo — that was good."

Fig. 4. The surprise kangaroo.

5.3 Quality goal *show presence*

Immediacy: Maintaining the feeling of presence in renewing the contact within a brief time frame.

The grandchildren would lose interest if it took longer than a few hours until a box was sent back while with one grandmother it increased the anticipation.

"If I didn't have a message I was quite disappointed. When the seal was on I was wondering: 'what are they doing'? That made me think much more often of them during the day than it would have otherwise." [grandparent during interview].

"Sometimes when the grandparents would not send back a box and the kids got really impatient I would give them a call: 'I think a box is awaited pretty urgently over here'." [parent during interview].

Giving updates: Having the urge to communicate to a close person all topical news as soon as possible.

Grandparents and grandchildren showed a similar frequency pattern in using the electronic Magic Box as before. A grandmother and her granddaughter with a lot of contact to each other used the application more regularly than the other families where the contact was not as frequent:

"I don't think you can get them interact more than they do. There was always an open flowing communication between the two of them and it was always positive. Now it was a bit different. Instead of Andrea coming home and telling what happened at school it would be about the computer or the photos: 'Have you opened the box? You haven't opened the box!'." [parent during interview].

This shows how some aspects of the relationship have influence on the use of the electronic Magic Box. Interest is based on a stable and loving relationship.

5.4 Quality goal *show affection*

Show weaknesses: Family members are comfortable not only showing the best side, but also failures and weak points, because there is a loving trust within the relationship.

A challenge for most of the grandparents was the managing of the technologies itself. Uncovering this kind of weakness is an intimate act in itself. Problems dealing with the electronic Magic Box were often communicated in a humorous way or loaded with self-irony making the technology handling a shared episode itself. A nice example was one grandmother sending this message:

"Dear Andrea, in trying to send this photo to you I burnt my steak I am having for dinner, yuk!!!"

After this she took a photo of her burned frying pan as well and sent it:

"When I tried to send this message Thursday the machine told me to try again, so here I am. This is the pan I burned while trying to enter the project!!!"

We only tend to show our failures to people we trust and love. Therefore, to the researchers, this grandmother assured in the interview that the pan was "all clean again" and that she had no more disasters. In a similar way one grandmother sent a photograph of her messy desk:

"This is my messy desk. I am trying to catch up with office work."

The granddaughter took it up immediately as something funny and kept saying in one of the interviews:

"Granny you are messy as well — you sent me this photograph of your desk."

That the grandparents admit to weaknesses being adults and "should know better" was received as something special by the grandchildren.

Share grief: The electronic Magic Box was particularly well suited in mediated shared emotions. There was sometimes an urge to transfer something important and emotional. One example was when the granddaughter's dog got really sick and died. The granddaughter wrote her Nanna accompanied with a really sad picture of herself:

"I really miss Sam — really really!"

Her granny answered:

"I have been thinking of her too, but she was very sick & you wouldn't want her to suffer, would you?"

The electronic Magic Box was also used for more complex emotions or situations that required context information in the sense that they were exceeding a simple state of a relationship, but telling longer stories with the aim to be comforted or understood.

5.5 New quality goal *build up confidence*

While we were interested to group the sub-qualities to our existing quality goals, in order to substantiate them with our field data, we permitted new main quality goals emerge, and hence allow changes to our overall goal model. As part of the method, in the event important activities or themes evolve for which we cannot find a home, we define new quality goals. Qualities emerging that we could not group with our existing quality goals were themes surrounding the technology use itself still being close connected to positive feelings - often explicitly described as fun. The new quality goal that emerge is *build up confidence*, shown in Figure 5.

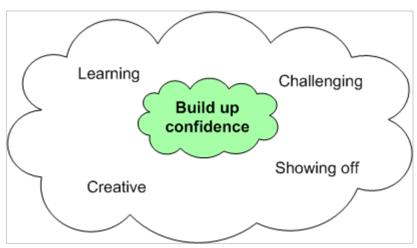


Fig. 5. Quality cloud for "build up confidence".

Learning: One important aspect was being able to continuously improve managing the technology:

"It is quite interesting to see where we started: 'I didn't find a photo, but here is the text'. [an early message from the grandparent]. Next time I was able to send the text as well. It is a bit of fun." [grandparent during an interview].

"I guess I have to get into email now with some kicking and screaming I am enough of a dinosaur. I think I am ready." [grandparent during an interview].

One grandchild could not get enough of the kangaroo magically jumping out of the box and could not figure out how this had happened; the grandmother was proud that she was able to do something unexpected with the technology what links in with building up confidence.

Showing off: Showing the application to people like neighbours, friends and other family members with a feeling of pride.

This theme is a clear sign that confidence indeed had been built up and another example or measure for validating the success of the application that is closely tied to a complex quality goal and not to a certain piece of functionality. One mother said about her daughter:

"Showing them something cool: 'this is what I've got this is mine' — this is my phone and I can send pictures." [parent during an interview].

The showing off effect was in particular interesting with the grandparents. There was a new role the grandparents suddenly had among their peers. They became advocates for new technologies, while they would have never anticipated themselves as champions of new technologies. They found confidence in the technology in a way that we had not planned.

6 Discussion

The quality representations of the field data helped to formulate high-level requirements for a design of a more complex and refined technology concept for grandparents-grandchildren interactions that we are currently building. These requirements are to a large extent influenced by the new quality cloud *build up confidence*. Building confidence is part of the intergenerational interaction and it has implications on how the technology should be designed: not put everything in an application at once, because it scares the grandparents away. We now maintain simple screen views and a layered application instead of a packed one with functionality.

Aiming for simplicity is not only based on the lack of confidence of many grandparents to deal with complex technology, but is suggested in the nature of strong-tie relationships themselves. In the sub-themes becomes apparent that these technologies rely on an existing rich and loving relationship. A lot is carried by these strong tie relationships that we can observe in the home. To support the long-term interactions between grandparents and grandchildren with technology, it is not necessary to build complex technologies. The technologies serve as a mediator of these subtle and complex relationships in the family context and routines

Another important insight was discovering "the other side of fun". Certain value sets have so far been marginalised to date such as disclosing weaknesses or failure — and laughing about them — or the demonstration of grief and openly dealing with it. The grandmother does not try to brush the grief away with some happy comment, but she honestly acknowledges that the loss of the loved dog indeed is sad. According to our results, the dealing with these kinds of emotions is just as important for a strong tie relationship as demonstrating love, play together and laugh about a joke. It is no contradiction that technologies for intergenerational fun also allow and even aim for activities that deal with aspects we would normally avoid to show openly.

In this sense the quality goals represent the essence of an intergenerational relationship independent of a specific implementation or even technology use at all. The motivation model tells us something more general about the values between grandparents and grandchildren. While we have chosen the domestic domain as a challenging example for demonstrating our approach, we believe that this approach could also be relevant for defining and substantiating the main quality goals and values important for communities. The AOSE models are able to represent the values and desired outcomes of social life and can serve as a shared source of discussion and decision making in community and government led projects.

7 Conclusion

As social ICTs become more, and more relevant for the home and families, software engineering needs processes to cater for and understand these complex and sensitive social goals. We propose a replicable process of interleaving motivational models and lightweight technologies to be able to analyse, substantiate and evaluate quality goals in the light of these mutually influenced artefacts. We based our exploration of family life on field data and advanced our understanding about the intended social outcomes (quality goals) in using a technology probe. This approach allowed us firstly the novel use of agent-based methods in building a bridge between ethnographers and software engineers and secondly to show how the notion of goal-oriented analysis, in particular the notion of quality goals, can be useful for the interpretation of ethnographic data.

Quality goals allow a focus on understanding the reasons why people do things or the essence of a relationship rather than describing a state of the world or an action. With quality clouds, we were creating a set of new testing artefacts for lightweight evaluation. They were useful in the process to validate associations between activities and high-level goals and evaluate the degree of the match between the two. The proposed method helped us to substantiate quality goals for social interactions for the development of meaningful domestic technologies, helping us to bridge the gap between the agent-oriented models, and the ethnographic data.

Traditional quantitative evaluation methods do not apply for the evaluation of socio-technical systems and require new approaches. We proposed a time intensive user study to ensure that users' evaluation of the system is based on real social experiences with the system. We made use of an iterative and qualitative evaluation process as we do not see a way to use traditional software engineering metrics to measure having fun mediated by a system. Our evaluation is engineered in the sense that we evaluated applications comparing and evaluating them against the original agent-based model. In our future research we aim to look more closely into finding a more formal, less descriptive, and less time intensive evaluation process for social quality goals. While we have chosen the domestic domain to evaluate our appoach, we believe that this approach is applicable for defining and substantiating quality goals in other domains.

References

1. R. Ballagas, J.J. Kaye, M. Ames, J. Go, and H. Raffle. Family communication: phone conversations with children. In *Proceedings of the 8th international Conference on Interaction Design and Children*, pages 321–324. ACM, 2009.
2. K. Blakemore. *Social Policy: an Introduction*. Open University Press, 1998.
3. L. Chung, B.A. Nixon, E. Yu, and J. Mylopoulos. Non-functional requirements in software engineering. 2000.
4. B. Evjemo, G.B. Svendsen, E. Rinde, and J.A.K. Johnsen. Supporting the distributed family: the need for a conversational context. In *Proceedings of the NordiCHI'04*, pages 309–312. ACM, 2004.
5. A. Garcia and N. Medinilla. The ambiguity criterion in software design. In *International Workshop on Living with Uncertainties (IWLU'07)*. ACM, 2007.
6. D. Gause. User driven design – the luxury that has become a necessity, a workshop in full life-cycle requirements management. In *ICRE 2000, Tutorial T7*, 2000.
7. R. Guizzardi and A. Perini. Analyzing requirements of knowledge management systems with the support of agent organizations. *Journal of the Brazilian Computer Society*, 11(1):51–62, 2005.
8. S. Howard, J. Kjeldskov, and M.B. Skov. Pervasive computing in the domestic space. *Personal and Ubiquitous Computing*, 11(5):329–333, 2007.
9. H. Hutchinson, B. Westerlund, B.B. Bederson, A. Druin, M. Beaudouin-Lafon, H. Evans, and N. Roussel. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24. ACM New York, 2003.
10. R. Iqbal, A. James, and R. Gatward. Designing with ethnography: An integrative approach to CSCW design. *Advanced Engineering Informatics*, 19(2):81–92, 2005.
11. I. Jureta and S. Faulkner. Clarifying goal models. In J. Grundy et al., editor, *Proceedings of ER*, volume 83 of *CRPIT*, pages 139–144, 2007.
12. C. Leonardi, C. Mennecozzi, et al. Knocking on elders' door: investigating the functional and emotional geography of their domestic space. In *Proceedings of OZCHI'09*, pages 1703–1712. ACM, 2009.
13. J. Paay, L. Sterling, F. Vetere, S. Howard, and A. Boettcher. Engineering the social: The role of shared artifacts. *International Journal of Human-Computer Studies*, 67(5):437–454, 2009.
14. M.Q. Patton. *Qualitative research and evaluation methods*. Sage Publications, Thousand Oaks California, 2002.
15. J. Pavón, M. Arroyo, S. Hassan, and C. Sansores. Agent-based modelling and simulation for the analysis of social patterns. *Pattern Recognition Letters*, 29(8):1039–1048, 2008.
16. N. Romero, P. Markopoulos, J. Baren, B. Ruyter, W. Ijsselsteijn, and B. Farshchian. Connecting the family with awareness systems. *Personal and Ubiquitous Computing*, 11(4):312, 2007.
17. L. Sterling and K. Taveter. *The Art of Agent-Oriented Modelling*. MIT Press, 2009.
18. S. Viller and I. Sommerville. Ethnographically informed analysis for software engineers. *International Journal of Human-Computer Studies*, 53(1):169–196, 2000.
19. A. Zafiroglu and M. Chang. Digital homes on wheels: designing for the unimagined home. *Personal and Ubiquitous Computing*, 11(5):395–402, 2007.

The Benefits of Agent-based Motivation Models in Policy Formulation and Implementation

Sonja Pedell[1] and Leon Sterling[1]

[1]Faculty of ICT, Swinburne University of Technology
3022, Hawthorn, Australia
{spedell, lsterling}@swin.edu.au

Abstract. Our aim is to demonstrate how agent-based motivation models can play a role in the process of policy making and implementation. In this paper, we describe how the motivation models support the description of desirable outcomes and help to develop relevant high-level goals in policy making in particularly complex areas. We give two example domains where we have developed agent-based models. The first is data management policy for university researchers. The second one focuses on sustainable households, and how to provide relevant guidance for educating people and helping them to understand how to behave in a more environmentally friendly manner. Our two examples demonstrate that the agent-based models are able to help to ask the right questions for coming up with relevant quality goals and identifying the right stakeholders in these two multifaceted and abstract domains. We aim to enable people involved in policy making to focus on and understand the relevant goals, quality goals and activities in order to formulate effective and supportive policies that can accomplish the intended outcomes.

Keywords: policy making, agent-based modelling**,** motivation models, high-level goals.

1 Introduction

We are using agent-oriented models from software engineering (AOSE) differently from goal models in the past. In former projects we have used the AOSE models for eliciting socially-oriented requirements and associated qualities specifically for the development of socio-technical systems [1, 20]. We concluded that high-level goal models are well suited as an initial basis for shared understanding independent from a specific implementation. These models represent the important characteristics of a domain. This can be really useful when we try to get a handle on very complex topics, as agents help people to think and focus on relevant aspects. When making policies many stakeholders should be engaged to come up with processes that provide real guidance to users. Note also that we are not developing the models with an insistence that the implemented system be agent-based, which is the case with methodologies such as Prometheus [16].

One area we have investigated is policy making within *eResearch*, particularly with respect to research data management. The Australian Research Council has placed a demand on researchers to keep data from funded research for seven years. Universities and governments have an increasing need to act in a compliant way to these demands. Typically, policy documents and processes to ensure such behaviour from researchers are hard to understand, not embedded in existing practice and often very confusing for the reader. An example of an institution grappling with the issues is the University of Melbourne (http://www.unimelb.edu.au/records/manual.html). We suggest that AOSE models can help to involve multiple stakeholders in developing relevant goals for data management policy that are connected to how people are actually conducting research.

Here, we are interested in high-level goals for policy making independent from specific technical solutions. We aim to direct the focus on relevant outcomes and necessary activities to achieve these outcomes before we think about technical solutions that might support these kinds of outcomes. For example, in an intergenerational relationship maintained over distance, goals such as *playing* and *gifting* and associated qualities goals such as *showing presence* and *share fun* are high-level goals. There are still many ways as to how this can be supported, but the social relationship stands at the centre and plays a crucial part that is independent of the technologies implemented [e.g. 17]. Our credo is that when we want to influence people to show desirable behaviour we have to understand what is truly relevant for these people within the respective domain and make suggestions for regulations from there.

Agent-oriented models are suitable for modelling the social domain because they represent the goals and motivations of roles and individuals, and quality goals can be used to discuss high-level outcomes relevant for policy making such as *providing data access* or *saving water*. Furthermore, the domains we investigate are truly socio-technical, and agent models allow us to represent human behaviour as well as software system behaviour. Our process leads to the development of high-level quality goals that are shared by everyone, but can be substantiated and adapted for an individual context in a meaningful way.

We use the construct of quality goals attached to functional goals as a way of representing quality attributes of socio-technical systems. Quality goals are essentially non-functional and are designed to encapsulate aspects of the context into discussions. Garcia and Medinilla [6] describe high-level quality goals as a specific form of uncertainty that can be used as a descriptive complexity reduction mechanism and to model and discuss uncertainties in the environment. High-level goals associated with activities can act as a point of reference for discussing the usefulness of alternative activities to achieve these goals. Instead of using the agent-based models in requirements elicitation for the development of a system we use them as shared artefacts for discussion [15] in the process of developing a shared understanding that can be used for policy formulation and implementation. The multi-agent paradigm offers benefits over other paradigms because the concepts used in modelling, such as roles, goals, and interactions, are part of everyday language and make it accessible for different stakeholders [18].

Here, we give two examples in which on the one hand policy making plays a major role as a larger institution or government has the need to reach certain predictable outcomes, but on the other hand relations are complex and the individuals being expected to follow these policies are in very different situations. Therefore, people need clear guidance and a good understanding of these relations in order to be motivated to follow policies. The first example is from policy making in data management and the second example is to encourage sustainable behaviour at home to fulfil long-term environmental goals. Before these examples are described we provide the foundations of our approach and important definitions.

2 An agent-Based Process of Policy Formulation

Our process builds on the work of Sterling and Taveter [20]. Their work has focused on how to make high-level AOSE models palatable in design discussions. They define goal and role models that build part of a motivation layer. An agent is actively situated in an environment and is assumed as being purposeful in this environment. The models of goals and roles refer to knowledge about the problem domain. At the motivation layer, such knowledge is represented as a set of domain entities and relationships between them. A *goal* can be defined as *"a situation description that refers to the intended state of the environment"* (p. 30). Goals are based on motives and can have sub-goals. A *quality goal* is a non-functional or quality requirement of a socio-technical system.

We aim to come up with a process that helps us to find out how conversations and policy making supported by agent-based diagrams is done best to involve multiple stakeholders e.g. the brainstorming of goals, roles and activities to develop a shared understanding. In this regard, we want to learn more about a repeatable process and not only the outcome. Here, we propose to use motivational models, roles and responsibilities as an easy way to represent the complex relations that are subject of regulations and long-term goals for larger communities such as researchers or citizens. AOSE models are very suitable as they

- are a good way to represent complex topics on a role and goal diagram level,
- provide a good overview which people (should) have which roles, behaviour, and attitudes depending on their current situation,
- help us to find out what we need to focus on when discussing complex topics such as sustainability.

We suggest several levels for using agent-based diagrams. Some of them are more general (level one and level two) and some are project specific (level three and level four). Starting with general high-level goals helps us to focus on the motivations for different roles. From there we can narrow our attention to the specific context of individuals such as families aiming for a more sustainable life or researchers with a certain research project.

The different levels or steps include:

1) High-level motivational goal model (goals, quality goals and roles)
2) More detailed roles – responsibilities and constraints
3) Agent types for a specific area describing actual activities in more detail
4) Activity plan – specific policy, regulations or guidelines

The next sections describe and define the different levels and the procedure of using agent-based models in more detail.

2.1 High-Level Goals and Quality Goals

Our starting point is a simple model of motivations of a socio-technical system including goals and quality goals. By capturing and representing quality goals in AOSE models we make a commitment to important aspects of socio-technical systems. By externalizing them in a simple format the models become shared artifacts [15] that are able to sustain multiple interpretations across disciplines. Quality goals allow a focus on understanding the reasons why people do things, or the essence of an attitude rather than describing a concrete action. In doing so, quality goals capture something that is more dynamic and fluid than other mechanisms found in usual software engineering practices. Non-functional goals do not generally have a direct relationship with functional goals [2]. In our approach there is a direct pairing between system goals and quality goals. Relating an abstract and unresolved quality attribute to a system goal enables a focus on social goals within the process of policy making.

In order to create shared AOSE models we use a straightforward syntax and semantics. Goal models are useful at early stages of requirements analysis to arrive at a shared understanding [7, 12, 13]; and the agent metaphor is useful as it is able to represent the concepts that we want to capture for socio-technical systems, such as agents taking on roles associated with goals. These goals include quality attributes that are represented in a high-level pictorial view used to inform and gather input from stakeholders. In Sterling and Taveter's notation [20], goals are represented as parallelograms, quality goals are clouds, and roles are stick figures. These constructs can be connected using arcs, which indicate relationships between them (see Figure 1).

2.2 Roles with Responsibilities and Constraints

Sterling and Taveter [20] define a *role* as some capacity or position that facilitates the system to achieve its goals. In their view, roles express functions, expectations and obligations of agents enacting them. They encompass these senses in the term *responsibilities*, which determine what an agent or set of agents enacting the role must do in order for a set of goals and quality goals to be achieved. In addition, a role may also have some *constraints* specifying conditions that the role must take into consideration when performing its responsibilities.

2.3 Agents and Activities

An *agent* is an entity that can act in the environment, perceive events, and reason. Reasoning means drawing inferences appropriate to the situation. Events that an agent perceives are caused by agents or other entities in the environment. Conversely, through acting, agents can affect entities in the environment. Agents can be humans as well as specialised hardware or software such as sensors.

3 Example (1) Data Management Policy

Data management is a socio-technical problem that we want to tackle with the help of the AOSE models to support policy making at universities. Motivational diagrams from AOSE are a good way to represent the challenges of data management because they provide a good overview which people have which roles. We conducted four rounds discussions of the models with different stakeholders involved with data management at two different universities; two discussions with people from the IT department who are building the infrastructure for long-term data storage, and two with librarians who are in the process of coming up with a data management policy at another university. Each discussion took approximately 1.5 hours. We had two aims for these discussions:

- Firstly, how does the data workflow of a project look like when using an agent-oriented view on data management? This also raised questions of what to do in specific project situations (e.g. the main researcher leaves after a year, one collaborating researcher is in another country, who has the responsibility for secure data storage?).
- Secondly, what is the actual process? What are we doing with 20 GB of mixed project data and is a policy able to support this process?

We have developed agent-oriented models, described in the following sections.

3.1 Challenges for Research Data Management

There are policies in place that regulate the correct handling of research data during and after a research project has finished. These policies often do not consider the individual situation of the researcher or provide the necessary infrastructure to be able to behave according to the policy. Questions that arise are: Where can I store a large amount of data? Who is responsible for it when the main researcher has left the institution? How can older data be retrieved and who is allowed to access them? These and many other practical questions arise. It becomes even more complex when the national research council is interested in granting access to former data to researchers nationwide as links between grants & data and publications & grants are needed.

The models were created based on the experience of the researchers with research projects and with one research project used as a case study. The project involved the following kind of data: fieldwork notes (handwritten and digital), interview data (MP3s), observations of humans, designed objects (diaries), photographs (digital) and other images (non-digital), analysis and coding of texts, case studies, software and code, audio and video recordings, and consent documents with signatures. All together the data that needed to be managed were 20 GB and were collected over a period of four years. Additionally, data files are in different formats and there are multiple digital files, organised into a folder structure that must be retained.

In an iterative process the models were discussed with stakeholders currently involved in data management and the creation of data management policies. After every discussion the models were updated. Figure 1 and figure 2 and show the version after the first two discussions with the system supporters from the first university. These figures represent level one and level two described in section 2.

3.2 High-Level Motivation Model

High-level goals for conducting research are *collecting data*, *analysing data*, *managing data* and *writing data* (figure 1). It is important to include, for example, the data collection as this has a major impact on what kind of data and data formats need to be managed later on. The quality goals provide information about the attributes of the research data. For example it is crucial that during the whole process of dealing with the research data, they are handled in a *secure* manner as it might be sensitive data and participants should not be recognisable.

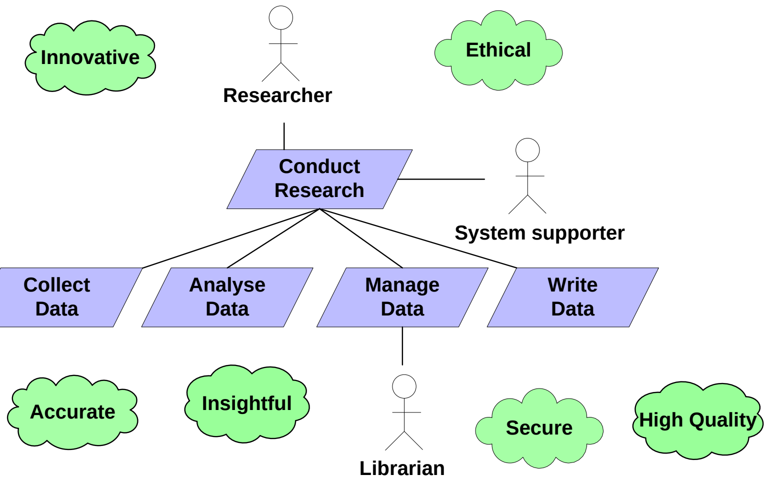
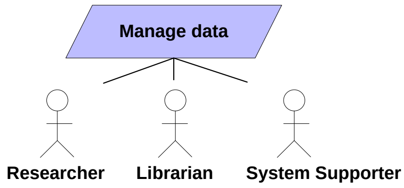


Fig. 1. High-level motivation model for data management.

Besides the *researcher*, the *system supporter* and the *librarian* will handle the data. The librarians we spoke to put a high emphasis on the quality attribute *insightful* and *accurate*. From their experience they know that it is merely useful to store data when these can be retrieved and sighted in an easy manner by a person interested in the data. The system supporter is responsible for providing access rights that are in accordance with *ethics* and the sensitivity of the data.

3.3 Responsibilities and Constraints

Level 2 (figure2) looks in more detail in the different roles, their responsibilities and constraints when specifically managing data.



N (Name)	Researcher+	Librarian	System Supporter
D (Description)	*Generates raw data*	*Organises data*	*Provides technical infrastructure*
R (Responsibilities)	Provide Data	Curate Data	Provide Storage
	Access Data	Archive Data	Secure Data (incl. protect)
	Verify Data	Add Metadata	Provide Access
	Update Data	Retrieve Data	Implement Policy
	Store Data	Monitor Policy	Monitor Policy
			Delete Data
C (Constraints)	Behave Ethically	Meet Data Standards	Be Consistent with Policy
		Needs access	

Fig. 2. Responsibilities and constraints for different roles.

3.4 Agent-Based Activities

This level (table 1) describes in more detail the agent-based activities for the different roles. The role of the researcher is further detailed in more specific agent-types such as *research fellow*, *student researcher*, *chief investigator* and *collaborator*.

Table 1. Agent-based activities for different roles

Research Fellow: Access data, provide information about data, retrieve data, copy data, upload data, structure data and locate data.
Student researcher: Access data.
Chief Investigator: Provide information about data management, remove data
Collaborator: Access data.
Librarian: Move data, access records (where to find data, contact resp. researcher etc.), track data/records.
System Support: Destruct data; track data/records, link data, provide log in, check access rights, register data, manage access, provide space/limit space, retrieve password, provide instructions about data, store/retain data, monitor time lines, check password.

3.5 Specific Data Management of One Project

After two discussions with the IT system supporters, we were able to agree on the data management structure for our specific project. This project served as an example for data management of all research projects implemented at the university. This structure includes a workspace to keep data of 20 GB for data storage and data access. Data can be pre-structured according to needs and the researcher is able to upload the data to a central server where she can access it herself, as well as manage access for international collaboration needs. This part of the structure is in place and has already been provided before the research project had finished. The data management structure will be extended to more permanently store data, enable the creation of data permalinks and association with relevant software as well as links to non-digital data (consent forms, etc.). The extended structure will provide fine grain control over private access, public availability, anonymisation of data, and security. The AOSE models helped to discuss this structure and focus on the most relevant needs for data management of our project.

3.6 Feedback and Evaluation

We used the models for discussion in another university so as to receive feedback and evaluate the models regarding their generalisability for the purpose of data management at different universities. This time our discussion partners were librarians with the task of coming up with a data management policy. Similar to when we were discussing the models with system supporters before, we were interested if our high-level models in particular are able to reflect the goals the librarians had in mind for creating a data management policy. After two more rounds of discussions we only made slight changes to our high level goal model and to the responsibilities of the role of a librarian.

The librarians saw a clear distinction in making decisions about, for example, deleting data and creating links between data and implementing these changes in a technical system. Therefore, we added these aspects of *decision making* to the responsibilities and agent activities. Another aspect that was very important to the librarians was that researchers could provide them in an efficient way with the information about the data to be managed in order to save time. This includes to pre-populate forms with researcher's details and only to ask relevant questions tailored to the first entries of an online form for capturing research data. At the moment the librarians take a lot of time to retrieve information – up to one hour interviews with researchers are needed to structure the data in a suitable manner. Therefore we added *efficient* as a quality goal. A further problem in capturing the research data is that there are research areas that are so specific that it is hard to articulate even the right questions to get a suitable data structure implemented. In these research areas the librarians depend on close collaborations with researchers. They need to be provided with the right keywords so that data can be recorded and retrieved later in a useful manner. Interestingly the librarians also spoke about a needed shift in the perception of relevance of data management. If easy retrievable data was something that could add to the prestige of a university and their researchers then it would be more valued and the attitude towards data management would change. Research prestige as a consequence of carefully conducted data management is a really important outcome that also was then captured as quality in the goal model (*prestigious*) in the high-level goal model.

The AOSE models helped us to ask the right kind of questions when talking to different stakeholders and to focus on relevant goals and their associated qualities. The different models enabled us to add new information - depending on its level of detail - to activities, responsibilities, high-level goal or associated quality. Additionally the models helped to uncover new goals directly relevant to policy formulation for data management.

The discussion around the models made it possible to find the natural points in the researchers' workflow or the life of a project for the formulation and implementation of data management policies. That means that the researchers do not feel that additional organisational work is created, but that they can already profit during the research project from the provided data management structure and storage provision. One important metaphor that came up during one of the discussions was: How do you get people to wear a lab coat? – You do not place the hook for it at the exit.

4 Example (2) Sustainable Households

4.1 Challenges for Guidelines on Sustainable Domestic Behaviour

The whole area of sustainability is even more abstract and complex than the previous example on data management. How does a family motivated to live a sustainable lifestyle know what kind of sustainable behaviour realistically can be expected from them as part of a wider community? And on the other hand, how can a government that is interested in citizens behaving sustainably [3, 21] educate and encourage the right activities? One approach to support people in behaving sustainably is to showcase increasingly available applications and devices for e.g. monitoring energy consumption [4, 5, 11]. These only help when people understand what they are aiming for, have the right infrastructure in place and get more individualised feedback [8]. In addition, monitors measure against a statistical average and people that are below this average often feel encouraged to use the resources that "they are entitled to" [4]. Competitiveness can lead to saving of resources, but the question is if people are not more successful if the main goal is *living sustainably* instead of comparing oneself to one's neighbour. There is also a plethora of publications for environmentally friendly behaviour available for families. Most of these publications explain the need for sustainable behaviour and give concrete advice how to save, for example, energy and water at home. While all of this advice is useful and successful to some extent, it does not take into account the individual situation of different households, the climate, and personal preferences. Therefore, it is difficult for the individual to decide which actions are effective. When dealing with a complex area such as sustainability we would like concrete and simple advice. Yet, if the advice is too simplified it lacks relevance for the single household and its specific socio-economic situation. Reasons that make the domain for *sustainable behaviour* so challenging are that:

- it is a very complex topic (e.g. "green coal"),
- it is a value loaded topic ("if you don't do this you don't care about…"),
- the topic includes many perspectives and sometimes controversial advice,
- results are hardly visible and in the overall context only a minimal contribution to a large goal,
- there is a complex relation between water, food, energy and waste,
- it is a very fashionable topic and there is a danger of people "getting over it".

Consequently we need to

- give good examples and explanations,
- allow a range of opinions and approaches,
- define the right or overarching goals,
- set small personal mile stones,
- look at aspects of sustainability in a dependent way,
- work against inertia and the feeling of helplessness.

Other specifics we have to consider when we focus within the large topic of sustainability on utility use in the home: firstly, set targets have to be supported by all people living in one home. Secondly there is a different level of insight into the topic (e.g. children might not understand the need straight away). Finally, non-house owners have not the possibility to make certain decisions on sustainability. This means we have to consider several stakeholders within one home.

Again we see a role for the AOSE models in mapping the roles and responsibilities to overcome some of these challenges. The diagrams presented in the following sections show the different levels of abstraction: the first two diagrams are specific to sustainability – here we aim to keep a light touch and stick to more general descriptions valid for all homes. The two last levels are specific to one area/utility. The models are translated into specific activities and a management plan for family homes with specific characteristics.

The models in this second example are based on a body of literature on sustainability accessible to the general public in libraries. The models were created as a condensed version of the main and overlapping advice found in this literature – some of them containing 101 tips for sustainable living [e.g. 9, 10]. These tips differ largely in their effectiveness and their costs. For example, one book on water management advised on the same page "to cook vegetables in the microwave to save water" and "installing a rain water tank" for the same purpose [10]. If the high-level goal is *save water* then the latter advice is certainly more efficient unless it never rains in the region the rain water tank was installed. We suggest using the models for teaching people how to accomplish best high-level goals in utility management at home in accordance with the individual living situation.

4.2 High-level motivation model

Figure 3 shows the high-level goal model for sustainable utility management at home. Again the model consists of goals, quality goals and roles. In this example the quality goals take into consideration that while people want to live sustainable they are often not able to spend too much time and money on it (*manageable*). The quality aspect *innovative* encourages people to look actively for the latest solution and to keep up-to-date with sustainable developments. The different roles *decision maker* and *habitant* take into account that not every habitant in a home can make decisions on sustainability (e.g. children and people renting). The *rule maker* is an official body such as the city council or the regional government making rules on e.g. garden watering or recycling.

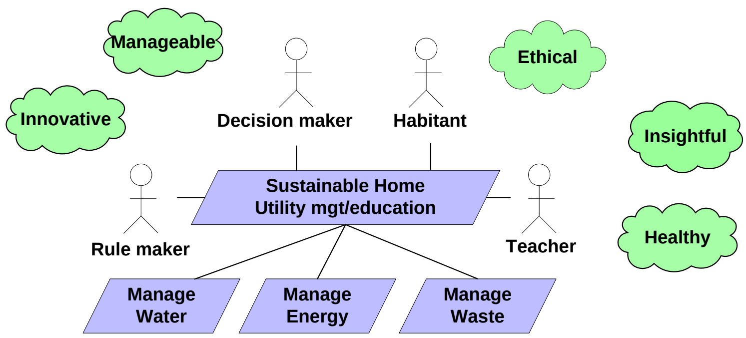


Fig. 3. High-level motivation model for sustainable behaviour.

This motivation model breaks down the high-level goals into sub-goals. Here we show the sub-goals for *manage energy* and *manage waste* (figure 4 and figure 5). The advantage of using a hierarchy of goal models is that no single model contains too much information. The high-level goals describe general activities such as *turn off* and *insulate* people can think through for their specific situation at home.

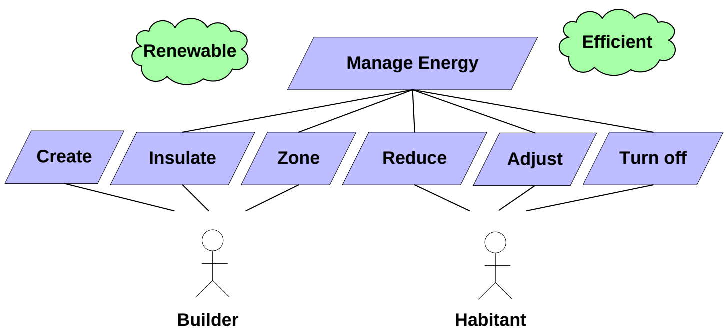


Fig. 4. Goal model specifically for *managing energy* at home

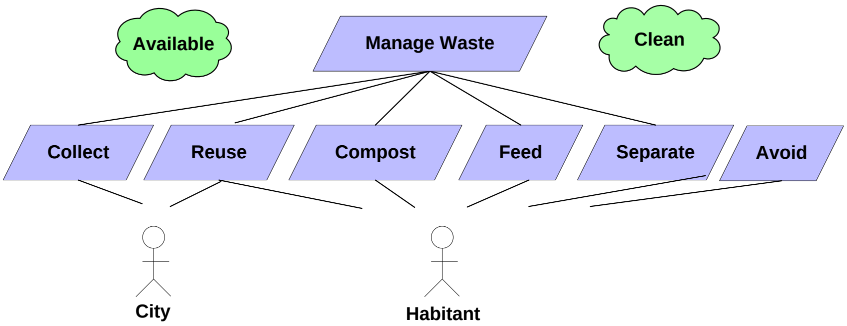
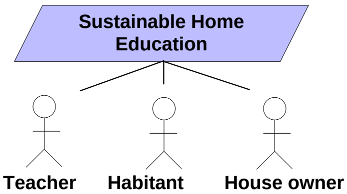


Fig. 5. Goal model specifically for *managing waste* at home

4.3 Responsibilities and constraints

Figure 6 describes the responsibilities and constraints for the different roles involved in educating about sustainable behaviour.



N (Name)	**Teacher**	**Habitant**	**House owner / Decision maker**
D (Description)	*Teaches other agents sustainable behavior*	*Lives in house; Learns about sustainability and acts on it*	*Owns house. Makes informed decisions and enables habitant to act sustainable*
R (Responsibilities)	Provide examples	Get informed	Purchase sustainable products
	Provide strategies	Match info & home	Provide sustainable infrastructure
	Explain relations	Change behavior	Set home rules
	Set values	Weigh alternatives	Monitor household members
	Match info & home		Update all of above
C (Constraints)	Good/correct examples	Needs infrastructure	Cost - benefit analysis
	Be up-to-date	Needs knowledge	Be Consistent with rules
	Behave Ethically	Activity needs to be safe	Access to policies and knowledge
	Maintain top level view	Approach teacher	
	Contact with other agents		

Fig. 6. Responsibilities and constraints or sustainable behaviour at home.

4.4 Agent-based activities for *energy management*

Specific agent types and actual activities for *energy management* are shown in table 2.

Table 2. Agent-based activities for different roles

Teacher: Teach knowledge about alternative energies (solar, wind, thermal…) and energy saving (e.g. insulation), give examples on consequences (…), provide best practice examples, communicate rules and policies, and discuss compromises.
Decision maker: use and subscribe to renewable and cleaner (non-carbon dioxide producers), insulate (seal windows, doors, roof), buy energy efficient appliances (fridge & freezer, washing machine).
Consumer/habitant: avoid fossil fuels, travel (travel together, walk, car pool, use public transport, cycle), eat and buy local products, save energy being energy smart (turn heat up early, turn down 1% to save 10%, switch off all lights, stand-by & computer).

4.5 Specific Energy Management Plan for One Household

Here we describe briefly the specific energy management plan for one household that is located in rather hot climate, is built with a lot of open areas and has single pane windows. This energy management plan is based on the high-level goal model from figure 4.

- Switch to hot water system with solar gas boost (eligible to governmental rebates)
- Secondary glazing on windows for insulation
- Draft proofing (windows, doors, seal garage/office, self-sealing exhaust fans)
- Compartmentalize rooms so heating/cooling is minimized (retrofit zoning on gas heating)
- Switch devices off during night (stand-by), when not used (computer, lights)
- Use winter/summer settings on fans (pushing air up or down depending on desired effect and temperature).

5 Conclusions

The role of the goal models is not simply the typical formal process of modelling to lead to the development of a system as in the traditional domain of software engineering. For us, they have become a way to think through problems, and to reach agreements. However, a body of literature that looks at software engineering from a social science perspective recognises that models and other documentation in software engineering have been used for a long time as a way to think through problems, to reach agreements, and to elaborate the needs of stakeholders in a different way than simply feeding into a formal process of modelling for system design [14, 19]. In this sense it is not completely novel to use models as tools that are not directly connected to the development of a system. We use AOSE models to facilitate discussions around complex socio-technical systems.

Agent-based models can play an important role in the process of policy making and implementation. The models helped us in discussions with externalising and making explicit the perspectives of different stakeholders on data management. In particular the quality goals helped to explore different perspectives in a distinguishable manner.

Sustainability is the result of different behaviours based on multiple perspectives, varying sometimes contradictive knowledge and social values that needs to find a balance. In short, it is very complex and even though a lot of people aim for a more sustainable lifestyle it is difficult for them to find the more efficient and cheapest way to do so. Our discussions demonstrate that the AOSE models are easier to read than process descriptions and focus better on relevant aspects. The high-level goal model needs to be consistent with the actual activities to ensure desirable outcomes. We are aiming for a match between the goal models and peoples' behaviour. Our discussions helped us to include every stakeholder's perspective and include this perspective into the description of responsibilities and constraints. We also include qualities such as affordability (that is how *manageable* can be interpreted) of environmentally friendly solutions that is crucial for families but often not directly linked to the discussion of sustainability. This gives us a more realistic account if people can and will adapt their behaviour to live more environmentally friendly.

We are planning to present the models for sustainability to educators to see if they can be embedded in educational material on sustainability.

Acknowledgments. We would like to thank all people involved in our discussion workshops giving us valuable insights into their daily life and work and providing us generously with critical feedback on our AOSE models for policy formulations and implementation.

References

1. Boettcher, A.: Moving From Cultural Probes to Agent-Oriented Requirements Engineering", OZCHI 2006, pp. 20--24, (2006)
2. Chung L.K., Nixon B.A., Yu E., Mylopoulos J.: Non-Functional Requirements in Software Engineering. Kluwer Publishing, (2000)
3. Department of Sustainability and Environment. Annual Report 2006. http://www.dse.vic.gov.au/CA256F310024B628/0/C4C97FE3979E7A88CA2575C4000A6869/$File/Energy+Efficiency+Action+Plan+2006.pdf
4. Fischer, C.: Feedback on household electricity consumption: a tool for saving energy? Energy Efficiency, vol. 1, 1, pp. 79--104 (2008)
5. Froehlich, J.: The Design of Eco-Feedback Technology. In Proceedings of the CHI (Atlanta, GA, USA, 10-15 April, 2010), ACM Press (2010)
6. Garcia A., Medinilla, N.: The ambiguity criterion in software design. In International Workshop on Living with Uncertainties. ACM (2007)
7. Guizzardi R., Perini, A.: Analyzing requirements of knowledge management systems with the support of agent organizations. Journal of the Brazilian Computer Society (JBCS)- Special Issue on Agents Organizations, 11(1), pp. 51-- 62 (2005)

8. He, H. A., Greenberg, S., Huang, E. M.: One size does not fit all: Applying the Transtheoretical Model to Energy Feedback Technology Design. Department of Computer Science, University of Calgary, Calgary, Alberta, Canada (2009)
9. Healey J. (editor): Our energy future. Issues in Society, 295. Thirroul, N.S.W. Spinney Press (2009)
10. Healey J. (editor): Water management. Issues in Society, 288. Thirroul, N.S.W. Spinney Press (2009)
11. Holmes, T.: Eco-visualization: Combining Art and Technology to Reduce Energy Consumption. In Proceedings of the C&C (Washington, DC, USA, 13-15 June, 2007), ACM Press (2007)
12. Iqbal, R., James, J., Gatward, R.: Designing with ethnography: An integrative approach to CSCW design. Advanced Engineering Informatics, vol.19, pp. 81--92 (2005)
13. Jureta I. J., Faulkner S.: Clarifying goal models. Proc. ER 2007, pp.139--144 (2007)
14. MacLean, A., Bellotti, V., Young, R. M.: What rationale is there in design? In D. Diaper, D. J. Gilmore, G. Cockton, and B. Shackel, editors, Proceedings of the IFIP TC13 Third International Conference on Human-Computer Interaction, pp. 207 -- 212 (1990)
15. Paay, J., Sterling, L., Vetere, F., Howard, S., Boettcher, A.: Engineering the Social: The Role of Shared Artifacts, IJHCS, vol. 67(5), pp. 437-- 454 (2009)
16. Padgham, L., Winikoff, M.: Developing Intelligent Agent Systems: A practical guide, Wiley (2004)
17. Pedell, S., Miller T., Vetere, F. et al. Having fun at home: interleaving fieldwork and goal models, Proc. of OZCHI 2009, pp. 309--312 (2009)
18. Pedell, S., Vetere, F., Howard, S. Miller T., Sterling, L.: Shared artefacts as participatory Babelfish. Proceedings of PDC 2010, pp. 167--170. ACM (2010)
19. Randall, D., Hughes, J., Shapir, D.: Steps toward a partnership: ethnography and system design. In M. Jirotka and J. Goguen, editors, Requirements Engineering: Social and Technical Issues, pp. 241-- 254. Academic Press (1994)
20. Sterling, L., Taveter, K.: The Art of Agent-Oriented Modelling. MIT Press (2009)
21. The Australian Government’s sustainability portal: www.livinggreener.gov.au

Leveraging Multiple Influence Mechanisms for Information Propagation

Andrew W. Wicker and Jon Doyle

Department of Computer Science,
North Carolina State University,
Raleigh, NC 27695-8206
{AWWicker,Jon_Doyle}@ncsu.edu

Abstract. We address the problem of how social influence affects the spread of information across a population of individuals. Most extant models have approached such problems through the use of simple models of influence that utilize a single influence mechanism for inducing changes in a population of individuals. We here present a new model of social influence that recognizes and leverages multiple influence mechanisms and multiple types of relations among individuals. Our model substantially increases expressivity and extensibility over that of existing related models and facilitates analysis of influence effects in a multitude of social contexts (e.g., marketing, trends, decision support, computer security).

Keywords: social networks, influence, information propagation

1 Introduction

We address the problem of how social influence affects information propagation in a population. The study of population dynamics, in particular, the propagation of information throughout a population, has been an increasingly active area of research. There are a variety of different models that have been used to study such problems as information propagation. Researchers are quick to point out the simplicity and computational ease of extant models of information propagation, but often do not provide any evidence from the social psychology literature to support their formulations.

We have developed a new model for studying information propagation in a population. Our model leverages multiple influence mechanisms for spreading information between individuals and multiple types of relations connecting individuals, which reflects more accurately the many mechanisms for transmitting information between individuals in the many social networks in which individuals exist.

The primary results we present pertain to how the topological structure of a population impacts solutions to the influence maximization problem [10] when studied within the formal framework of our model. By gaining insight into how a variety of influence mechanisms affects the evolving mental attitudes of individuals, we are able to make more informed decisions about what individuals we should target for initial influence.

1.1 Motivating Example

Consider a set of eight individuals that are labeled $\{m_1, \ldots, m_8\}$. Each of these individuals may be related to another through one of two relations: *is-coworker-of* denoted by $r_{coworker}$ and *is-friend-of* denoted by r_{friend}. For this example, assume that coworkers tend to be influenced by an authoritarian influence mechanism (e.g., "accept this because I said so"), whereas friends tend to be influenced by group conformity influence mechanisms since they desire acceptance by their friends (e.g., "all of my friends play musical instruments, so I will too"). We denote these influence mechanisms by h_{auth} and h_{conf}, respectively.

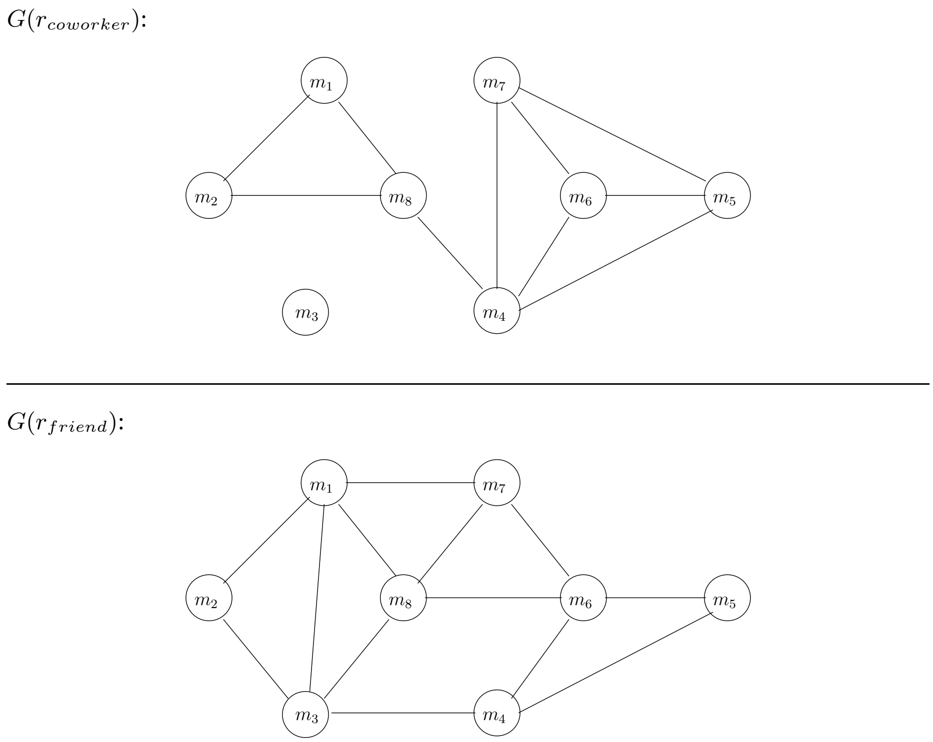


Fig. 1. Graphs $G(r_{coworker})$ and $G(r_{friend})$ representing two different (symmetric) relations *is-coworker-of* and *is-friend-of*, respectively, over the same set of individuals. Using both relations with one or more influence mechanisms may produce very different results than using only a single relation.

Figure 1 depicts the graph representations of relations $r_{coworker}$ and r_{friend}. These relations are assumed to be binary and symmetric, but this may not be the case in general. Notice that many individuals share both relations, whereas others are related only through a single relation. Such graphs may consist of more than one component, as in graph $G(r_{coworker})$.

If we use only a single influence mechanism, then we are limited to influence only within a single graph representing the relation for which that influence mechanism is applicable. For example, assume that individuals in $\{m_5, m_6, m_7\}$ share an equivalent individual state that is different from the rest of the population and exert influence on m_4 using the authoritarian influence mechanism h_{auth}. If these individuals are unable to influence m_4 (e.g., none command sufficient authority), then the remaining subpopulation has no chance of being influenced to adopt the individual state shared by those in $\{m_5, m_6, m_7\}$.

The use of both mechanisms h_{auth} and h_{conf} may increase the potential for influencing the remaining population. For example, individuals m_5 and m_6 may be influenced by m_7 using h_{auth} in $G(r_{coworker})$. Individual m_4 may subsequently be influenced by m_5, using h_{conf} in $G(r_{friend})$. Next, individuals m_4 and m_7 may influence m_1, m_8, and m_3 using h_{conf} in $G(r_{friend})$. Finally, m_8 may influence m_2 using h_{auth} in $G(r_{coworker})$. This simplified example demonstrates the significant impact of multiple influence mechanisms and relation types on the potential for influencing individuals in a population.

Restriction to a single relation and a single influence mechanism may produce very misleading results. Using only relation $r_{coworker}$ (depicted in $G(r_{coworker})$), we may conclude that individuals on the cut-edge (m_8, m_4) must necessarily be influenced in order to influence more than half of the population. Moreover, we may conclude that m_1 is unable to exert influence on m_7, or that m_3 is unable to exert influence on any other individual (or be influenced itself) since it is disconnected in $G(r_{coworker})$. However, the population structure and potential for exerting influence changes significantly with the inclusion of relation r_{friend} and mechanism h_{conf}. With both mechanisms and relations, there are numerous potential pathways for propagating the effects of influence.

In the remainder of this paper, we give an overview of our model that leverages multiple influence mechanisms and social relations, and support our claim that it represents an improvement over extant models. We focus on problems of maximizing the influence of sets of individuals that have been targeted for influence and how the topology of a population impacts the selection of such targeted individuals. Due to space constraints, we omit proofs for our theoretical results. We begin with an overview of related work.

2 Related Work

Research on influence models is a multidisciplinary pursuit and an increasingly active area of research. The fields of sociology, economics, physics, mathematics, biology, and computer science have each contributed to modeling the process and effects of influence. Despite being conceived out of tangential fields, these various influence models have much in common.

The general voter model (see [5, 7]) is one of the earliest formal models of influence and is often cited in later developments of influence models. The primary appeal of the voter model is mathematical simplicity. However, there are many simplifying assumptions made in the voter model that are inconsistent with reality. The voter model captures the effects of only a single influence mechanism, requiring that a single random individual completely adopts the full opinion of a random neighbor. Although some situations do conform to this all-or-nothing adoption, it does not translate well into other models of influence, such as on preferences or beliefs. For example, an individual may revise its beliefs to become more similar to its neighbor without adopting all of its neighbor's beliefs.

Another related influence model is Asavathiratham's Influence Model [1] that was developed primarily to model the influences among the various entities in a power grid. This model of influence is quite similar to our model in that it uses Markov chains as the underlying probabilistic state transition model, but contains some distinct differences.

The network of connections among the entities captures only a single, semantically vague notion of relation. That is, entities are only viewed as being related or not. There is no notion of different types of relations, except from what is expressed in the edge weights. Moreover, only a single influence mechanism is used in the Influence Model. This influence mechanism gives a probability of adopting the state of one's neighbors as a weighted sum of the entities that are in a particular state along with the edge weight connecting the two entities. Such an influence mechanism is derived from the mechanism used in the related voter models.

Cascade models [10, 11, 3, 14] have received a lot of attention recently from those developing influence or information diffusion models. We here focus on what is referred to as the *independent cascade model* since it was used in [10] to study the influence maximization problem. There are some obvious shortcomings of the independent cascade model as a model of influence, but for our purposes it suffices to observe that they do not use multiple influence mechanisms. By using only a weighted sum of active neighbors, these cascade models are limited to a single type of influence mechanism.

3 Individuals and Populations

Social influence is not independent of the relations maintained between individuals (see [12]). For example, the mechanism(s) by which a mother influences her child (relation being *is-mother-of*) may be very different from the mechanism(s) by which a Twitter user influences a follower (relation being *is-twitter-follower-of*). Such evidence of differences in the mechanisms of influence across different relations accentuates the need for a formal statement of these relations held among individuals in a population.

We assume a finite set of individuals $\mathcal{M}$, which we refer to as the *population*. An *individual* is denoted by $m \in \mathcal{M}$ and a subset of individuals, or *sub-population*, by $M \subseteq \mathcal{M}$. We assume an enumeration of individuals in the population given by $\langle \mathcal{M} \rangle = \langle m_1, m_2, \ldots, m_{|\mathcal{M}|} \rangle$.

Let $\mathcal{R}$ be a finite set of relations over $\mathcal{M}$. We write $r \in \mathcal{R}$ for a particular type of relation over $\mathcal{M}$ and $R \subseteq \mathcal{R}$ for any subset of relations defined such that $R = \bigcup_{r \in R} r$. The symmetric closure of a relation r is given by $r^{\leftrightarrow} = \{(m, m'), (m', m) \in \mathcal{M}^2 \mid (m, m') \in r \vee (m', m) \in r\}$. The transitive closure of a relation r is given by $r^{+} = \{(m, m') \in \mathcal{M}^2 \mid \text{there exists a path from } m \text{ to } m' \text{ in } r\}$. We write $r^{\leftrightarrow +} = (r^{\leftrightarrow})^{+}$ for the relation that is obtained from r by first symmetrizing r and then taking its transitive closure.

We do not require that all relations $r \in \mathcal{R}$ satisfy the same properties. For example, *is-a-family-member-of* is a symmetric binary relation, but *is-boss-of* is an asymmetric binary relation. We assume that all relations $r \in \mathcal{R}$ are static. That is, the relations do not vary over time. We may remove this restriction as part of future work to consider dynamic relations, but here restrict our work to static relations. We restrict $\mathcal{R}$ to *finitary relations*, or k-ary relations for which $k \in \mathbb{N}^+$. In the remainder of this paper, however, we deal only with binary (i.e., $k = 2$) relations in order to simplify the exposition.

Let $\mathcal{G} = \{G(r_1), \ldots, G(r_{|\mathcal{R}|})\}$ be the set of graphs representing binary relations $r_i \in \mathcal{R}$. Each $(m, m') \in r_i$ is a directed edge from m to m'. The graph properties are inherited from the relations that are represented by each specific graph. We write $G(R)$ for the graph representing the union of each $r \in R$ such that there is an edge (m, m') in $G(R)$ if and only if $(m, m') \in r$ for some $r \in R$.

4 Probabilistic State Dynamics

In this section, we describe Markov chains, which form the underlying probabilistic state transitions in our model. We introduce individual state transition probability measures for expressing the probabilistic individual state transitions and build upon this with population state transition probability measures.

4.1 States

Any model of social influence must specify what it is that changes under the exertion of influence. As in most related models, we view each individual as having a state, or mental attitude, that constitutes the material subject to the forces of influence.

Let Ω_M be the finite set of *population states* for individuals $M \subseteq \mathcal{M}$. If $M \subset \mathcal{M}$, then we sometimes refer to Ω_M as *sub-population states*. We write $\Omega = \bigcup_{M \subseteq \mathcal{M}} \Omega_M$ for the set of all states over all sets of individuals (including singletons $\{m\}$), where $\Omega_\emptyset = \emptyset$.

Each individual $m \in \mathcal{M}$ maintains a finite set of *individual states* $\Omega_{\{m\}}$. We typically write Ω_m instead of $\Omega_{\{m\}}$. An individual state captures the mental attitude of an individual, which may be as simple as binary states or more complex preferences and beliefs. We do not require that $\Omega_m = \Omega_{m'}$ for each $m, m' \in \mathcal{M}$.

We define $\Omega_M = \prod_{m \in M} \Omega_m$, with the factors of the product taken in ascending individual enumeration order $\langle \mathcal{M} \rangle$. Thus, Ω_M forms a *product space* containing all $|M|$-dimensional vectors whose elements are individual states.

A *population state* is a function $\omega \in (\mathcal{M} \to \Omega)$ such that $\omega(m) \in \Omega_m$ for each $m \in \mathcal{M}$. We say that $\omega(m)$ is an *individual state* for individual $m \in \mathcal{M}$. For any $M \subseteq \mathcal{M}$, we define $\omega[M] = \prod_{m \in M} \omega(m)$ such that $\omega[M] \in \Omega_M$. We often write $\omega = \omega[\mathcal{M}]$ and $\omega(\emptyset) = \emptyset$. We distinguish different individual states by writing $\omega_i(m) \in \Omega_m$, but often remove the subscript i when the distinction is unnecessary.

4.2 Markov Chains

Markov chains are a simple mathematical model for expressing probabilistic state transitions over a state space under certain basic assumptions. The simplicity of Markov chains and the abundance of corresponding analytical methods and results form the motivation for their use in our model.

Let $(\Omega_{\mathcal{M}}, \Im_{\mathcal{M}}, P_{\mathcal{M}})$ be a probability space (see [8, 13]. Let $T \subseteq \mathbb{N}$ be a countable set of contiguous *time instants* t. A *random state sequence* for some T is given by $(\omega^t)_{t\in T}$. We write ω_i^t to mean $\omega^t = \omega_i$ (i.e., the state at time t is ω_i). A sequence $(\omega^t)_{t\in T}$ satisfies the first-order *Markov property* if and only if:

$$P_{\mathcal{M}}(\omega_j^t \mid \omega_i^{t-1}, \omega_k^{t-2}, \ldots, \omega_l^1) = P_{\mathcal{M}}(\omega_j^t \mid \omega_i^{t-1}). \tag{1}$$

A sequence $(\omega^t)_{t\in T}$ is a discrete-time finite Markov chain with state space $\Omega_{\mathcal{M}}$ that is *time homogeneous*, if for any time instant $t \in T$ and any integer $k \geq 0$:

$$P_{\mathcal{M}}(\omega_j^t \mid \omega_i^{t-1}) = P_{\mathcal{M}}(\omega_j^{t-k} \mid \omega_i^{t-k-1}). \tag{2}$$

Otherwise, the Markov chain $(\omega^t)_{t\in T}$ is *time non-homogeneous*. For the remainder of this paper, we assume a time homogeneous Markov chain $(\omega^t)_{t\in T}$.

For each individual $m \in \mathcal{M}$, we assume a probability space $(\Omega_m, \Im_m, P_m)$ where $P_m \in (\Im_m \to [0,1])$ is an *individual state transition probability measure* for individual m on the measurable space $(\Omega_m, \Im_m)$ and $\Im_m = \mathcal{P}(\Omega_m)$ is the power set of individual states. We write $P_m(\omega_j^t(m) \mid \omega_i^{t-1}(m))$ for the conditional probability of individual m transitioning from state $\omega_i(m)$ at time instant $t-1$ to $\omega_j(m)$ at the next time instant t.

Let $M \subseteq \mathcal{M}$ be any set of individuals. We write P_M for a (sub-)population state transition probability measure on the measurable space $(\Omega_M, \Im_M)$. In the following discussion, we refer to a population instead of a sub-population, but the reader should note that we mean sub-population whenever $M \subset \mathcal{M}$. We define a population state transition probability measure P_M for any $\omega_i[M], \omega_j[M] \in \Omega_M$ as:

$$P_M(\omega_j^t[M] \mid \omega_i^{t-1}[M]) \stackrel{\text{def}}{=} \prod_{m\in M} P_m(\omega_j^t(m) \mid \omega_i^{t-1}(m)). \tag{3}$$

Population state probability measures are thus probabilistic product measures on the product space given by Ω_M.

Theorem 1. *For all $\omega_i, \omega_j \in \Omega_{\mathcal{M}}$, $M \subseteq \mathcal{M}$, and $t \in T$, if $[\![M]\!]$ is a partition of M, then $P_M(\omega_j^t[M] \mid \omega_i^{t-1}[M]) = \prod_{M'\in[\![M]\!]} P_{M'}(\omega_j^t[M'] \mid \omega_i^{t-1}[M'])$.*

Let $\pi \in (\mathcal{P}(\mathcal{M}) \to (T \times \Omega \to [0,1]))$ be a *state distribution function*. We write π_M for $\pi(M)$ and write simply π to mean $\pi(\mathcal{M})$. Moreover, for any $M \subseteq \mathcal{M}$, $t \in T$, and $\omega[M] \in \Omega_M$, we write $\pi_M^t(\omega[M])$ to mean $\pi(M)(t, \omega[M])$.

An *initial state distribution* denoted by π_M^0 satisfies $\sum_{\omega[M]\in\Omega_M} \pi_M^0(\omega[M]) = 1$. The state distribution π_M^t when starting with initial state distribution π_M^0 is given by

$$\pi_M^t = \pi_M^0 [P_M]^t, \tag{4}$$

where $[P_M]^t$ is the t^{th} power of the $|\Omega_M| \times |\Omega_M|$ state transition matrix induced by P_M.

This discussion of Markov chains forms the foundation in our influence model for analyzing the probabilistic state dynamics of individuals in a population that are influenced to transition between their individual states.

5 Influence Mechanisms

We have so far presented the foundations for the probabilistic transitions of our influence model. Although we have stated formally what is an individual transition measure, we have not yet specified how they capture a notion of social influence. We present in this section what we refer to as influence mechanisms, which capture the variety of ways that influence can be exerted among individuals.

Informally, an influence mechanism is a function that expresses the probabilistic effects of a certain type of influence exerted by a given sub-population on a particular individual. For example, a similarity metric over some preferences captured in individual states or a group size measure for particular individuals that have equivalent states (with respect to some information) are both influence mechanisms that can be used to define the transition probabilities of an individual (see also Section 1.1).

We assume a set $\mathcal{H} = (\mathcal{M} \times \Omega \to \mathcal{F})$ of *influence mechanism functions* and write $H \subseteq \mathcal{H}$ for a subset of influence mechanisms, where $\mathcal{F} = \bigcup_{m \in \mathcal{M}} (\Im_m \to [0,1])$. An *influence mechanism function* $h \in \mathcal{H}$ is defined such that for each $m \in \mathcal{M}$ and $\omega[M] \in \Omega$ we have $h(m, \omega[M]) \in (\Im_m \to [0,1])$ is a probability measure on the measurable space $(\Omega_m, \Im_m)$. The probability measure properties of $h(m, \omega[M])$ follow from the discussion in Section 4.

5.1 State Equivalence

For the influence problems that we consider, it is necessary to specify those individuals that share an equivalent state at a given time with respect to some notion of equivalence. Without this notion of equivalent states, we are unable to say anything about the particular information with which a population is being influenced. Moreover, we are not interested in sets of individuals that have the same complete mental attitude, but rather we focus on particular preferences, beliefs, and other such information that compose the mental attitudes.

For example, if we wish to model how individuals are influenced on their preferences for vacation spots, then their respective preferences for favorite color may be irrelevant. We restrict our focus to particular aspects of mental attitude by specifying equivalence relations on individual states.

We assume a finite set of *equivalence relations* $\mathcal{E}$ over the set of all individual states given by $\bigcup_{m \in \mathcal{M}} \Omega_m$. Each *equivalence relation* $e \in \mathcal{E}$ partitions $\bigcup_{m \in \mathcal{M}} \Omega_m$ into equivalence classes. We write $[\omega(m)]_e$ for the *equivalence class* that contains all individual states in $\bigcup_{m \in \mathcal{M}} \Omega_m$ that are equivalent to $\omega(m)$ with respect to the equivalence relation e. We write $\omega(m) \sim_e \omega'(m')$ if and only if $\omega(m)$ and $\omega'(m')$ are in the same equivalence class induced by e.

The set of all equivalence classes given by an equivalence relation $e \in \mathcal{E}$ is denoted by $Q_e = \{[\omega(m)]_e \mid \omega(m) \in (\bigcup_{m \in \mathcal{M}} \Omega_m)\}$. We extend this to the set of all equivalence classes over any equivalence relation by defining $Q = \bigcup_{e \in \mathcal{E}} Q_e$.

We build upon the notion of equivalence of individual states with respect to an equivalence relation by defining equivalence classes for (sub-)population states. For any $\omega, \omega' \in \Omega_{\mathcal{M}}$, $M \subseteq \mathcal{M}$, and $e \in \mathcal{E}$, we write $\omega[M] \sim_e \omega'[M]$ if and only if $\omega(m) \sim_e$

$\omega'(m)$ for each $m \in M$. This gives an equivalence class $[\omega[M]]_e = \{\omega'[M] \in \Omega_M \mid \omega[M] \sim_e \omega'[M]\}$. The set of all such equivalence classes forms a partition of Ω_M.

Definition 1. *Given $M \subseteq \mathcal{M}$, $\omega \in \Omega_{\mathcal{M}}$, and $e \in \mathcal{E}$, the state $\omega[M]$ is called a* homogeneous state *of M with respect to e if and only if $\omega(m) \sim_e \omega(m')$ for each $m, m' \in M$.*

Each influence mechanism $h \in \mathcal{H}$ is defined in terms of some particular information captured in the state of each individual. For example, h_{auth} may be defined only for some particular preferences or beliefs (e.g., ice cream preferences) and is undefined for all other individual state information (e.g., music preferences). This information is captured by an equivalence relation $e \in \mathcal{E}$.

The correspondence of each influence mechanism to a notion of state equivalence is given by a *mechanism state equivalence function* $g \in (\mathcal{H} \to \mathcal{E})$ that maps each influence mechanism $h \in \mathcal{H}$ to an equivalence relation $e \in \mathcal{E}$. We permit different influence mechanisms $h, h' \in \mathcal{H}$ to map to the same equivalence relation $e \in \mathcal{E}$ (i.e., $g(h) = e = g(h')$). For any set of influence mechanisms $H \subseteq \mathcal{H}$, we define $g[H] = \bigcup_{h \in H} g(h)$. For any $\omega \in \Omega_{\mathcal{M}}$, $m \in \mathcal{M}$, and $H \subseteq \mathcal{H}$, the equivalence class is given by $[\omega(m)]_{g[H]} = \{\omega'(m') \in (\bigcup_{m \in \mathcal{M}} \Omega_m) \mid (\omega(m), \omega'(m')) \in g[H]\}$.

For the remainder of this paper, we hold g constant. That is, all references to g are referring to the same mapping function.

5.2 Influence Neighborhoods

For each $m \in \mathcal{M}$, the set of individuals $M \subseteq \mathcal{M}$ that may exert influence on m through a mechanism $h \in \mathcal{H}$ is given by an influence neighborhood. Let $\delta \in (\mathcal{H} \to (\mathcal{M} \to \mathcal{P}(\mathcal{M})))$ be an *influence neighborhood function* defined such that for each $h \in \mathcal{H}$, $\delta(h)$ maps each $m \in \mathcal{M}$ to a set of individuals. For each $h \in \mathcal{H}$ and $m \in \mathcal{M}$, we refer to $\delta(h)(m)$ as the *influence neighborhood* of m under mechanism h.

For each $h \in \mathcal{H}$ and $m \in \mathcal{M}$, we define the influence neighborhood function as:

$$\begin{aligned}\delta(h)(m) \stackrel{\text{def}}{=} max\{M \subseteq \mathcal{M} \mid \forall \omega \in \Omega_{\mathcal{M}}, \forall M' \subseteq \mathcal{M}, \forall M'' \subseteq M, \text{if } M'' \neq \emptyset, \text{then}\\ h(m, \omega[M]) = h(m, \omega[M \cup M'])\\ \wedge h(m, \omega[M]) \neq h(m, \omega[M \setminus M''])\}. \end{aligned} \tag{5}$$

In other words, the influence neighborhood $\delta(h)(m)$ is the largest set of individuals $M \subseteq \mathcal{M}$ that exert influence on m such that inclusion of additional individuals to M does not change the influence mechanism function and the removal of any $m' \in M$ results in a different influence mechanism function than that defined using M.

For $H \subseteq \mathcal{H}$ and $m \in \mathcal{M}$, we write $\delta[H](m) = \bigcup_{h \in H} \delta(h)(m)$. For any $h \in \mathcal{H}$ and set of individuals $M \subseteq \mathcal{M}$, we write $\delta(h)[M] = \bigcup_{m \in M} \delta(h)(m)$. For $H \subseteq \mathcal{H}$ and $M \subseteq \mathcal{M}$, we write $\delta[H][M] = \bigcup_{h \in H} \bigcup_{m \in M} \delta(h)(m)$. We typically write $\delta_h(m)$ for $\delta(h)(m)$ and $\delta_H(m)$ for $\delta[H](m)$ for any $h \in \mathcal{H}$ and $H \subseteq \mathcal{H}$.

Each $h \in \mathcal{H}$ induces a population structure through the relational information provided by the influence neighborhoods for each $m \in \mathcal{M}$. Such structure is given by an *influence relation function* $f \in (\mathcal{H} \to \mathcal{R})$ that is defined for each $h \in \mathcal{H}$ as

$$f(h) \stackrel{\text{def}}{=} \{(m, m') \in \mathcal{M}^2 \mid m \in \delta(h)(m')\}. \tag{6}$$

For $h \in \mathcal{H}$, we refer to the relation $f(h)$ as the *influence relation* of h and the graph $G(f(h))$ as the *influence graph* of h.

Given an influence mechanism $h \in \mathcal{H}$, we write $h(m, \omega[\delta_h(m)])$ for the individual state transition probability measure resulting from influence exerted through h by the individuals in the influence neighborhood of m. We write $h(m, \emptyset)$ whenever $\delta_h(m) = \emptyset$. We often write h_i^m as shorthand for $h_i(m, \omega[\delta_h(m)])$, but sometimes remove the mechanism subscript to write only h^m.

Lemma 1. *For all* $h \in \mathcal{H}, m \in \mathcal{M}, \omega \in \Omega_{\mathcal{M}}$, *and* $M \supseteq \delta_h(m)$, *we have* $h(m, \omega[M]) = h(m, \omega[\delta_h(m)])$.

Lemma 2. *For all* $h \in \mathcal{H}$, $m \in \mathcal{M}$, *and* $\omega \in \Omega_{\mathcal{M}}$, *we have* $h(m, \emptyset)(\omega(m) \mid \omega(m)) = 1$.

6 Combining Multiple Influences

Just as different physical materials (e.g., wood, steel) can have different levels of susceptibility to different forces acting upon them (e.g., magnetic forces), so too can different individuals have different levels of susceptibility to different influence mechanisms (see [6, 4, 2]). Although one may argue that a single influence mechanism is sufficient for simplified problem domains, it is insufficient for modeling influence effects among individuals that may have different responses to the same influence mechanisms. This motivates the discussion in this section on combining the effects of multiple influence mechanisms.

A *combination method* is a function $c \in (\mathcal{P}(\mathcal{H}) \to (\mathcal{M} \times \Omega \to \mathcal{F}))$ that maps each set of mechanisms $H \in \mathcal{P}(\mathcal{H})$ to a function $c(H) \in (\mathcal{M} \times \Omega \to \mathcal{F})$, where $\mathcal{F} = \bigcup_{m \in \mathcal{M}} (\Im_m \to [0,1])$. Each $c(H)$ maintains the same properties as an influence mechanism function (see Section 5). We have from Lemma 1, for example, that $c(H)(m, \omega[M]) = c(H)(m, \omega[\delta_H(m)])$ for all $M \supseteq \delta_H(m)$. We often write c^H to mean $c(H)$.

A common method for combining multiple measures into a single measure is to take the average of the values given by each measure [15]. When applied to influence mechanisms, one may conceptualize this by viewing each influence mechanism as representing an expert's assessment about how an individual is influenced. The combination of these influence mechanisms can be viewed as being analogous to the problem of combining the influence probability assessments of experts. This problem has been widely studied in the area of group decision making (see, for example, [15, 9]). A related method for defining transition probabilities of individuals is used in [6].

The linear opinion pool combination method that we now present is a weighted average of the probabilities given under each single influence mechanism. Let $v \in (\mathcal{M} \to (\mathcal{P}(\mathcal{H}) \to [0,1]))$ be a *mechanism weight function* such that for each $m \in \mathcal{M}$ we have that $v(m)$ is a probability measure on the measurable space $(\mathcal{H}, \mathcal{P}(\mathcal{H}))$. For $m \in \mathcal{M}$, the value given by $v(m)(h)$ is the non-negative *mechanism weight* assigned to h by m. We typically write v_m instead of $v(m)$. We write $v_m^H(h) = \frac{v_m(h)}{\sum_{h \in H} v_m(h)}$ for the mechanism weight of h with respect to $H \subseteq \mathcal{H}$.

Let c_{lop} denote the *linear opinion pool combination method*. We define c_{lop} for an individual m, influence mechanisms $H \subseteq \mathcal{H}$, and states $\omega_i(m), \omega_j(m) \in \Omega_m$ as:

$$c_{lop}^{H}(m, \omega_i[\delta_H(m)])(\omega_j^t(m) \mid \omega_i^{t-1}(m)) \stackrel{\text{def}}{=} \sum_{h \in H} v_m^H(h) \cdot h^m(\omega_j^t(m) \mid \omega_i^{t-1}(m)). \quad (7)$$

If $\sum_{h \in H} v_m(h) = 0$, then we define $c_{lop}^{H}(m, \omega_i[\delta_H(m)])(\omega_i^t(m) \mid \omega_i^{t-1}(m)) = 1$. That is, an individual is not influenced to transition to a different state by H if they place zero weight on each mechanism $h \in H$.

7 Maximizing Influence

One such influence problem that has received considerable attention in recent years is the influence maximization problem [10]. We give a statement of this problem using the formal notation developed in this paper and that is sufficient for subsequent discussions of its application to our model.

The objects of concern to most of our analysis in the remainder of this paper are the population, targeted population, and set of influence mechanisms. The current population state (i.e., population state prior to targeted influence) and influence information do not impact our subsequent results. We simplify the exposition by defining a set $\mathcal{S} = (\Omega_{\mathcal{M}} \times (\bigcup_{m \in \mathcal{M}} \Omega_m))$ of all ordered pairs containing a population state and individual state. Specification of the set $\mathcal{S}$ permits a concise reference to a state pair $s \in \mathcal{S}$.

7.1 Specifying Influenced Individuals

In order to target a set of individuals for initial influence, the information that they are being influenced to adopt must be specified. We introduce an *initial state function* $\nu \in (\mathcal{P}(\mathcal{M}) \times \mathcal{S} \times \mathcal{E} \rightarrow \Omega_{\mathcal{M}})$ that is defined for $M \subseteq \mathcal{M}$, $(\omega_k, \omega_j(m)) \in \mathcal{S}$, and $e \in \mathcal{E}$ as the state $\omega_i \in \Omega_{\mathcal{M}}$ such that $\omega_i(m') \sim_e \omega_j(m)$ for each $m' \in M$, and $\omega_i(m'') = \omega_k(m'')$ and $\omega_i(m'') \not\sim_e \omega_j(m)$ for each $m'' \notin M$. We often write ω_i^0 for the initial state of a population given by the initial state function ν.

Definition 2. *Given $(\omega_k, \omega_j(m)) \in \mathcal{S}$ and $e \in \mathcal{E}$, a* targeted population *is a set of individuals $M^0 \subseteq \mathcal{M}$ such that $\omega_i^0(m') \sim_e \omega_j(m)$ for each $m' \in M^0$, and $\omega_i^0(m'') = \omega_k(m'')$ and $\omega_i^0(m'') \not\sim_e \omega_j(m)$ for each $m'' \notin M^0$.*

For a population M^0 targeted with state information $[\omega_j(m)]_e$, the initial population state is $\omega_i^0 = \nu(M^0, (\omega_k, \omega_j(m)), e)$, where ω_k is the population state prior to targeted influence. It follows directly from the definition of a targeted population M^0 that $\omega_i^0[M^0]$ is a homogeneous state with respect to e.

A population is assumed to be in a known state prior to specifying the targeted population. Specification of the initial individual state for each $m \in M^0$ thus entails a deterministic initial population state distribution π^0 such that $\pi^0(\omega_i) = 1$ for some $\omega_i \in \Omega_{\mathcal{M}}$, where $\omega_i[M^0]$ is a homogeneous state with respect to the equivalence relation e with which M^0 is targeted. If a population is currently in a state ω_k and M^0 are targeted to adopt states in $[\omega_j(m)]_e$, then $\pi^0(\omega_i) = 1$ where $\omega_i = \nu(M^0, (\omega_k, \omega_j(m)), e)$. Such a deterministic initial state distribution is consistent with assumptions in related models of influence that are applied to related problems, such as the influence maximization problem [10].

The *equivalent state function* $\rho \in (\mathcal{E} \to (\Omega \times \Omega \to \mathcal{P}(\mathcal{M})))$ gives the set of individuals whose states are equivalent to the states of some other individuals with respect to a particular equivalence relation. We write ρ_e to mean $\rho(e)$. For $e \in \mathcal{E}$, $\omega_i[M] \in \Omega$, and $\omega_j[M'] \in \Omega$ we define:

$$\rho_e(\omega_i[M], \omega_j[M']) \stackrel{\text{def}}{=} \{m' \in M' \mid \omega_j(m') \sim_e \omega_i(m) \text{ for each } m \in M\}. \quad (8)$$

The individuals in the set $\rho_e(\omega_i[M], \omega_j[M'])$ are those whose individual state in ω_j is in the same equivalence class with respect to $e \in \mathcal{E}$ as each $m \in M$ in state ω_i.

7.2 Expected Influence

Analysis of information propagation in a network requires a method for determining the expected number of individuals that have been influenced by some targeted population. This expected influence in a population is given by the *expected influence function* $\psi \in (\mathcal{P}(\mathcal{M}) \to (\mathcal{P}(\mathcal{M}) \times \mathcal{P}(\mathcal{H}) \times \mathcal{S} \to (T \to \mathbb{R}^+)))$ that gives the expected number of influenced individuals at each time $t \in T$. We define the expected influence at time $t \in T$ in a population $M \subseteq \mathcal{M}$ with a targeted population $M^0 \subseteq \mathcal{M}$, influence mechanisms $H \subseteq \mathcal{H}$, and state pair $s \in \mathcal{S}$ as:

$$\psi_M(M^0, H, s)(t) \stackrel{\text{def}}{=} \sum_{\omega[M] \in \Omega_M} |\rho_{g[H]}(\nu(M^0, s, g[H])[M^0], \omega[M])| \cdot \pi^t_M(\omega[M]), \quad (9)$$

where $\pi^t_M = \pi^0_M[P_M]^t$ is the population state distribution at time t for a Markov chain $(\omega[M])_{t \in T}$ defined on the state space Ω_M using H as the state transition functions. The initial state distribution π^0_M is defined such that $\pi^0_M(\omega_l[M]) = 1$ if and only if $\omega_l = \nu(M^0, s, g[H])$. The expected influence function gives the summation over each state $\omega_i[M] \in \Omega_M$ of the product of the probability of being in state $\omega_i[M]$ at time t and the number of individuals that are in a state in $\omega_i[M]$ that is $g[H]$-equivalent to that of individuals M^0 in ω_l.

7.3 Influence Maximization Problem

Given a number k, find k individuals that maximize the expected number of individuals that have adopted some particular information (i.e., individuals that have been influenced). A formal statement of the influence maximization problem follows:

Influence Maximization : Given $k > 0$, $s \in \mathcal{S}$, $H \subseteq \mathcal{H}$, and $t \in T$, what targeted population M^0 of size k maximizes $\psi(M^0, H, s)(t)$?

We do not seek to extend results achieved in [10] with respect to the problem itself, but rather will use the influence maximization problem to demonstrate some of the claimed improvements of our model over extant influence models (e.g., the independent cascade model used in [10]).

8 Improving Information Propagation

The primary contribution of our influence model comes from its leverage of multiple influence mechanisms for spreading information throughout a population. In this section, we provide results to support our claim that our model represents an improvement to extant influence models. Our results focus on the optimality of targeted populations and how the topology of a population affects or helps determine such optimal sets. We begin with a formal specification of certain topological features of a population that are required in order to formulate some of our results.

8.1 Connected Components

We do not assume that a population is connected under each relation. We here introduce a notion of connected components in a population in order to distinguish those individuals that are connected in the same connected component under one or more relations.

The individuals that are connected to $m \in \mathcal{M}$ by a path of any length in the transitive closure of the symmetrization of a relation $r \in \mathcal{R}$ is given by $[m]_r = \{m' \in \mathcal{M} \mid (m, m') \in r^{\leftrightarrow +}\}$. By definition, we have that $[m]_r \subseteq \{m\}$ for all $m \in \mathcal{M}$ that are isolated in $r \in \mathcal{R}$ (including self-loops). Moreover, if $m \in \mathcal{M}$ is not isolated in $r \in \mathcal{R}$, then $[m]_r \geq 2$. This follows from $m, m' \in [m]_r$ whenever $(m, m') \in r^{\leftrightarrow +}$. We define $Isolates(R) \stackrel{\text{def}}{=} \{m \in \mathcal{M} \mid [m]_R \subseteq \{m\}\}$ as the set of all individuals that are isolated (including self-loops) under $R \subseteq \mathcal{R}$.

These notions of individuals that are connected through a single relation can be extended to sets of relations and sets of individuals. We write $[m]_R = \bigcup_{r \in R} [m]_r$ for the set of individuals that are connected to m through any $r \in R$. For any $R \subseteq \mathcal{R}$ and $M \subseteq \mathcal{M}$, we write $[M]_R = \bigcap_{m \in M} [m]_R$ for the set of individuals connected to each $m \in M$ through any $r \in R$.

Definition 3. *A set of individuals $[M]_R$ is called a* maximal connected component *under relations $R \subseteq \mathcal{R}$ if and only if $[M]_R = [m]_R$ for any $m \in M$.*

Definition 4. *A connected component $[M]_R$ is called an* isolated component *if and only if $|[M]_R| \leq 1$.*

The set of all non-isolated maximal connected components for any $R \subseteq \mathcal{R}$ is given by $[\![\mathcal{M}]\!]_R = \{[M]_R \subseteq \mathcal{M} \mid |[M]_R| \geq 2\}$. For any $R \subseteq \mathcal{R}$, we write $\eta(R) = |[\![\mathcal{M}]\!]_R|$ for the number of distinct non-isolated maximal connected components under R. We include the number of isolated individuals under R by writing $\eta^+(R) = \eta(R) + |Isolates(R)|$. A set $[\![\mathcal{M}]\!]_R$ does not necessarily form a partition over $\mathcal{M}$, however, we do sometimes specify $[\![M]\!]$ for an arbitrary partition of $M \subseteq \mathcal{M}$.

Lemma 3. *For all $R \subseteq \mathcal{R}$, $\eta^+(R) \leq \sum_{r \in R} \eta^+(r)$.*

Lemma 4. *For all $R \subseteq \mathcal{R}$, $\eta(R) \leq \sum_{r \in R} \eta(r)$.*

Some individuals have a structural position in a population that has importance to our subsequent discussions of influence and information propagation.

Definition 5. *A pair $(m, m') \in \mathcal{M}^2$ is called a* bridge *between $r, r' \in \mathcal{R}$ if and only if $[m]_r \neq [m']_r$ and $(m, m') \in (r')^+$.*

In Figure 1, we can see that m_3 is isolated under relation $r_{coworker}$. Under relation r_{friend}, m_3 is not isolated and is connected to each individual in $\{m_1, m_2, m_4, m_8\}$. By definition of a bridge, we have, for example, that the pair (m_3, m_8) is a bridge between $r_{coworker}$ and r_{friend}.

Theorem 2. *For all $R \subseteq \mathcal{R}$, $\eta(R) = \sum_{r \in R} \eta(r)$ if and only if there does not exist a bridge between any $r, r' \in R$.*

Corollary 1. *For all $R \subseteq \mathcal{R}$, $[\![\mathcal{M}]\!]_R = \bigcup_{r \in R} [\![\mathcal{M}]\!]_r$ if and only if there does not exist a bridge between any $r, r' \in R$.*

8.2 Influence Optimality

Given $M \subseteq \mathcal{M}$, the measure ψ induces an ordering $\succsim_M$ over $\mathcal{P}(\mathcal{H})$ such that $H \succsim_M H'$ if and only if $\psi(M, H, s)(t) \geq \psi(M, H', s)(t)$ for all $s \in \mathcal{S}$ and $t \in T$. We define a similar ordering induced by ψ over sets of individuals. Given $H \subseteq \mathcal{H}$, the measure ψ induces an ordering $\succsim_H$ over $\mathcal{P}(\mathcal{M})$ such that $M \succsim_H M'$ if and only if $\psi(M, H, s)(t) \geq \psi(M', H, s)(t)$ for all $s \in \mathcal{S}$ and $t \in T$. These induced orderings give us notions of maximization and optimality.

Definition 6. *Given $M \subseteq \mathcal{M}$, H is* maximizing *with respect to M if and only if $H \succsim_M H'$ for all $H' \subseteq \mathcal{H}$.*

Definition 7. *Given $H \subseteq \mathcal{H}$, M is* maximizing *with respect to H if and only if $M \succsim_H M'$ for all $M' \subseteq \mathcal{M}$.*

We define the set of *maximizing targeted populations* for $H \subseteq \mathcal{H}$ as $O^*_H \stackrel{\text{def}}{=} \{M \subseteq \mathcal{M} \mid M \succsim_H M' \text{ for all } M' \subseteq \mathcal{M}\}$. The set of *minimal maximizing (or optimal) targeted populations* for $H \subseteq \mathcal{H}$ is given by $O^{**}_H \stackrel{\text{def}}{=} \{M \in O^*_H \mid |M| \leq |M'| \text{ for all } M' \in O^*_H\}$. Similarly, we define the set of *maximizing influence mechanisms* for $M \subseteq \mathcal{M}$ as $O^*_M \stackrel{\text{def}}{=} \{H \subseteq \mathcal{H} \mid H \succsim_M H' \text{ for all } H' \subseteq \mathcal{H}\}$. The *minimal maximizing (or optimal) influence mechanisms* for $M \subseteq \mathcal{M}$ is given by $O^{**}_M \stackrel{\text{def}}{=} \{H \in O^*_M \mid |H| \leq |H'| \text{ for all } H' \in O^*_M\}$. We write $O_H \in O^*_H$ and $O_M \in O^*_M$ for maximizing targeted populations and mechanisms, respectively. By definition, $O^{**}_H \subseteq O^*_H$ and $O^{**}_M \subseteq O^*_M$. We refer to each $M \notin O^{**}_H$ as *sub-optimal* with respect to H.

We restrict the sets of maximizing targeted populations to only those individuals that are connected under $f[H]$ for any $H \subseteq \mathcal{H}$ by defining $\hat{O}^*_H \stackrel{\text{def}}{=} \{M \subseteq \mathcal{M} \mid \exists M' \in O^*_H \text{ such that } M = M' \setminus Isolates(f[H]))\}$. The restatement in terms of optimal targeted populations is given by $\hat{O}^{**}_H \stackrel{\text{def}}{=} \{M \subseteq \mathcal{M} \mid \exists M' \in O^{**}_H \text{ such that } M = M' \setminus Isolates(f[H]))\}$. We write $\hat{O}_H \in \hat{O}^*_H$ for a particular (minimal) maximizing targeted population when using H. These restricted sets $\hat{O}^*_H$ and $\hat{O}^{**}_H$ are instrumental in stating of some of our subsequent results.

8.3 Impact of Topology on Information Propagation

There are certain properties of influence mechanisms that arise from distinguishing the maximal connected components of a relation. We here are particularly interested in the expected influence of mechanisms on a maximal connected component.

For each maximal connected component $[M_i]_{f[H]}$ under $f[H]$, we can determine an optimal targeted population $O_{H,i}$ that maximizes influence over M_i using $H \subseteq \mathcal{H}$. For any $M \subseteq \mathcal{M}$, $H \subseteq \mathcal{H}$, $s \in \mathcal{S}$, and $t \in T$, we write $\psi_{M_i}(M, H, s)(t)$ for the expected influence at time t on population M_i in the maximal connected component $[M_i]_{f[H]}$. The propagation of information being restricted to a maximal connected component $[M_i]_{f[H]}$ gives $0 \leq \psi_{M_i}(M, H, s)(t) \leq |M_i|$ for all $s \in \mathcal{S}$ and $t \in T$. Furthermore, $\psi_{M_i}(M, H, s)(t) = 0$ for all $M \subseteq \mathcal{M} \setminus M_i$, $s \in \mathcal{S}$, and $t \in T$ since each $m \in \mathcal{M} \setminus M_i$ exerts zero influence on any $m' \in M_i$. Similarly, $\psi_{M_i}(\emptyset, H, s)(t) = 0$ for all $1 \leq i \leq \eta^+(f[H])$, $s \in \mathcal{S}$, and $t \in T$.

Theorem 3. *For all $M, M', M'' \subseteq \mathcal{M}$, $H \subseteq \mathcal{H}$, $s \in \mathcal{S}$, and $t \in T$, if $M \cap M' = \emptyset$, then $\psi_M(M'', H, s)(t)$ is independent of $\psi_{M'}(M'', H, s)(t)$.*

Definition 8. *A mechanism $h \in \mathcal{H}$ is said to exert* non-vanishing influence *if and only if for all $s \in \mathcal{S}$ and $t \in T$, $\psi_{M_i}(M, \{h\}, s)(t) \geq 1$ for each $1 \leq i \leq \eta^+(f(h))$ whenever $M \cap M_i \neq \emptyset$.*

That is, a mechanism h is called non-vanishing if and only if there is an expectation that at least one individual will be influenced in each component that contains at least one individual from the targeted population.

Definition 9. *A set of mechanisms $H \subseteq \mathcal{H}$ is called a* non-vanishing set (of mechanisms) *if and only if each $h \in H$ is non-vanishing.*

Lemma 5. *For all $H \subseteq \mathcal{H}$ and $O_H \in O^*_H$, if H is a non-vanishing set of mechanisms, then $Isolates(f[H]) \subseteq O_H$.*

It should be clear that isolated individuals do not provide any benefit to the propagation of information in a population. This is due to the fact that isolated individuals must be contained in the initial targeted population in order to contribute to an increase in expected influence. As a result of Lemma 5, we can safely ignore the isolated individuals in most of our discussions about optimality of targeted populations.

We are able to place a lower-bound on the size of optimal sets for any non-vanishing set of mechanisms.

Theorem 4. *For all $H \subseteq \mathcal{H}$ and $O_H \in O^*_H$, if H is a non-vanishing set of mechanisms, then $|O_H| \geq \eta^+(f[H])$.*

Corollary 2. *For all $H \subseteq \mathcal{H}$, if H is non-vanishing set of mechanisms and $k < \eta^+(f[H])$, then the influence maximization problem using mechanisms H and parameter k will give a sub-optimal solution.*

For any set of mechanisms, we can show that the union of the optimal sets for each component forms an optimal set for the entire population.

Theorem 5. *For all $h \in \mathcal{H}$, if h is a non-vanishing mechanism, $n = \eta^+(f(h))$, and $O_{\{h\},i} \in O^{**}_{\{h\},i}$ for each $1 \leq i \leq n$, then $\bigcup_{i=1}^{n} O_{\{h\},i} \in O^{**}_{\{h\}}$.*

Extension of Theorem 5 to sets of mechanisms follows simply from the independence of influence between distinct maximal connected components.

Corollary 3. *For all $H \subseteq \mathcal{H}$, if H is a non-vanishing set of mechanisms, $n = \eta^+(f[H])$, and $O_{H,i} \in O^{**}_{H,i}$ for each $1 \leq i \leq n$, then $\bigcup_{i=1}^{n} O_{H,i} \in O^{**}_{H}$.*

We heretofore have not made any assumptions about the connectivity between the population topologies as recognized by each different influence mechanism. There may be bridges between different maximal connected components of the influence relations induced by different influence mechanisms. Such bridges facilitate the propagation of influence information between mechanisms that exert otherwise independent influence.

Theorem 6. *For all $H \subseteq \mathcal{H}$ and $h, h' \in H$, if $h \neq h'$, H is a non-vanishing set, there does not exist a bridge between $f(h)$ and $f(h')$, and $\hat{O}_{\{h\}} \in \hat{O}^{**}_{\{h\}}$ for each $h \in H$, then $\bigcup_{h \in H} \hat{O}_{\{h\}} \in \hat{O}^{**}_{H}$.*

9 Conclusions and Future Work

In this paper, we have given a formal exposition of our influence model that leverages multiple influence mechanisms and relations for propagating information throughout social networks. Our results highlight some of the benefits offered by such a model. We have shown that selection of optimal targeted populations can be decomposed into the selection of optimal targeted populations within connected components induced by multiple relations. There are many other types of analysis with our model that we are pursuing in on-going work.

A particularly interesting area for future work deals with the inversion of our influence model. Our focus has been placed on modeling how individuals are influenced and understanding how such influence affects the expected spread of information throughout a population. We view our current pursuit as being highly instrumental in modeling the inverse process of determining the originating source(s) of influence. That is, given the current state of a population, how were the individuals influenced to transition into their current states and what individuals are the originators of exerting that influence? Such an invertible model has important applications in areas such as computer security.

References

1. Asavathiratham, C. (2000). *The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains*. Ph.D. thesis, Massachusetts Institute of Technology.
2. Campos, A., Dignum, V., Dignum, F., Signoretti, A., Magály, A., Fialho, S. (2006). A process-oriented approach to model agent personality (extended abstract). In Decker, Sichman, Sierra, & Castelfranchi (Eds.), *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Budapest, Hungary*, pp. 1141-1142. International Foundation for Autonomous Agents and Multiagent Systems.

3. Carnes, T., Nagarajan, C., Wild, S. M., Zuylen, A. v. (2007). Maximizing influence in a competitive social network: A follower's perspective. In *Proceedings of the Ninth International Conference on Electronic Commerce (ICEC), Minneapolis, Minnesota, USA*, pp. 351-360. ACM.
4. Cialdini, R. B., Goldstein, N. L. (2004). Social Influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591-621.
5. Clifford, P., Sudbury, A. (1973). A model for spatial conflict. *Biometrika, 60*(3), 581-588.
6. Friedkin, N. E. (1986). A formal theory of social power. *Journal of Mathematical Sociology, 12*(2), 103-126.
7. Holley, R. A., Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Annals of Probability, 3*(4), 643-663.
8. Howard, R. A. (1971). *Dynamic Probabilistic Systems*, Vol. 1: Markov Models of *Series in Decision and Control*. John Wiley & Sons, Inc.
9. Jacobs, R. A. (1995). Methods for combining experts' probability assessments. *Neural Computation, 7*(5), 867-888.
10. Kempe, D., Kleinberg, J., Tardos, E. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the Ninth International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 137-146. ACM Press.
11. Kempe, D., Kleinberg, J., Tardos, E. (2005). Influential nodes in a diffusion model for social networks. In *Proceedings of the Thirty-Second International Colloquium on Automata, Languages and Programming (ICALP), Lisbon, Portugal*, Vol. 3580 of *Lecture Notes in Computer Science*, pp. 1127-1138. Springer Berlin / Heidelberg.
12. McPhee, W. N. (1963). *Formal Theories of Mass Behavior*. The Free Press of Glencoe.
13. Norris, J. R. (1997). *Markov Chains*. Cambridge Series on Statistical and Probabilistic Mathematics. Cambridge University Press.
14. Saito, K., Nakano, R., Kimura, M. (2008). Prediction of information diffusion probabilities for independent cascade model. In *Knowledge-Based Intelligent Information and Engineering Systems*, Vol. 5179 of *Lecture Notes in Computer Science*, pp. 67-75. Springer Berlin / Heidelberg.
15. Stone, M. (1961). The opinion pool. *Annals of Mathematical Statistics, 32*(4), 1339-1342.

Author Index