The 20th International Workshop on Multi-Agent-Based Simulation

MABS 2019

Montreal, Canada

May 13, 2019

PROGRAM CHAIRS

- Jaime Simão Sichman (University of São Paulo, Brazil)
- Mario Paolucci (ISTC, Italy)
- Harko Verhagen (Stockholm University, Sweden)

THE MABS STEERING COMMITTEE

- Frédéric Amblard (University of Toulouse, France)
- Luis Antunes (University of Lisbon, Portugal)
- Paul Davidsson (Malmö University, Sweden)
- Nigel Gilbert (University of Surrey, UK)
- Tim Gulden (George Mason University, USA)
- Emma Norling (Manchester Metropolitan University, UK)
- Mario Paolucci (National Research Council, Italy)
- Jaime Simão Sichman (University of São Paulo, Brazil)
- Takao Terano (Tokyo Institute of Technology, Japan)

PROGRAM COMMITTEE

- Diana Adamatti (FURG, Brazil)
- Frederic Amblard (Université Toulouse 1, France)
- Luis Antunes (University of Lisbon, Portugal)
- Tina Balke (Vanderlande Industries, The Netherlands)
- Joao Balsa (University of Lisbon, Portugal)
- Federico Bianchi (University of Brescia, Italy)
- Cristiano Castelfranchi (ISTC-CNR, Italy)
- Sung-Bae Cho (Yonsei University, Korea)
- Paul Davidsson (Malmö University, Sweden)
- Frank Dignum (Utrecht University, the Netherlands)
- Graçaliz Pereira Dimuro (FURG, Brazil)
- Bruce Edmonds (Centre for Policy Modelling, UK)
- Francisco Grimaldo (University of Valencia, Spain)
- Laszlo Gulyas (University of Budapest, Hungary)
- Rainer Hegselmann (University of Bayreuth, Germany)
- Marco Janssen (Arizona State University, USA)
- Jean-Pierre Muller (CIRAD, France)
- Luis Gustavo Nardin (Brandenburg University of Technology, Germany)
- Emma Norling (University of Sheffield, UK)
- Paulo Novais (Universidade do Minho, Portugal)
- Mario Paolucci (ISTC-CNR, Italy) **(co-chair)**
- William Rand (North Carolina State University, USA)
- Juliette Rouchier (CNRS/LAMSADE), France)
- Jaime Sichman (University of Sao Paulo, Brazil) **(co-chair)**
- Samarth Swarup (University of Virginia, USA)
- Klaus Troitzsch (University of Koblenz, Germany)
- Natalie Van der Wal (Vrije Universiteit Amsterdam, the Netherlands)
- Harko Verhagen (Stockholm University, Sweden) **(co-chair)**
- Neil Yorke-Smith (Delft University of Technology, The Netherlands)

CONTENTS

Modelling policy shift advocacy

Antoni Perello-Moragues[1,2,3], Pablo Noriega[2], Lucia Alexandra Popartan[4], and Manel Poch[4]

[1] Aqualia, Spain
[2] IIIA-CSIC, Barcelona, Spain
{tperello, pablo}@iiia.csic.es
[3] Universitat Autònoma de Barcelona, Spain
[4] LEQUIA, Universitat de Girona, Spain
{luciaalexandra.popartan, manuel.poch}@udg.edu

Abstract. In this paper, we propose to enrich standard agent-based social simulation for policy-making with an affordance inspired by *second-order emergent social phenomena*. Namely, we explore the inclusion of agents who have means to perceive, aggregate and respond to emergent collective outcomes and demand political intervention. Given this purpose, we work on a subclass of socio-cognitive technical systems that we called *value-driven policy-making systems*. We inspire and illustrate our proposal with a model of urban water management.

Keywords: agent-based social simulation · socio-cognitive technical systems · policy-making · values · second-order emergent phenomena

1 Introduction

Agent-based social simulation (ABSS) has been shown to be appropriate tool for policy-making [5]. Nonetheless, it has been suggested that in order to increase the usability for policy-making, standard ABSS may be enriched with some specific socio-cognitive affordances [11].

In this spirit, we proposed to afford some type of ethical reasoning and means to promote and assess moral behaviour [13]. The rationale being that, on the one hand, policy-makers draw on their political views and principles to design a policy intended to bring about a better state of the world, and deploy policy instruments that are consistent with such aim; and on the other hand, those agents who are subject to one such policy act according to their own principles, interests and motivations [17,3].

With this claim in mind, we characterised a type of agent-based simulators of public policies, as a subclass of socio-cognitive technical systems (SCTS) [9], that we called *value-driven policy-making systems*. They involve *values* as a first class notion and propose their operationalisation through *policy-schemas*, which consist of sets of *policy means* and *policy ends* [13].

In this paper, we extend that work with an affordance that we find specially relevant in some policy domains; namely, means to perceive, aggregate and respond to emergent collective outcomes. This affordance is inspired by the notion *second-order emergent social phenomena* (EP2) [11,4]. In order to illustrate our proposal, we model the management of urban water and, more specifically, the interplay between influential stakeholders (e.g. political factions) and their target groups in the process of advocating policy changes.

For these purposes, we start with a brief overview of our previous work and the type of second order phenomena simulation we propose (Sec. 2). In Sec. 3 we outline the core components of the enhanced framework. In Sec. 4 we present a model of the example and discuss some results. We close with remarks on further work (Sec. 5).

2 Background

Our aim is to define a framework for the simulation of *second-order social phenomena* in a sub-class of SCTS that have been characterised as *value-driven policy-making systems* (*VDPMS*) [13]. We build on the following ideas:

1. Socio-cognitive technical system (SCTS) are situated, on-line, hybrid, open regulated multi-agent systems [9]. They are composed by two first class entities: a *social space* and participating *agents*, who have socio-cognitive (opaque) decision models that guide their actions.

In previous work [13], we explored the role of values in the regulation of the social space and in the decision-making of agents. We proposed a core meta-model for that class that includes five components that we discuss in more detail in Sec. 3: (i) At least two *agent roles*; (ii) A *policy-schema* composed of *means* —that aim to produce a behavioural change on policy-subjects so as to drive the system towards a desirable world-state— and *ends* —that define those desirable world-states; (iii) A finite set of *values* that are projected onto the policy-schema; and (iv) *Satisfaction functions* for agents.

In addition, we assumed that policy-making presumes a socio-ecological context that determines the relevant part of the physical world and a policy domain that informs values, policy-schemas and satisfaction functions (Fig. 1).

2. Values. We assume a cognitive notion of value that may be used to model value-based reasoning for individuals, and value-based assessment of a state of the world [8,12,7,16,19]. Thus, we assume that values have the following six properties [16]: (P1): Values are beliefs; (P2): Values refer to desirable goals; (P3): Values serve as standards or criteria; (P4): Values are ordered by importance; (P5): The relative importance of multiple values guides action; (P6): Values transcend specific actions.

3. Second order emergence social phenomena (*EP2*) refers to the idea that agents may recognise an emerging macro-phenomenon and, as a consequence, they may intentionally support or hinder the phenomena or the emerging process itself [4,14,11]. Castelfranchi [4] approached *EP2* as the *cognitive emergence* of the macro-phenomena in the agent's mind, and afterwards a process of *cognitive immergence* that changes its behaviour. He discusses examples where the awareness of the phenomenon can promote or discourage it (e.g. urban segregation, ghetto formation). Other examples can be found in [11,18].

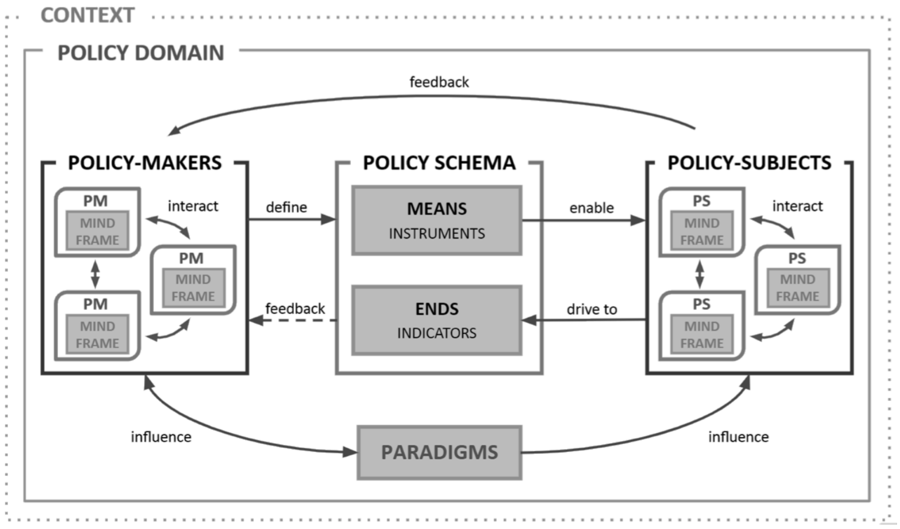


Fig. 1: Distinctive features of policy-making as a *value-driven socio-cognitive system* [13]

3 A conceptual framework

Due to space limitations, we cannot go into the detail needed to have a formal metamodel to represent SCTS with an appropriate level of accuracy. Nonetheless, we work on the basic components used in social coordination frameworks [1] and *VDPMS* [13], such as *socio-economic environment*, *values* and *policy-schemas*.

Rather, we address some features that should afford modelling some relevant second order emergence phenomena. Namely, we focus on the *perception model*, that refers to what and how agents perceive the environment. In the case of EP2, we distinguish between two types of perception and two types of evaluation:

(p.1) ***Micro-level*** **perception**, of those "local" variables in the state of the world that are observable by any individual agent and are part of its decision-making (e.g. a household's income, water use, etc.).
(p.2) ***Macro-level*** **perception**, of variables in a "global" environment, processed, and then stored as aggregated data (e.g. total population, average income of the neighbourhood, trends in water conservation practices, etc.).

(e.1) ***Micro-level*** **evaluation**, which would correspond to the assessment of the state of the local environment and
(e.2) ***Macro-level*** **evaluation**, which would correspond to the evaluation of a global environment, according to the socio-political understanding handled by the agent.

These distinctions characterise two types of agent-roles in the artificial society:

a. *Policy-subjects*, who only have micro-level perception, but evaluate the world-state both at the micro and macro levels, as they have ethical and political interests.
b. *Public Influencer* (*PIs*), who are political stakeholders in the domain —therefore they have value profiles and political agendas (i.e. their own values and goals)— and perceive and evaluate the world-state at the macro-level.

The rationale is that it is not usual that citizens draw conclusions directly from raw data of the system, that is, at best, open and accessible (for instance, from national statistics institutes). In general, most citizens do not have enough resources (e.g. time, attention, motivation, economic, technical, etc.) to process and reason about data and contexts that concern macro-scales. Nonetheless, citizens may include some features that affect the system as a whole, that involve other subjects and which may not be directly accessible by them. Rather, this information is often provided by *trusted* stakeholders that provide a framing discourse supported by *relevant* data (biased). These stakeholders have access to macro-level data —either because individuals have transferred directly their data to them, or because they use aggregated data from other entities— and are capable of processing and analysing it as well, which makes them capable of observing emergent phenomena (e.g. gentrification, demographic change, water use trends, etc.). Presumably, their *trustworthiness* and *relevance* arise from holding similar value profiles and socio-cognitive biases to the citizen that look upon them. Eventually, the citizen trusts the stakeholder, who shares the citizen's values and has its own political agenda, provides him with *useful* information and a *sound* framing (see [6]).

PIs in the real world are usually collectives (e.g. mass media companies, NGOs, think tanks, political parties, etc.) that one could notice they represent *political factions*. Nonetheless, notice that this representation can also fit for hybrid systems where agents may represent either human or artificial entities. In the hyper-connected society, human agents interact with artificial entities (e.g. virtual assistants, recommendation systems, etc.), and both of them may communicate with different *PIs*. For instance, in the context of household, a human agent is provided with a service through an appliance (e.g. laundry and washing machine). This device would take and transfer data of the user or the environment, and could also retrieve aggregated data or additional instructions from higher artificial entities (e.g. an order to delay wash program to avoid peak flows). In some way, there is an exchange of information (and resources) between agents of different levels.

Outlying a meta-model. As mentioned, *PIs* act in the social space: they have their own political agenda, perceive and evaluate the world-state, and interact with policy-subjects, who have also a political mind, in the sense that they reason about the collective life and its state.

In terms of the meta-model, we consider that there is an exchange of information between *PIs* and policy-subjects. The latter transfer micro-level data, that serves to produce aggregated data (able to be analysed at the system level in a political sense). The former communicate macro-level data and messages to

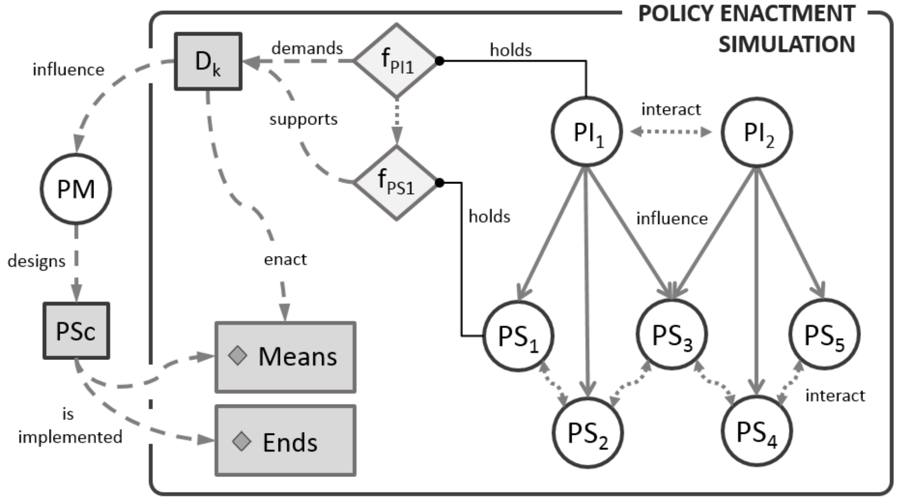


Fig. 2: Diagram of public influencers (*PIs*) raising political demands for producing policy shifts

policy-subjects, who will be used to evaluate the world-state. This affordance is added to the social interactions among policy-subjects (e.g. social influence).

It is assumed that *PIs* have their own functions of political satisfaction, which may be transferred to policy-subjects (basically by framing and making sense of the world-state, and providing this information).

In more precise terms, we implement *second-order social phenomena* in *VDPMS* as explained below (Fig. 2):

1. The agent *PI₁* has a function of political satisfaction f_{PI1}. These functions could be modelled as aggregation functions of variables (e.g. $f = \sum_j WjVj$) or satisficing conditions (e.g. thresholds).
2. The *policy-subject* agent PS_1 "delegates" the function of political satisfaction to the PI_1. It receives f_{PI1} and adapts its own function f_{PS1}, which takes into account micro-level and macro-level evaluation.
3. The *policy-subject* agent PS_3 looks upon two *PIs*, so it receives functions f_{PI1} and f_{PI2}. It can take both for its own f_{PS3} (e.g. combining them, discarding one, etc.).
4. If a *PI* agent is not politically satisfied (that is, the *desired world-state* and the *current world-state* are discrepant enough), it may suggest *political demands* D_k. Agents PS_i may support these (according on their own satisfaction).
5. The *policy-maker* designs a *policy-schema PSc* (i.e. *means* and *ends*) according to its values and presumably to the political demands raised in the social space.
6. Other demands raised by *PIs* may intervene directly on the social space bypassing *policy-makers* (e.g. adoption of social norms by *policy-subjects*). It can be interpreted as new, enacted *means*.

7. An updated *policy-schema* is eventually enacted in the social space.

Agents *evaluate* the world-state by comparing the desired state with the current one. This will serve to infer their *political satisfaction*. Thus, political participation is motivated by dissatisfaction with current political situations [10]. Presumably this will be implemented with evaluation functions that reflect the values of the agent (which must be represented by variables). These functions can take many forms, computationally speaking: weighted aggregation functions, thresholds, satisficing combinations (scrutinising all relevant values or only a subset of them), etc. There are diverse sophisticated methods to elicit functions empirically (e.g. [2,15]) and values (e.g.[16,19]).

3.1 Observations

Irrationality. We say that satisficing combinations are *irrational* when they are *unfeasible*. For instance, an agent wants to satisfy two values at the same time that are opposed, making the situation impossible.

It is true, however, that *unfeasibility* is hard to be demonstrated. In fact, some agents may even invoke unfeasibility as a political argument to attack the demands of other agents (in which case, it would not be *politically unfeasible*, but rather *politically undesirable* from the argument-maker's point of view).

When a *PI* holds political functions that are unfeasible, this would lead to a perpetual state of dissatisfaction, no matter the policy enacted. If these functions are communicated to policy-subjects, they are likely to be perpetually *unsatisfied* as well. This would entail unstable scenarios, as the *PI* (or even other *PIs*) could take advantage of the situation to make irrational political demands (irrational because they do not address the variables evaluated, either intentionally or unwittingly).

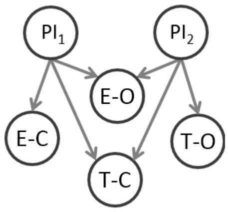
Policy misalignment. Policy-subjects might be incapable of perceiving the attainment of policy targets at the macro-scale, either because they do not receive the information by trusted *PIs* or because they do not have the values to make it relevant. Nonetheless, policy-subjects can be aware of the local effects of the policy. If the local effects are viewed as negative (e.g. restrictions or taxes) and the policy-subject is unable to perceive/value the macro-effects, this can lead to unstable social situations.

***PIs* and policy-makers' agendas.** Notice that *PIs* may evaluate the world-state using variables that policy-makers in charge may not consider relevant because of their values, thus being the latter incapable of perceiving the *same world-state*. Consequently, policy-makers will only receive the political demands once they have been raised due to the political dissatisfaction of *PIs* and policy-subjects. For this reason, the administration of the social space may become unstable.

4 A model for policy shift advocacy

Picture a neighbourhood of a city: each household houses a family with a certain income level, water needs, conservation practices, etc. There is a water utility

Table 1: Specification of the meta-model in a simple computational model

Network	**Demand**	**Set by**	**Supported by**	**Type**
	Suppress social tariff	PI1	E-C, E-O	Influence PM
	Create social tariff	PI2	T-C, T-O	Influence PM
	Change management	PI1, PI2	E-O, T-O	Influence PM
	Control demand	PI1, PI2	E-C, T-C	Enacted

company (a public or a private company) that supplies water, service that is supported by a fee. That fee may be adjusted according to additional features; thus, for example, some households may get a subsidy because of low income.

Citizens assess the service they get and may at some point want to have better conditions. Their satisfaction depends on what they believe is important (values), and they may identify some ways of adapting in order to increase their level of satisfaction. However, it is based not only in those variables that affect them directly and can assess on their own, but also it may depends on some features that affect the neighbourhood as a whole. This information is often provided by public influencers who assess the world-state with respect to their own values and promote adjustments in the way water is being managed. In order to get support, public influencers try to persuade households, yet their success would depend on the affinity with the values of the household.

Model. We model a crude urban environment to simulate the enactment of simplistic policies in a space formed by policy-subjects and *PIs* that hold different value profiles. The point of the exercise is to exemplify the notions of *EP2* in *VDPMS* and to illustrate the interplay of policy-subjects (i.e. households), *PIs* (i.e. influencers), and the policy-maker. This specification of the meta-model is summarised in Table 1.

The model represents an urban region and it is focused on the water supply public service and how its policies affect the world-state. On the one hand, citizens make use of the water supply for their basic needs, but they want the service to be managed according to their understanding of justice and welfare as well. On the other hand, public influencers (i.e. *PIs*) may demand political measures if they consider that the world-state is not aligned with their public values. The purpose of the model is to test policies (introduced as norms and actions) and then observe their effects on the socio-economic environment and on the acceptance of *PIs*.

Agents. We consider two type of agents in the model: (a) *households* (i.e. policy-subjects) and (b) *public influencers* (i.e. *PIs*). *Households* are characterised by (i) value profile; (ii) number of dwellers; (iii) income; (iv) water demand; (v) conservation practices; (vi) service satisfaction; and (vii) political satisfaction. Elements (ii–iii) are based on real-data (from the Spanish Statistical Office), (iv–vii) evolve as results of the simulation, and (i) is an input of the model. *PIs* are characterised by (i) value profile; and (ii) political satisfaction.

Scales and process overview. The model simulates one decade of activity through discrete time steps of one month. Each month households demand water, receive the water bill, may adopt conservation practices (to protect the environment or to have more wealth to spend in other goods), assess their satisfaction, and may support political demands. Likewise, *PIs* evaluate the world-state and may push political demands. The model represents an urban district in Barcelona.

Value profiles. *PIs* hold *public values*, related to public affairs and organisational settings, for which we use the work on Public Service Values [19]. In contrast, households hold *motivational values*, related to needs and goals, for which we use the Schwartz Theory of Basic Values [16]. It is recognised that they also hold *public values*, which are defined by their interaction with the *PIs*. For the sake of simplicity, values are static during the simulation.

The typologies of households are defined according to the classification of value sets in the Schwartz Theory of Values. There are two pair of opposite dimensions. On the one hand, the pair of *self-enhancement*, which focus on self-esteem and the pursuit of self-interests; and *self-transcendence*, that concern for the welfare and interests of others. On the other hand, the pair of *conservation*, which stress resistance to change, order, self-restriction, and subordination of oneself in favour of socially imposed expectations; and *openness to change*, that emphasises the independent behaviour and readiness for new experiences.

There are four typologies according to the value dimensions that are predominant in the household:

– **E-O:** *self-enhancement* and *openness to change. Households E-O* value *power* (i.e. social power, wealth, authority); *achievement* (i.e. ambition, influence, capability); and *self-direction* (i.e. freedom). Focusing on its own welfare, these households think that the service must ensure the autonomy of households (they consider it is well represented by wealth and budget) and, then, its own (financial) sustainability.
– **E-C:** *self-enhancement* and *conservation. Households E-C* value *achievement* (i.e. ambition, influence, capability); *power* (i.e. social power, wealth, authority); and *security* (i.e. social order), *tradition*, and *conformity* (i.e. compliance). Focusing on its own welfare, these households do not want any shock/policy that can put the institutions and the public service at risk (for example, a social subsidy, that they think that may jeopardise the financial sustainability of the service).
– **T-O:** *self-transcendence* and *openness to change. Households T-O* value *benevolence* and *universalism* (i.e. equity, environment, social justice, peace); and *self-direction* (i.e. freedom). Focusing on social welfare, these households think that the service must protect the access to households while ensuring the preservation of the environment.
– **T-C:** *self-transcendence* and *conservation. Households T-C* value *benevolence* and *universalism* (i.e. equity, environment, social justice, peace);

and *security* (i.e. social order), *tradition*, and *conformity* (i.e. compliance). These households believe that the service must provide support to vulnerable households and must not waste resources on respect to others who also need them.

There are two *PIs*: *PI1*, whose values are economic responsibility, then citizen autonomy, and finally equality (understood as equal treatment); and *PI2*, whose values are conservation of the environment, social justice, and that the access of households to the service is assured and protected.

PIs' satisfaction. *PI1* focuses first on the financial sustainability of the public service. When the service is covered entirely by the water tariff, it examines the number of households whose impact on their budget due to the tariff is significant. If the service is not sustained by the total billing, it checks whether a social tariff —which is a subsidised tariff for those households categorised as vulnerable— is active; in case there are many vulnerable households, it may blame that policy for hindering the sustainability of the service. *PI2* audits first the average water demand of households and then focuses on the number of households whose impact on their budget due to the tariff is significant. In any case, a policy that includes a social tariff mitigates its discontent.

Households' satisfaction. Households' satisfaction is divided into two components depending on the context: *service satisfaction* (i.e. household context) and *political satisfaction* (i.e. system context). On the one hand, households use local variables within the context of using the service at home to decide whether the water service meets their standards or not. So far, as the model is basic, they only perceive the impact on their budget, and evaluate the service accordingly. For more sophisticated models, they could include other locally-perceivable variables as access to the service, interruption of supply, water quality, water pressure, company intrusion, etc. On the other hand, households evaluate the world-state according to the values they hold. The political satisfaction components and framing is delegated to the *PIs*, as they are capable of perceiving the whole world-state. *Households E-O* and *T-C* look upon both *PIs* and make an aggregation, while *households E-C* and *households T-O* take into account only one *PI* (*PI1* and *PI2*, respectively). Eventually, they make a mean of the two components to elicit their global satisfaction.

Political demands. Both *PIs* may try to convince policy-subjects to diminish their water demand by releasing information about the water use at the society level and appealing to be within a *normal* range. Notice that it is done due to different motivations depending on the *PI* (i.e. citizen autonomy and protection of the environment, respectively). Only *households E-C* and *T-C* can support and follow this advise (because they want to abide by social norms). When the service is not sustainable, the *PI1* may advocate for suppressing the social tariff if it is active, or even support to change the management model in case it is not (e.g. privatisation or terminate the contract for the concession). In contrast, the *PI2* may advocate for establishing a social tariff to protect vulnerable households, or even demand to terminate the management when the protection of the environment is unacceptable. As mentioned, households may support these demands depending on their value profile and their level of global satisfaction.

4.1 Simulation examples

Gentrification scenario: The population starts constituted by 200 households: 25 % *households E-O*, 50 % *households E-C* and 25 % *households T-C*. Each month, households with the lowest income are forced to move out and are replaced by wealthy *households T-O*, causing that the original population is entirely replaced in 5 to 8 years. Additionally, a policy that establishes financial aid for vulnerable households —those whose water bill exceeds a defined threshold— is active (i.e. social tariff).

As former residents are replaced by new residents with environmentalist behaviour, the collective water demand becomes lower over time. Consequently, the service is becoming financially unsustainable, since it has been designed so that a minimum water amount must be demanded by each person (Fig. 3). Apart from this, new residents are wealthy and therefore not categorised as vulnerable households (Fig. 4). This results in a period in which *PI1* is completely dissatisfied (Fig. 5). The service has become financially unsustainable, there are too many households that are not autonomous —in the sense they have to face too high water bills in comparison to their income—, and too many households receive financial aid. This leads this *PI1* to demand for the social tariff, proposal that is supported (between 10 and 20 %) during a period of 2 years (Fig. 6). Nonetheless, its support decreases over time because newcomers' values do not align with the proposal —in Fig. 5 the average household satisfaction reflects *PI1*'s assessment. Anyway, as the population is being replaced, and although the financial situation of the service is only partially acceptable, the new residents are solvent and do not need social assistance, which satisfies the values of *PI1*. This world-state is acceptable enough to dissuade *PI1* to demand for the suppression of social aid. *PI2* is satisfied because environmental protection is ensured —households use an acceptable amount of water—, there is a policy of social tariffs for vulnerable households, and newcomers have access to the service (actually, they are wealthy enough to not be in a precarious situation); this political assessment is communicated to the new population, causing the average household satisfaction to increase again (Fig. 5).

5 Closing remarks

In this paper, we have proposed to enrich simulation of policy-making taking inspiration from *second-order emergent social phenomena*, and we outlined a meta-model for it. Further work should enhance the meta-model by considering interactions between *PIs* in political arenas and more dynamic networks of relationships between *PIs* and *policy-subjects*.

Furthermore, applications of this meta-model should be complemented with sophisticated the satisfaction models (for instance, represented by knowledge

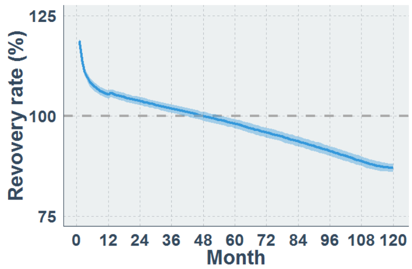


Fig. 3: Recovery rate (%) under a gentrification scenario.

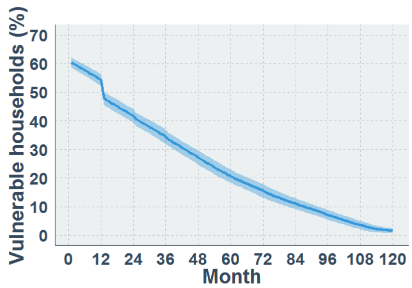


Fig. 4: Households categorised as vulnerable (%) under a gentrification scenario.

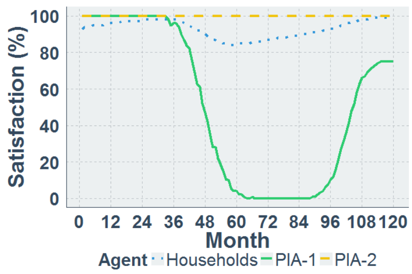


Fig. 5: Satisfaction of agents during the gentrification scenario

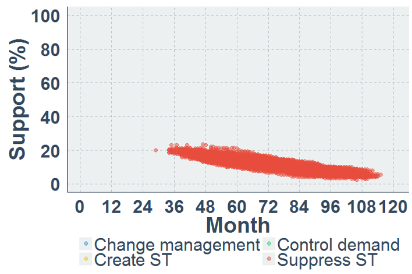


Fig. 6: Demands and support during the gentrification scenario

based systems with production rules or aggregation functions), which requires field-data to empirically elicit values, hierarchies, indicators and associated political demands. As far as we know, participatory approaches of ABSS consists of gathering the stakeholders to explore together the dynamics of a socio-ecological system, but not modelling and implementing their own behaviour in the simulation.

Apart from this, we open the possibility to include an autonomous, artificial policy-maker agent, that is capable of enacting *means* according to the world-state at a macro-level and taking into account the satisfaction of *PIs* and policy-subjects and their demands. Also, this could be further explored with *machine learning* so as to train the artificial policy-maker.

Acknowledgements. The first author is supported with the industrial doctoral grant 2016DI043 of the Catalan Secretariat for Universities and Research (AGAUR), sponsored by FCC AQUALIA, IIIA-CSIC, and UAB

References

1. Aldewereld, H., Boissier, O., Dignum, V., Noriega, P., Padget, J.: Social Coordination Frameworks for Social Technical Systems. No. 30 in Law, Governance and Technology Series, Springer International Publishing (2016)
2. Bennett, J., Blamey, R.: The choice modelling approach to environmental valuation. Edward Elgar Publishing (2001)
3. Campbell, J.L.: Ideas, politics, and public policy. Annual Review of Sociology **28**(1), 21–38 (2002)
4. Castelfranchi, C.: Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In: Sichman, J.S., Conte, R., Gilbert, N. (eds.) Multi-Agent Systems and Agent-Based Simulation. pp. 26–44. Springer Berlin Heidelberg (1998)
5. Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., Narasimhan, K.P., Wilkinson, H.: Computational modelling of public policy: Reflections on practice. Journal of Artificial Societies and Social Simulation **21**(1), 14 (2018)
6. Lakoff, G.: Why it matters how we frame the environment. Environmental Communication **4**(1), 70–81 (2010)
7. Mercuur, R., Dignum, V., Jonker, C.: The use of values for modeling social agents. In: Quan Bai, Fenghui Ren, M.Z.T.I. (ed.) Proceedings of the 3nd International Workshop on Smart Simulation and Modelling for Complex Systems (2017)
8. Miceli, M., Castelfranchi, C.: A Cognitive Approach to Values. Journal for the Theory of Social Behaviour **19**(2), 169–193 (1989)
9. Noriega, P., Padget, J., Verhagen, H., d'Inverno, M.: Towards a framework for socio-cognitive technical systems. In: Ghose, A., Oren, N., Telang, P., Thangarajah, J. (eds.) Coordination, Organizations, Institutions, and Norms in Agent Systems X. pp. 164–181. Lecture Notes in Computer Science 9372, Springer (2015)
10. Oishi, S., Diener, E., Lucas, R.E.: The optimum level of well-being: Can people be too happy? In: Diener, E. (ed.) The Science of Well-Being, pp. 175–200. Springer Netherlands (2009)
11. Pablo Noriega, Jordi Sabater-Mir, H.V.J.P., d'Inverno, M.: Identifying affordances for modelling second-order emergent phenomena with the WIT framework. In: Sukthankar, G., Rodriguez, J.A. (eds.) AAMAS 2017 Workshops Visionary Papers. Springer International Publishing (2017)
12. Parks, L., Guay, R.P.: Personality, values, and motivation. Personality and Individual Differences **47**(7), 675–684 (2009)
13. Perello-Moragues, A., Noriega, P.: Using Agent-Based Simulation to understand the role of values in policy-making. In: Proceedings of the Social Simulation Conference 2018 (SSC2018) (In Press)
14. Rosaria Conte, Giulia Andrighetto, M.C., Paolucci, M.: Emergent and immergent effects in complex social systems. In: Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence. pp. 8–11 (2007)
15. Saaty, T.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
16. Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In: Zanna, M.P. (ed.) Advances in Experimental Social Psychology, vol. 25, pp. 1–65. Academic Press (1992)
17. Stewart, J.: Public Policy Values. Palgrave Macmillan UK, London, 1st edn. (2009)
18. Vanhée, L., Ferber, J., Dignum, F.: Agent-Based Evolving Societies. In: Kamiński, B., Koloch, G. (eds.) Advances in Social Simulation. pp. 283–294. Springer Berlin Heidelberg (2014)
19. Witesman, E., Walters, L.: Public Service Values: A new approach to the study of motivation in the public sphere. Public Administration **92**(2), 375–405 (2014)

Reinforcement Learning of Supply Chain Control Policy using Closed Loop Multi-Agent Simulation

Souvik Barat[0000−0002−4677−1957], Prashant Kumar[0000−0002−3850−383X], Monika Gajrani[0000−0002−8858−1748], Harshad Khadilkar[0000−0003−3601−778X], Hardik Meisheri[0000−0002−9014−1098], Vinita Baniwal[0000−0002−7329−3540], and Vinay Kulkarni[0000−0003−1570−1339]

TCS Research
souvik.barat@tcs.com

Abstract. Reinforcement Learning (RL) has achieved a degree of success in control applications such as online gameplay and autonomous driving, but has rarely been used to manage operations of business-critical systems such as supply chains. A key aspect of using RL in the real world is to train the agent before deployment by computing the effect of its exploratory actions on the environment. While this effect is easy to compute for online gameplay (where the rules of the game are well known) and autonomous driving (where the dynamics of the vehicle are predictable), it is much more difficult for complex systems due to associated complexities, such as uncertainty, adaptability and emergent behaviour. In this paper, we describe a framework for effective integration of a reinforcement learning controller with an actor-based multi-agent simulation of the supply chain network including the warehouse, transportation system, and stores, with the objective of maximizing product availability while minimising wastage under constraints.

1 Introduction

Business-critical systems such as supply chain networks require continual evaluation and adjustment to stay competitive and economically viable in a dynamic environment. Reinforcement Learning (RL) [35, 29] is a class of machine learning algorithms that can be used for controlling such complex systems in an adaptive and flexible manner. The goal of the system controller (also called *RL agent*) is to learn to take the best possible control actions in each possible state of the system, in order to maximise long-term system objectives. A crucial aspect of RL is the computation of next state and associated rewards for the chosen action(s), in a closed loop to enable learning. In compact systems with well-understood behaviour such as software-based games or vehicle dynamics, the action-driven state transition is simple to model, at least in terms of a probabilistic description. This is not the case for complex networked systems with a large number of entities that have their own micro-behaviour, and where the individual interactions build into emergent (and sometimes unpredictable) macro-behaviour. In such scenarios, top-down modelling allows for only a coarse approximation of the next states and rewards, hampering the training process of the RL agent.

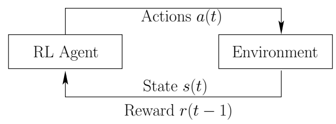


Fig. 1. Interaction of RL agent with an environment (actual system or simulation).

A more accurate representation of the next state and reward (see Fig. 1), and consequently a better estimate of the long term consequences of a series of actions for an *exceedingly complex* [33] business-critical system such as a supply should be possible using a bottom-up multi-agent simulation approach. Fundamentally, these systems are *open* as they exchange messages with their environment, and *complex* as they contain multiple non-linear feedback loops [3]. Moreover, these systems are not monolithic deterministic automatons, but are complex (scale-free) networks [5] or *systems of systems*, where the global behaviours emerge from the interactions of *autonomous*, *adaptable*, and *self-organising* sub-systems and constituent *agents* [15]. These characteristics pose obstacles to the application of alternative control approaches such as adaptive control and approximate dynamic programming. While the former requires an analytical representation of the control and adaptation laws, the latter requires at least a one-step rollout of a significant subset of actions, followed by a functional approximation of the subsequent value function.

We postulate that the use of analytical expressions for modelling (the method of choice for simpler RL applications [11, 19]), is infeasible for complex systems, and instead advocate an agent/actor based modelling abstraction [1]. The paper presents a framework that uses reinforcement learning for exploring policies and deciding control actions, and an actor-based modelling and simulation technique to perform accurate long-term rollouts of the policies, in order to optimise the operation of complex systems. The key attraction of RL is that online decision making is a one-shot forward pass through (typically) a set of neural networks, and does not require online search. We use the domain of supply chain replenishment as an illustrative example to demonstrate the proposed modelling abstraction and its impact on training RL agent prior to its deployment.

2 Problem formulation

Generic RL problem: A reinforcement learning problem is described by a Markov Decision Process (MDP) [35] represented by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, P, \gamma)$. Here, $\mathcal{S}$ is the set of states of the system, $\mathcal{A}$ is the set of control actions that the RL agent can choose from, $\mathcal{R}$ is the set of possible rewards, P is the (possibly stochastic) transition function from $\{\mathcal{S}, \mathcal{A}\} \rightarrow \mathcal{S}$, and γ is a discount factor for future rewards. In several cases, the agent is unable to observe the state space entirely, resulting in a partially-observable MDP or POMDP [35]. A set of observations $\mathcal{O}$ is derived from $\mathcal{S}$ to represent what the agent can sense. The goal of the RL agent is to compute a policy $\mathcal{O} \rightarrow \mathcal{A}$ that maximises the future discounted long-term reward. Clearly, an accurate representation of the transition function $P : \{\mathcal{O}, \mathcal{A}\} \rightarrow \mathcal{O}$ is a critical aspect of this effort.

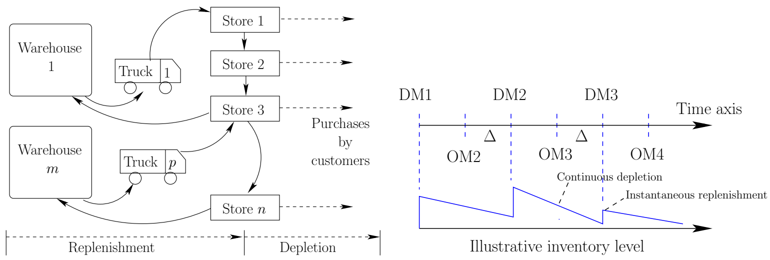


Fig. 2. [Left] Schematic of supply chain replenishment use case. [Right] Schematic of the periodic replenishment cycles. OM are ordering moments when the actions are computed, while DM are delivery moments when the inventory is actually delivered.

Specific instance: We illustrate the generic RL problem in the context of supply chain replenishment, which presents well-known difficulties for effective control [22, 30]. The scenario is that of a grocery retailer with a network of stores and warehouses served by a fleet of trucks for transporting products. The goal of replenishment is to regulate the availability of the entire product range in each store, subject to the spatio-temporal constraints imposed by (i) available stocks in the warehouses, (ii) labour capacity for picking and packaging products in the warehouses, (iii) the volume and weight carrying capacity of the trucks, (iv) the transportation times between warehouses and stores, (v) the product receiving capacity of each store, and (vi) available shelf space for each product in each store. A schematic of the product flow is shown in Fig. 2 [Left].

A temporal illustration of the replenishment process is shown in Fig. 2 [Right]. The replenishment of inventory is assumed to take place periodically (typically every 6 hours), at the time instants marked as DM (Delivery Moments). However, since it takes a non-zero amount of time to procure the new inventory within the warehouse, to transport it to the store, and to load it onto the shelves, the replenishment quantities of each product are computed at the time instants marked OM (Order Moments). There is a lead time Δ provided between each OM and the subsequent DM. The inventory of products is a monotonic non-increasing function between delivery moments, and there is a step increase at every DM when new inventory is provided to the stores.

Processes to be modelled: The warehouses stock a range of products and supply them to the stores as described in Fig. 2. This involves packing the products (using trolleys), loading packed products to the trucks/carriers and delivering them to respective stores on predefined routes. Each sub-process contains constraints such as the warehouse labour capacity, machine capacity, number of trucks, and the truck volume/weight capacities. The uncertainties that emerge due to the probabilistic behaviours of the individual elements are: unavailability and varying productivity of the labours, sub-optimal machine throughput and unavailability and unaccounted delays of the trucks. Trucks are constrained by the volume and weight capacities, often they are suited for specific types of products and each of them has probabilistic characteristics, such as: propensity for transportation delay and break down.

Let us assume that there are m warehouses, p trucks, and n stores in the system. From operational perspective, each store stocks $i = \{1, \ldots, k\}$ unique varieties of products, each with a maximum shelf capacity $c_{i,j}$ where $j \leq n$ is the index of the store. Further, let us denote by $x_{i,j}(t)$ the inventory of product i in store j at time t. The replenishment quantities (*actions*) for delivery moment d are denoted by $a_{i,j}(t_d)$, and are to be computed at time $(t_d - \Delta)$. The observation $O(t_d - \Delta)$ consists of the inventory of each product in each store at the time, the demand forecast for each product between the next two delivery moments, and meta-data such as the unit volume v_i and weight w_i, and its shelf life l_i.

Note that the *states* differ from the *observations* in this case because the actual inventory at the time of replenishment is $x_{i,j}(t_d)$, which must be estimated based on the current inventory $x_{i,j}(t_d - \Delta)$ and some forecast of the depletion in the remaining time Δ. The inventory $x_{i,j}(t)$ depletes between two delivery moments $(d-1)$ and d, and undergoes a step increase by amount $a_{i,j}(t_d)$ at time t_d. The actions are constrained by the various capacities in the system, including those within warehouses, in the transportation network, and in the store. The reward $r(t_{d-1})$ is a function of the previous actions $a_{i,j}(t_{d-1})$ and the evolution of inventory states $x_{i,j}(t)$ in $t \in [t_{d-1}, t_d)$. From a business perspective, of particular interest are: (i) the number of products that remain available throughout the time interval $[t_{d-1}, t_d)$, and (ii) the wastage of any products that remain unsold past their shelf lives. Mathematically, we define this as,

$$r(t_{d-1}) = 1 - \frac{\text{count}(x_{i,j} < \rho)}{k\,n} - \frac{\sum_{i=1}^{k} \sum_{j=1}^{n} w_{i,j}(t_{d-1})}{\sum_{i=1}^{k} \sum_{j=1}^{n} X_{i,j}}, \tag{1}$$

where $\text{count}(x_{i,j} < \rho)$ is the number of products that run out of inventory (drop below fraction ρ) at some time $t \in [t_{d-1}, t_d)$, $w_{i,j}(t_{d-1})$ is the number of units of product i in store j that had to be discarded in the time interval because they exceeded their shelf lives, and $X_{i,j}$ is the maximum shelf capacity for product i in store j. Since both negative terms in (1) fall in the range $[0, 1]$, we see that $-1 \leq r(t_{d-1}) \leq 1$. The goal of the control algorithm is to compute actions $a_{i,j}(t_{d-1})$ that maximise the discounted sum of these rewards from the present moment onwards, $\sum_{z=0}^{\infty} \gamma^z r(t_{d+z})$.

3 Related work

Control approaches: Supply chain control, including inventory management, has been a problem of interest for a long time [32]. Theoretically, such problems can be solved using mixed-integer linear programming [34, 7], but this is infeasible at real-world scales. Instead, actual implementations typically use approximate methods such as threshold policies [9]. Other traditional methods such as state feedback [6] and model predictive control [24] have similar scaling issues. Adaptive critics [31] and reinforcement learning [12, 17, 27] have also been used in literature, but primarily for single-product scenarios. However, these methods along with approximate dynamic programming (ADP) are likely to be the best suited for our problem, because they are known to work in related areas.

ADP has been used for task allocation problems in transportation networks [13, 37], but has the inherent restriction of requiring analytical descriptions of value functions and at least a one-step rollout of the policy. Imitation Learning has been used in robotics [10], but is only feasible where an expert policy is available to imitate. Reinforcement learning has been used in the computation of torque commands for robotic applications [28, 20]. In these cases as well as other complex systems [11, 19], the system models are analytically defined, thus simplifying the computation of step rewards and next state of the system. This is because RL is effective only when the (stochastic) transition functions closely approximate the real system to be controlled. In situations where the system cannot be described analytically, algebraic expressions cannot be used to compute rewards and transitions. Where RL has been used in supply chain management [12, 17, 38, 27], it tends to focus on single-product scenarios. An experimental approach can be used for training the RL agent when the system is non-physical (for example, is itself a piece of software as in the case of computer games) [26]. However, applying experimental approach on the actual system is not feasible in the case of business-critical systems. Therefore, the development of (and integration with) a high-fidelity simulation model is crucial for effective training of the RL agent and controlling complex systems.

Modelling approaches: Complex systems are typically modelled using two broad categories of approaches: *top-down* approach and *bottom-up* approach [36]. A *top-down* approach visualises a system from a higher scale and specifies aggregated macro-behaviour. This approach uses a range of models, such as *mathematical/analytical model* and *enterprise model* (EM), to represent and analyse the system as a whole. The *analytical models*, *e.g.*, Mixed Integer Linear Programming, represent a system using mathematical formulae and use rigorous mathematical and statistical problem solving techniques for system analysis. The *operational research techniques* are the specialised form of analytical models. The *enterprise models*, such as ArchiMate [16], BPMN [39] and System Dynamics [25], also serve a wide range of modelling and analysis needs by representing aggregated system behaviour. However, these approaches are found to be inadequate to represent systems (and their transition functions P) that contain multiple adaptable, self organising and uncertain entities (such as warehouses, trucks, products and store), individualistic behaviour (such as product expiry) and exhibit emergent behaviours (such as availability, unavailability and damages of products that are significantly influenced by several uncertain spatio-temporal aspects: transportation delay, inappropriate packaging with certain class of products, availability of other similar products, *etc.*).

The *bottom-up* approaches, such as *actor model of computation* [1] and multi-agent systems [23], capture the micro-behaviours of a system using a collection of interacting *actors* [1, 14] or *agents* [23] (henceforth *actor*) and help to observe emergent macro-behaviour at a higher level of abstraction. The agent and actor based technologies, such as Erlang [4] and Akka [2], realise system as set of *autonomous*, *self-contained*, and *reactive actors*.

4 Solution considerations and proposed approach

The proposed framework contains a RL agent based controller and two control loops as shown in Fig. 3. The model-based simulation loop helps to train RL agent and evaluate of new policies prior to their implementation in a real system, and real time control loop controls the real system using tranned RL agent. As shown in the figure, the controller decides an action based on its policy, *state* of the system and observed *rewards*. The model-based simulator consumes an action, which is produced by controller, as an external *event* and derives its impact by computing the *state* and *rewards* when a specific action is applied to the model. The model-based simulation loop iterates over multiple such actions to complete the necessary training. We adopt actor based simulation to specify the micro-behaviours of a system and compute emerging macro behaviours (*i.e.*, overall system *state* and *rewards*).

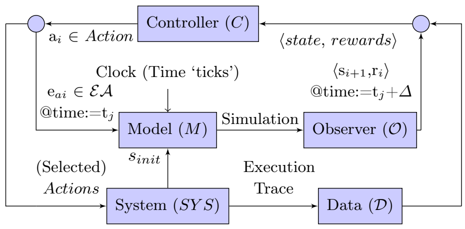


Fig. 3. Proposed approach.

A meta-model to represent systems using an extended form of *actor* is shown in Fig. 4. Here, a system is a set of `Actors`, whose characteristics can be described using a set of variables or `Properties`. Each `Actor` has its own `State` and `Behaviour`. They interact with other `Actors` by consuming and producing `Events`, where an incoming (*i.e.* consumed) `Event` may trigger a `Behaviour` unit that can change the state of an `Actor`, send `Events` and create new `Actors`.

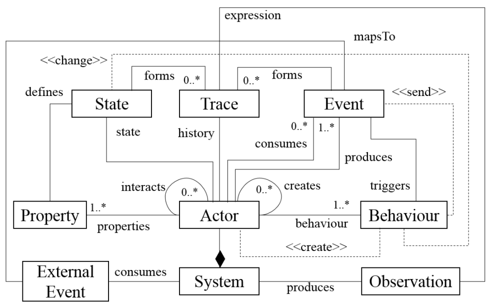


Fig. 4. Meta model to represent complex system using agents termed as 'actor'.

We extend this canonical form of an `Actor` with a notion of `Trace` (a sequence of information about `State` and `Events`) and an explicit construct to describe uncertainty in behavioural specification (as shown in Fig. 5). Formally, an actor (ACT) is a five-tuple: $\langle \mathcal{S}$, $\mathcal{EC}$, $\mathcal{EP}$, Tr, $\mathcal{B}\rangle$, where

- $\mathcal{S}$ A set of labels that represent `States` of an `Actor`.
- $\mathcal{EC}$ A finite set of `Events` that an `Actor` can consume.
- $\mathcal{EP}$ A finite set of `Events` that an actor can produce. Here, $\mathcal{EC}$ and $\mathcal{EP}$ are not disjoint set (`Events` $\mathcal{EC} \cup \mathcal{EP}$ are consumed within an actor).
- Tr A finite sequence of a tuple, where each tuple captures consumed `Event`, corresponding `State` and produced `Event`, *i.e.*,

stmt ::= **become**(state$_{new}$)   State change of an actor
| **send**(event$_i$, ACT_k)   Send event to an actor
| **create** ACT(state$_{init}$)   Create new actor
| e$_1$:stmt$_1$+...+e$_n$:stmt$_n$ Guarded statements
| **probably**(e$_{prop}$):stmt   Probabilistic statement

Fig. 5. Abstract Syntax of Behavioural Statements.

$\mathrm{s}_0 \xrightarrow{\mathrm{ec0}} \langle \mathrm{s}_1, \mathrm{E}_{ps1} \rangle \xrightarrow{\mathrm{ec1}} \langle \mathrm{s}_2, \mathrm{E}_{ps2} \rangle ... \xrightarrow{\mathrm{ec(k\text{-}1)}} \langle \mathrm{s}_k, \mathrm{E}_{pk} \rangle$, where $\{\mathrm{e}_{c1}, ..., \mathrm{e}_{c(k-1)}\} \in \mathcal{EC}$ and $\mathrm{E}_{ps1}, ..., \mathrm{E}_{pk} \subset \mathcal{EP}$

$\mathcal{B}$ A set of behavioural units. We consider that every behavioural unit $B \in \mathcal{B}$ is a set of programs that contain a sequence of stochastic statements. An abstract syntax to specify these programs is presented in Fig. 5.

In conformance with the meta model presented in Fig. 4, the system can be formally described as a quadruple $M = \langle \mathcal{ACT}, \mathcal{EA}, CLK, \mathcal{O} \rangle$, where $\mathcal{ACT}$ is a finite but dynamic set of actors; $\mathcal{EA}$ is a fixed and finite set of external `Events`, which are triggered from external sources; CLK is a clock that triggers virtual time `Events` or simulation 'ticks' (as external events) to progress simulation; and $\mathcal{O}$ is a set of observations. An observation is a tuple $\langle AS, Fact \rangle$, where AS is a set of actor states and $Fact$ are temporal expressions on actor traces (*e.g.*, occurrence of events over time). Two critical components of the control setup are described below.

1. Computation of $\mathcal{O} \rightarrow \mathcal{A}$: The observations $\mathcal{O}$ consist of the inventories at time $(t - \Delta)$, the order forecast (expected depletion) $f_{i,j}$ in the next time period, and product meta-data v_i, w_i, and l_i. There are five input variables per product, leading to a total input size of $5kn$. The output size is kn, and each output can take any value between 0 and $X_{i,j}$. The number of variables in computing such a mapping directly using RL is infeasibly large. Therefore, we compute this mapping iteratively, one product at a time. We use a form of RL known as Advantage Actor Critic (A2C) [21] to compute the actions (not to be confused

Algorithm 1 Compute observations using Actor Simulation

procedure SIMULATE(M_{init}, D, Δ)
  Duration: D, observation interval: Δ, $interval$:= 0, $time$:= 0, state := $Active$
  $\forall$ $ACT_i \in \mathcal{ACT}$: create $ACT_i(s_0)$ ▷ Initiate simulation by instantiating actors
  while (time != D) **do** ▷ Simulate for D time duration
    receive($event_{ext}$) ▷ M receives an external event
    if ($event_{ext}$ is 'tick') **then** ▷ If event is a time event
      $time := time + 1$
      if (state := $Active$) **then**
        $\forall$ $a_i \in \mathcal{ACT}$: send($event_{ext}$,a_i) ▷ Broadcast time event
      if ($interval = \Delta$) **then**
        $interval := 0$
        O := observe(M) ▷ Compute O from all S and Tr of $\mathcal{ACT}$
        notify(O, C) ▷ Notify $\langle state_i, reward_i \rangle$ to C
        state := $Pause$ ▷ Pause time event for RL agent
      else
        $interval := interval + 1$
    if ($event_{ext} \in \mathcal{EA}$) **then** ▷ If event is a RL Agent action
      state := $Active$ ▷ Restart time event for next observation
      for $a_i \in \mathcal{ACT}$ (of M) **do**
        if ($\exists$ e_x such that $e_x \in \mathcal{EC}$ of ACT_i **And**
          map(e_{ext},$event_x$) $\in \mathcal{MAP}$) **then**
          send(e_x,a_i) ▷ Send event to relevant actors

with 'actor' in the simulation context) . The Actor is a neural network with 5 inputs and 11 outputs, representing discretised and normalised actions between 0 and 1. The Critic is a neural network with the same 5 inputs but a single output representing the expected *value* of the current observations. The basic reward function given in (1) was augmented for the purposes of training the agent, by a penalty on any actions that could not be feasibly implemented in the system because of capacity exceedance. This allows the RL agent to relate the effect of individual decisions on the net system reward.

2. Computing $\mathcal{A} \rightarrow \mathcal{O}$: The updates to $O \in \mathcal{O}$, the actor states, and events, are computed through simulation. As shown in Algorithm 1, all actors of an initial model (M_{init}) are instantiated to a random state or a system state to initialise a simulation. Actors execute their behaviours in their respective threads, interact with each other through actor events, change their states (possibly several times to respond to external and internal events) and create new actors. The external events that include time 'tick' and events corresponding to the RL actions are propagated to all relevant actors and allowed them to react for time duration Δ before the emergent states and traces are observed. The observed and computed O is then notified to the controller for the *next* RL action.

5 Illustration, validation and discussion

We use an actor language [8] that supports proposed actor abstraction to model a supply chain network as shown in Fig. 6. The simulation dashboard depicting the emerging reward related factors in a hypothetical retail chain is shown in

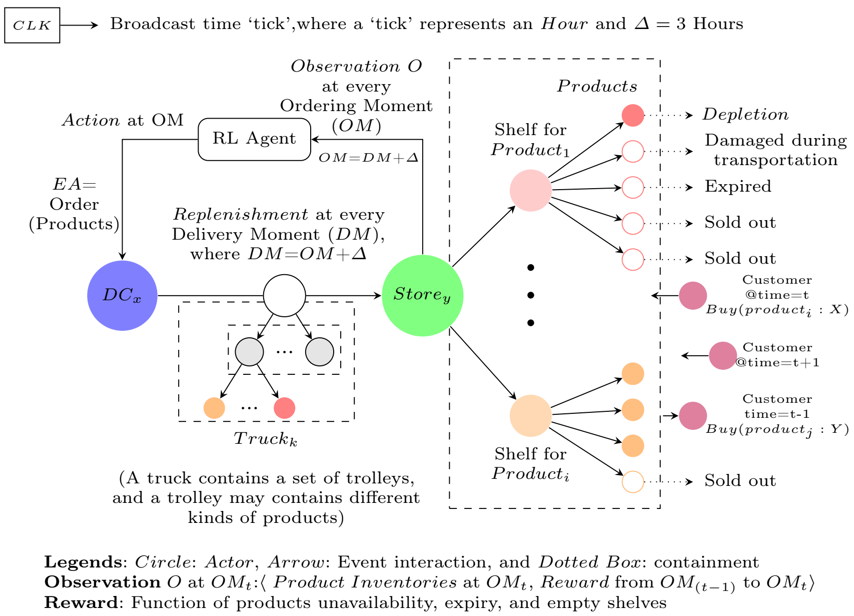


Fig. 6. An implementation of supply chain replenishment case study described in Fig. 2

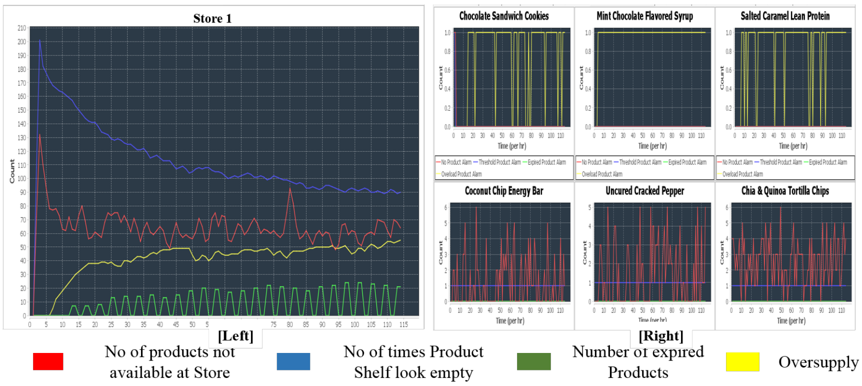


Fig. 7. [Left] Instance trace of product unavailability, emptiness of shelves, product wastage and product over-supply for a shop. [Right] Traces of individual products.

Figure 7. We use a data set spanning one year derived from a public source [18]. A total of 220 products were chosen from the data set, and their meta-data (not available originally) were defined manually. A single store and a single truck are used in the results presented here (see Sec 6 for extensions). Forecasts were computed using a uniform 10-step trailing average for each product. The store capacity, truck volume and weight capacity, and labour counts were computed based on the order volumes seen in the data.

The time between successive delivery moments set to 6 hours (leading to 4 DM per day). The lead time Δ was 3 hours. The truck volume constraint was set lower than the average order numbers in order to stress the system. The initial normalised inventory level for each product is set to 0.5, and the level below which penalty is imposed is set to $\rho = 0.25$. Of the order data set, the first 225 days (900 DM) were used for training, while the remaining 124 days (496 delivery moments) were retained for testing. Fig. 8 [Left] shows the training over 15 episodes, each spanning the 900 DM in the training data set. The average reward over all 220 products is seen to increase as training proceeds. The reward is compared with an industry-standard replenishment heuristic adapted from prior literature [9]. We see that the reward at the end of training exceeds the heuristic performance, and this advantage is retained on the test data set as well (plotted using separate markers at the ends of the curves). Also shown in Figure 8 [Left] is the 'exploration rate' of RL, which is the probability with which the RL algorithms takes randomised actions.

The performance advantage is due to the nature of Fig. 8 [Right], which plots the inventory levels of products on the test data set (496 delivery moments). Both algorithms begin with an initial (normalised) inventory level of 0.5 for all products. However, RL is able to maintain a higher average inventory level than the heuristic. The characteristics of the Critic and Actor networks of the RL agents are illustrated in Fig. 9. The value attached by the Critic network is shown in Figure 9 [Left], as a function of the inventory level (after averaging over all other feature values). The peak value is near the penalty threshold $\rho = 0.25$. The value drops quickly below this level. There is also a decrease in the estimated

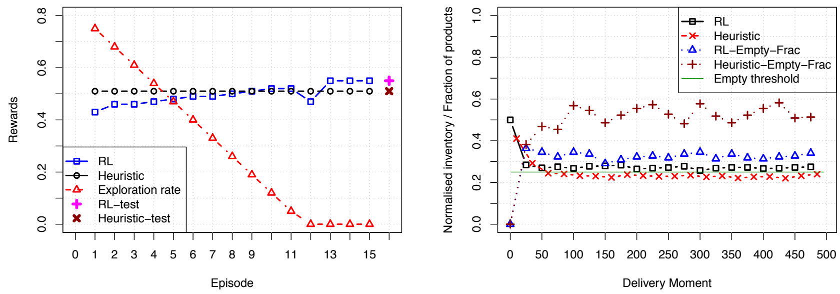


Fig. 8. [Left] Evolution of rewards during training, in comparison with the heuristic algorithm. [Right] Trend of average inventory levels across all 220 products over time.

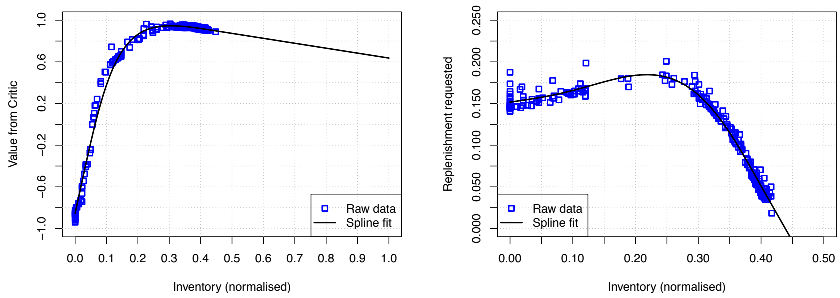


Fig. 9. [Left] Estimate of state value as a function of inventory. [Right] Replenishment requested by the RL agent as a function of current inventory level.

value at very high inventory levels, due to the higher rates of product wastage. Fig. 9 [Right] shows the replenishment quantity requested by the Actor network, again as a function of the inventory level of the product after averaging over all other features. We note that the requested replenishment rises as inventory levels drop from 0.5 towards the penalty threshold of 0.25.

6 Conclusion

We described the use of a realistic closed loop multi-agent simulation model for training a reinforcement learning based control policy, as opposed to the traditional use of analytical expressions for rewards. Initial tests show that training using proposed approach is both feasible and effective. The use of the proposed actor based simulation as an environment to understand the overall implication of multiple RL actions (produced for different parts of a network) and locally optimised solutions for subsystems in a global system context, can also be viewed as a viable option. The ultimate goal of this work is to develop a closed-loop simulation and reinforcement learning framework, that allows us to deploy the trained agent on a real system with minimal subsequent adjustments.

A trained version of the reinforcement learning algorithm for computing replenishment orders is expected to become operational in a grocery retail network with approximately 10^3 stores and 10^5 products per store, in December 2019. The current setup has been tested on a single store scenario with nearly 10^4 products, which generates approximately 4×10^5 Actors in the simulation. The challenge in the coming year is to extend the capability to full system simulation while retaining computational feasibility. We believe this is feasible, based on the following considerations. First, the current implementation of the simulation and learning loop works on a single laptop. There is thus scope for increasing the computational power as necessary. Second, the decision-making portion ($\mathcal{O} \rightarrow \mathcal{A}$ map) works independently for each product, allowing us to parallelize the online workflow. Finally, it may be possible to loosely partition complex supply chain networks into sub-networks, further reducing the computational complexity.

References

1. Agha, G.A.: Actors: A model of concurrent computation in distributed systems. Tech. rep., DTIC Document (1985)
2. Allen, J.: Effective akka. O'Reilly Media, Inc. (2013)
3. Anderson, P.: Perspective: Complexity theory and organization science. Organization science **10**(3), 216–232 (1999)
4. Armstrong, J.: Erlang - a survey of the language and its industrial applications. In: Proc. INAP. vol. 96 (1996)
5. Barabási, A.L., Bonabeau, E.: Scale-free networks. Scientific american **288**(5), 60–69 (2003)
6. Bouabdallah, S., Noth, A., Siegwart, R.: Pid vs lq control techniques applied to an indoor micro quadrotor. In: Proc. of The IEEE International Conference on Intelligent Robots and Systems (IROS). pp. 2451–2456. IEEE (2004)
7. Caro, F., Gallien, J.: Inventory management of a fast-fashion retail network. Operations Research **58**(2), 257–273 (2010)
8. Clark, T., Kulkarni, V., Barat, S., Barn, B.: ESL: An actor-based platform for developing emergent behaviour organisation simulations. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 311–315. Springer (2017)
9. Condea, C., Thiesse, F., Fleisch, E.: Rfid-enabled shelf replenishment with backroom monitoring in retail stores. Decision Support Systems **52**(4), 839–849 (2012)
10. Duan, Y., Andrychowicz, M., Stadie, B., Ho, J., Schneider, J., Sutskever, I., Abbeel, P., Zaremba, W.: One-shot imitation learning. In: Proc. of Conference on Neural Information Processing Systems (NIPS). vol. 31 (2017)
11. Gabel, T., Riedmiller, M.: Distributed policy search RL for job-shop scheduling tasks. International Journal of Production Research **50**(1), 41–61 (2012)
12. Giannoccaro, I., Pontrandolfo, P.: Inventory management in supply chains: a reinforcement learning approach. International Journal of Production Economics **78**(2), 153–161 (2002)
13. Godfrey, G.A., Powell, W.B.: An ADP algorithm for dynamic fleet management, i: Single period travel times. Transportation Science **36**(1), 21–39 (2002)
14. Hewitt, C.: Actor model of computation: scalable robust information systems. arXiv preprint arXiv:1008.1459 (2010)

15. Holland, J.H.: Complex adaptive systems. Daedalus pp. 17–30 (1992)
16. Iacob, M., Jonkers, D.H., Lankhorst, M., Proper, E., Quartel, D.D.: Archimate 2.0 specification. Van Haren Publishing (2012)
17. Jiang, C., Sheng, Z.: Case-based reinforcement learning for dynamic inventory control in a multi-agent supply-chain system. Expert Systems with Applications **36**(3), 6520–6526 (2009)
18. Kaggle: Instacart market basket analysis data. `https://www.kaggle.com/c/instacart-market-basket-analysis/data` (Retrieved August 2018)
19. Khadilkar, H.: A scalable reinforcement learning algorithm for scheduling railway lines. IEEE Transactions on ITS (2018)
20. Kober, J., Bagnell, J.A., Peters, J.: Reinforcement learning in robotics: A survey. The International Journal of Robotics Research **32**(11), 1238–1274 (2013)
21. Konda, V.R., Tsitsiklis, J.N.: Actor-critic algorithms. In: Advances in neural information processing systems. pp. 1008–1014 (2000)
22. Lee, H.L., Padmanabhan, V., Whang, S.: Information distortion in a supply chain: The bullwhip effect. Management science **43**(4), 546–558 (1997)
23. Macal, C.M., North, M.J.: Tutorial on agent-based modelling and simulation. Journal of simulation **4**(3), 151–162 (2010)
24. Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.: Constrained model predictive control: Stability and optimality. Automatica **36**(6), 789–814 (2000)
25. Meadows, D.H., Wright, D.: Thinking in systems. Chelsea Green Publishing (2008)
26. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.: Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602 (2013)
27. Mortazavi, A., Khamseh, A.A., Azimi, P.: Designing of an intelligent self-adaptive model for supply chain ordering management system. Engineering Applications of Artificial Intelligence **37**, 207–220 (2015)
28. Powell, W.B.: Ai, or and control theory: A rosetta stone for stochastic optimization. Princeton University (2012)
29. Russell, S.J., Norvig, P.: Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited, (2016)
30. Sabri, E.H., Beamon, B.M.: A multi-objective approach to simultaneous strategic and operational planning in supply chain design. Omega **28**(5), 581–598 (2000)
31. Shervais, S., Shannon, T.T., Lendaris, G.G.: Intelligent supply chain management using adaptive critic learning. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans **33**(2), 235–244 (2003)
32. Silver, E.A.: Operations research in inventory management: A review and critique. Operations Research **29**(4), 628–645 (1981)
33. Simon, H.A.: The architecture of complexity. In: Facets of systems science, pp. 457–476. Springer (1991)
34. Smith, S.A., Agrawal, N.: Management of multi-item retail inventory systems with demand substitution. Operations Research **48**(1), 50–64 (2000)
35. Sutton, R., Barto, A.: Reinforcement learning. MIT Press, 2 edn. (2012)
36. Thomas, M., McGarry, F.: Top-down vs. bottom-up process improvement. IEEE Software **11**(4), 12–13 (1994)
37. Topaloglu, H., Powell, W.B.: Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. INFORMS Journal on Computing **18**(1), 31–42 (2006)
38. Valluri, A., North, M.J., Macal, C.M.: Reinforcement learning in supply chains. International journal of neural systems **19**(05), 331–344 (2009)
39. White, S.A.: BPMN modeling and reference guide (2008)

Agent based simulation of the dengue virus propagation

L. da Silva Rodrigues[1], S. G. Mamedio Oliveira[1][0000-0003-1409-5103], L. F. Lopez[1], and J. S. Sichman[1][0000-0001-8924-9643]

(1) Universidade de São Paulo (USP), Brazil
{leticia2.rodrigues,sostenes.mamedio,lopez,jaime.sichman}@usp.br

Abstract. This work aims to implement a model to simulate the dengue virus propagation, which is one of the main public health problems in Brazil. In order to do it, we adopted a multi-agent based simulation (MABS) approach. The agent model is inspired by the idea of compartments, widely used in classical models of epidemiology. The model was implemented in the GAMA platform, as well as a classical model based on ordinary differential equations. Although adopting some simplifying assumptions, comparing the output of the two models made it possible to validate our approach and to indicate that our model may serve in the future as a basis for the development of more refined models.

Keywords: Dengue · Epidemiological models · Multi-agent based simulation.

1 Introduction

The main purpose of this work is to design and implement a multi-agent-based model of the dengue virus' propagation, and to compare the model results with some traditional deterministic epidemiological models. There are two benefits from modeling an epidemic spread using simulation: to understand the mechanisms of propagation and to predict how the disease spread will behave in the future, given a current state and possible sanitary actions. Moreover, we seek to understand the effect of local conditions, like human typical trajectories, in the disease spread. A multi-agent based simulation approach perfectly fits these requirements, due to its dynamics, and this is the reason why it has already been used by epidemiologists in many studies.

Next section gives a brief introduction to dengue, followed by the the most common approaches to simulate epidemics in section 3. Then, in section 4 we detail our model. Implementation details and experimental results of its application are discussed in section 5. Finally, we present in section 6 our conclusions and further work.

2 Dengue

The World Health Organization (WHO) estimates that 40% of the global population is exposed to the dengue virus [10]. In Brazil, the first documented case was in 1923, in Rio de Janeiro [7]. Although in most parts of Brazil there has been a decrease in the number of reported cases, in the state of São Paulo there was a big increase from 2017 (4044 cases) to 2018 (8979 cases), according to data from Sinan (Sistema de Informação de Agravos de Notificação), a Brazilian institution that collects this kind of data.

The government and other organizations have been struggling in containing the spread of the virus, since its behavior depends on the its location, making it difficult to correctly apply the most effective policies against the mosquito and the virus infection. The main measures promoted are usually campaigns explaining how to avoid mosquito bites, how to prevent its reproduction and also how to take the vaccine shots. However, the mosquito and the population behaviors differ not only from one country to another, but also from rural to urban areas [6]. Therefore, it is indispensable the existence of tools that make possible to estimate the impact of a given sanitary action in a certain area. In this context, many epidemiological models are being developed to assist in the comprehension of the spread and resurgence of the virus and hence helping in decision making.

The dengue virus needs a biological vector to be transmitted. The main vector of dengue is the mosquito of the genus Aedes and the main species involved in the transmission in the West is the Aedes Aegypti. The disease is considered tropical because its proliferation is favored by the hot and humid climate, which are the ideal conditions for the vector reproduction. It is present in several regions of the world, such as Africa, Asia and the Americas [10]. The probability of occurrence of this disease around the globe is shown in Figure 1.

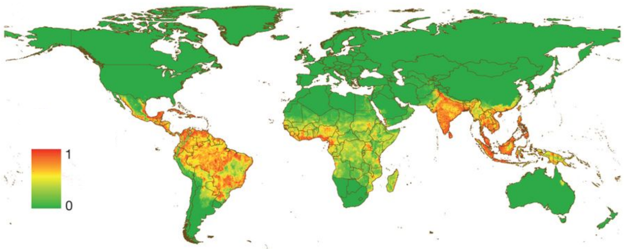


Fig. 1. Map shoeing the likelihood of getting dengue. The closer to 1 (orange), more exposed is someone who lives there. Source: [3].

The mosquito that transmits dengue and is well adapted mainly to the urban environment [7], it has a little less than 1cm in length and its bite does not cause pain or itching reactions. The female of this species is responsible for the transmission of the virus, since the male does not feed on blood.

The mosquito has a habit of biting in the early hours in the morning and in the late afternoon and usually does not get too far from its place of birth [9]. The mosquito is infected by biting an infected person as the virus can not be transmitted between mosquitoes or between humans. Symptoms include high fever, headaches, muscle aches, among others, and may progress to hemorrhagic fever and shock syndrome in more severe cases, where the patient may even die [6].

3 Epidemiological models

Epidemiological models usually try to answer two questions: how a certain disease spreads and how it reappears in places not even considered in first place as susceptible ones. However, all models present a limitation regarding the reality, as they commonly try to represent a phenomenon either in a global (macro) or local (micro) scale.

For either of these scales, we can use the compartments theory. The idea of compartments is to represent the states of individuals. The most common types, which will be used in this project are susceptible (S), exposed or latent (E), infected (I) and recovered (R). Individuals in the state S are not carriers of the disease, but can be infected. In state E, they carry the virus but do not transmit it or show symptoms yet, which will only appear when they reach the state I, when they can also transmit the infection. In the R state, individuals are cured and immune to new infections. A number of different models, that consider different types of compartments, were proposed in the literature, such as SIR, SEI, SEIR, SEIRS, among others. The flow in the compartments happens in the order in which the letters are displayed in the model name. The model to choose will depend on the intrinsic characteristics of a disease. For instance, a SIR model states that individuals begin susceptible to the disease, then become infected and finally are recovered. The transition to the next state is usually not mandatory and depends on internal parameters of the model, which tries to resemble the reality.

3.1 Macro simulation models

This type of simulation focuses on a global level representation and makes use of mathematical equations, hence being a deterministic approach. This kind of approach ignores the individuals characteristics and the interaction among them. For instance, information such as age, sex, address and others cannot be used. According to [6], when using such a macro model there must be awareness concerning these limitations, since some of the individual characteristics may be of crucial relevance to the spread of some diseases. On the other hand, they are easier to implement and interpret. In [4], the authors concluded that it was not necessary to extinguish the vector in order to stop the dengue virus propagation.

The classical mathematical models are mostly based on compartments theory, and the change between one state to another is given by a system of ordinary differential equations (ODE).

3.2 Micro simulation models

In contrast with macro simulation models, in micro simulation models each individual is represented separately. The micro simulation approach adopted in this work is multi-agent-based simulation (MABS) [8]. A multi-agent system generally represents a complex system that has some characteristics such as non-linearity and multiple levels of abstraction. Such systems contains agents that perceive the environment and act on them. These agents can be merely reactive to situations and also could take decisions based on their cognitive abilities. In MABS, there is no predefined algorithm that can predict the global behavior of the system. Hence, the interactions of agents at the individual lead to an emergent global structure. Therefore, the simulation needs to be performed several times, since it is non-deterministic. From the analysis of the distribution of these various results, it becomes possible to conclude something about the overall behavior of the system. Since the relationship between the inputs and outputs of MABS systems can not be explicitly defined, it is hard to verify and validate such models. In [2] and [6], the authors opted to consult a specialist in the field of epidemiology to validate the models.

Some previous works have reported multi-agent systems approaches in epidemics. In [6], the authors created a module able to trace a relationship between the spread of dengue virus and commercial routes in Asia. In [2], a model that encapsulated the characteristics of the north of Vietnam was used to describe the spread of the H1N5 virus. Hence, for the study of dengue in Brazil, it seems reasonable to use the same approach.

4 Our approach

The main purpose of this work is design and implement a MABS to represent the spread of dengue virus, and to compare and validate its results with a mathematical macro simulation model. Both models will follow the compartments theory approach [1], in which we consider the SIR model for human beings, and the SEI model for the mosquitoes. Moreover, we are not considering any other life forms of the mosquito other than the adult, like eggs or worms. In addition, the birth and death rates during the simulation are considered to be zero, keeping the same size of mosquito and human populations from the beginning to the end. Figure 2 illustrates the compartment schematics for this study.

It can be observed that the passage of humans from susceptible to infected, and mosquitoes from susceptible to exposed, does not depend on those infected from their own species, but rather from the other one.

4.1 Mathematical model

The equations set 1 describes the behavior of the model, represented in Figure 2:

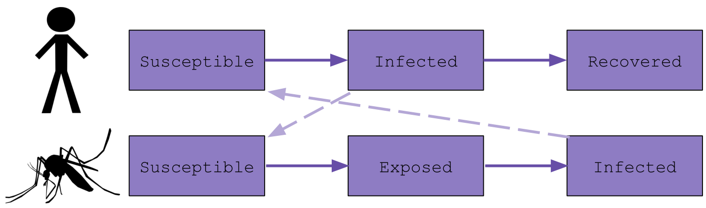


Fig. 2. SIR and SEI compartment flowchart for Human and Mosquito agents.

$$
\begin{cases}
\frac{dS_H}{dt} = -abS_H \frac{I_M}{N_H} \\
\frac{dI_H}{dt} = abS_H \frac{I_M}{N_H} - \gamma I_H \\
\frac{dR_H}{dt} = \gamma I_H \\
\frac{dS_M}{dt} = -acS_M \frac{I_H}{N_M} \\
\frac{dE_M}{dt} = acS_M \frac{I_H}{N_M} - \gamma E_M \\
\frac{dI_M}{dt} = \gamma E_M \\
N_H = S_H(t) + I_H(t) + R_H(t) \\
N_M = S_M(t) + E_M(t) + I_M(t)
\end{cases}
\tag{1}
$$

where S, I, E and R represent the number of individuals in a given time t, whether human (H) or mosquito (M), in the respective compartment symbolized by the letter. Table 1 describes the other model parameters, as well as their chosen values.

Table 1. Parameters of the model, their biological meanings and values.

Parameter	**Meaning**	**Value**
a	Daily rate of bites	0,168
b	Fractions of infectious bites	0,6
γ_H	Human recovery rate per day	0,143
γ_M	Mosquito latency rate per day	0,143
c	Mosquito susceptibility to dengue	0,526

These values were extracted from [1] and [4]. The human recovery rate corresponds to the mean time that a person recovers from the virus. The latency rate of mosquitoes is the average time to pass to the infected state.

4.2 MABS model

In the MABS model, there are two types of agents: the mosquito and the human agent, each one with its own set of actions, described next.

i) Mosquito agent The mosquito agent is able to:

- *Move*: each mosquito moves in a random direction, and travels a random distance from its current location. Its position will always be close from the starting point.
- *Change to exposed state*: the change from the susceptible to the latent state can only occur if there are infected humans in the vicinity of the mosquito. The larger this number, the greater the chance to change its state. Being e the average bite rate per day and c the probability of a mosquito to be infected with the virus, we have $a * c$ the probability of the mosquito to bite an infected human and get infected as well. In addition, n represents the number of humans infected near the mosquito. Therefore, a mosquito changes from susceptible to exposed state with a probability p given by:

$$p = 1 - (1 - a * c)^n$$

- *Change to infected state*: the change from exposed to infected state depends only on the parameter γ_M, and the probability p for this to occur is:

$$p = \gamma_M$$

ii) Human Agent The human agent is able to:

- *Move*: the displacement of humans aims to represent a common routine, from home to another destination, which may be school, work, among others. So every human has a residence and a destination and moves between these two points.
- *Change to infected state*: similar to the mosquito agent, changing the susceptible state to the infected state in humans can only occur if there are mosquitoes infected nearby. The higher that number, the greater the chance of this transition to occur. Being a the average bite rate per day and b the likelihood of a human being infected by getting bitten by a mosquito with the virus, we have $a * b$ the likelihood of a human getting bitten by an infected mosquito and becoming infected. In addition, n represents the number of mosquitoes infected nearby. The probability p of changing from susceptible to infected state is given by:

$$p = 1 - (1 - a * b)^n$$

- *Change to recovered state*: the change from infected to recovered state depends only on the parameter γ_H, and the probability p for this to occur is given by:

$$p = \gamma_H$$

4.3 Simulation cycle

Each simulation cycle represents a duration of $12h$ in real time. In the mathematical model, the number of humans in each of the states (susceptible, exposed or infected) is calculated at each cycle. In the agent-based model, at each cycle all agents perform their specific actions. In addition, the human agent can also have its destiny changed (from residence to school/work or vice-versa). The destination of humans is represented by a third type of agent, called immobile agent. Thus, at each cycle, an immobile agent checks if there is any human agent within its area. If it is the case, the immobile agent changes the objective destination of the human agent for the next cycle.

5 Implementation and experiments

5.1 Technical details

The implementation of the models was developed in the GAMA platform, which is specifically designed for MABS [5]. This platform makes it possible to visualize geographically the behavior of the agents on a map, as shown in Figure 3.

Fig. 3. A view in the GAMA platform using a simulated geographical localization.

In the map, green means susceptible, yellow represents exposed, red indicates infected and blue means recovered. In order to analyze the output data from both models, we used R as the programming language in RStudio. After every simulation, we also validated the model with some specialists in the field of epidemiology.

5.2 Experiments

Three tests were performed, varying only the initial number of infected mosquito-type agents at the beginning of the simulation. Tables 2 and 3 show the initial amounts of each agent that were used in the simulations.

The comparative analysis was based considering the number of infected human agents.

Table 2. Model input variables for the Human agent.

Human (State)	**Initial amount**
Susceptible	495
Infected	5
Recovered	0

Table 3. Model input variables for the Mosquito agent.

Mosquito (State)	Initial amount		
	Test 1	Test 2	Test 3
Susceptible	9000	9900	9990
Exposed	0	0	0
Infected	1000	100	10

Scenario 1: High rate of infected of mosquito agents In this first test, 10% of all mosquitoes were infected from the start. Figure 4 shows the simulation results for the mathematical model (ODE) and Figure 5 for the multi-agent based simulation model (MABS).

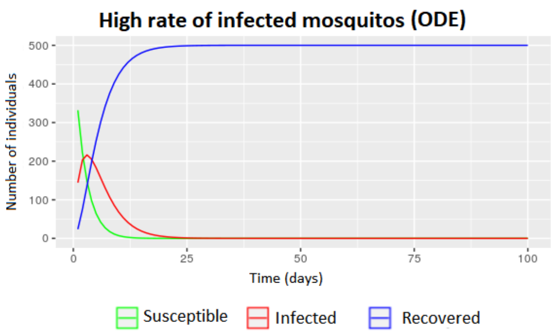


Fig. 4. Result of the ODE model with high rate of infected mosquitoes.

In Figure 5, as in the next ones to be displayed for the other tests, the full line represents the mean of the simulations and the dashed lines, the standard deviation. In the case of the ODE simulation, shown in Figure 4, the full lines show the result of the equations set 1. We can observe that in this experiment, all the curves are similar, indicating that all individuals were infected and later entered the recovered state. The peak that occurs is characteristic of an epidemic that occurred in both models. Thus, the global epidemic behavior of the two models may be considered equivalent.

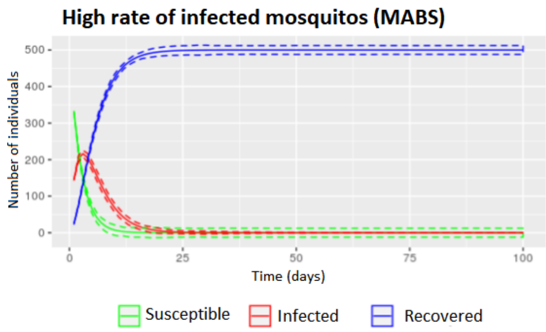


Fig. 5. Result of the MABS model with high rate of infected mosquitoes.

Scenario 2: Medium rate of infected of mosquito agents In this second test, only 1% of all mosquitoes were infected from the start. Figure 6 shows the simulation results for the mathematical model (ODE) and Figure 7 for the multi-agent based simulation model (MABS).

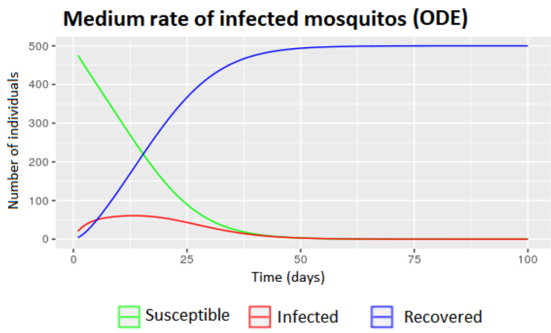


Fig. 6. Result of the ODE model with medium rate of infected mosquitoes.

The main difference between curves lies in the number of susceptible and recovered individuals. It is observed that all individuals pass from the first to the last state in the ODE model, but not in the MABS model. Observing the equations set 1, which represent the compartments, it is easy to see why this occurs in the ODE model. As shown in Figure 2, which represents this model, for the human agents there is no exit transition from the recovered state . Hence, all individuals arriving in that state remain in it. Combining this with the fact that there are no entries of new individuals nor exits during a simulation, and that the transitions occur towards the recovered state, we have that at some point all human agents will reach that state.

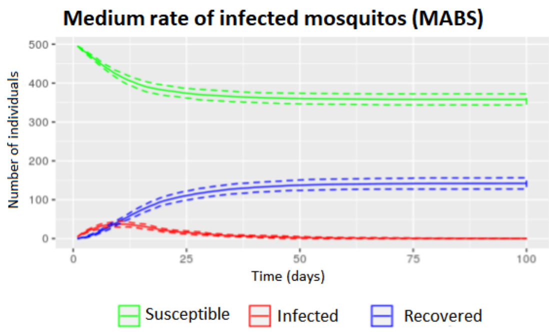


Fig. 7. Result of the MABS model with medium rate of infected mosquitoes.

The same behavior, however, does not occur in the MABS model. One possible justification is due to the agents' movement behavior and the possibility of including their location information at any time of the simulation. As they always move within the same region, vectors do not always spread the disease to uninfected areas nor do humans contaminate mosquitoes from other regions after a certain interval since the beginning of the experiment. This result is compatible with the one obtained by [6]. In this case, a positive correlation was observed between the increase in the number of dengue cases and the increase in commercial relations between some Asian countries. The hypothesis is that greater human displacements enable the virus to reach a larger region. Thus, by doing the opposite, which is, by limiting the area covered by the agents, we would be limiting also the spread of the virus. In addition, comparing the MABS results produced in this test with those produced in the first experiment, we can notice that the final amount of recovered humans is lower when the initial number of infected mosquitoes is also lower. This can be considered as an expected result, since fewer people were infected. In addition, one can see again an equivalence between the curves representing the infected individuals. There is a small increase in the number of cases at the beginning, which soon stabilizes.

Scenario 3: Low rate of infected of mosquito agents In this last test, the number of mosquitoes that were infected from the start is even smaller, with a value of 0.1%. Figure 8 depicts the simulation results for the mathematical model (ODE) and Figure 9 for multi-agent based simulation model (MABS).

As in the previous case, the fact that the movement of agents is limited to one region also limits the proliferation of the disease. This effect is even more evident in this test, since the number of susceptible and recovered humans has very little variation during the simulation. In addition, the curves of infected humans are also similar, with few cases of the disease, which is expected given the small number of infected mosquitoes.

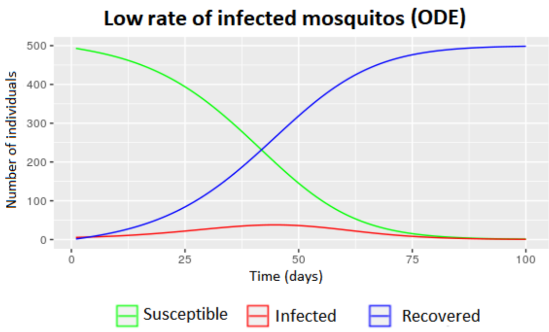


Fig. 8. Result of the ODE model with low rate of infected mosquitoes.

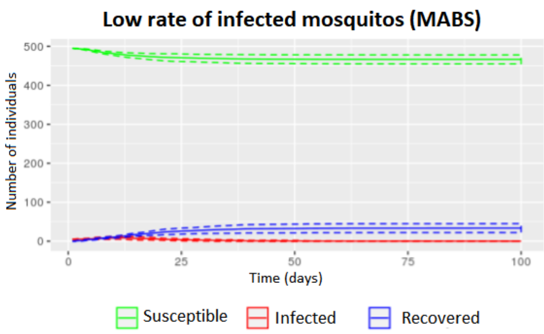


Fig. 9. Result of MABS model with low rate of infected mosquitoes.

6 Conclusions and further work

This work aimed to design a model of the spread of dengue virus using a multi-agent based simulation approach. The model was implemented, representing a basic and simplified version of how the virus propagates.

The comparison between the classical mathematical model and the multi-agent based model showed that the second technique can lead the system to the expected global behavior, since in the three studied cases, the profiles of the disease propagation (infected curves) were similar. Therefore, we obtained the same macro behavior when defining the interactions between the agents at the micro level.

This conclusion is significant because it indicates that multi-agent models can represent reality at least as well as the classical models. Moreover, as their capacity of representation is much more detailed, they can take into account the heterogeneity of the population and characteristics of the environment, among others. Hence, agent-based models have a greater chance of being more faithful than models solely based on mathematical equations.

Although being able to reach this conclusion, the current model is still very simplified. Taking into account its current state and drawing on the bibliography, the following items are proposed for a possible future evolution of the project: *(i)* inclusion of birth and death rates in humans and mosquitoes, *(ii)* inclusion of other mosquito life forms, such as eggs and larvae, *(iii)* use of real geographical data, *(iv)* use of geographical information (such as urban or rural areas) to adapt the mechanisms of propagation and *(v)* variation of the mosquito population due to the season.

Thus, we consider this work as an initial contribution to public policies, aiming to obtain in the future a model closer to the reality, and that can serve as a decision tool to mitigate the spread of this disease.

References

1. Amaku, M., Azevedo, F., Burattini, M., Coutinho, F., Lopez, L., Massad, E.: Interpretations and pitfalls in modelling vector-transmitted infections. Epidemiology & Infection **143**(9), 1803–1815 (2015)
2. Amouroux, E., Desvaux, S., Drogoul, A.: Towards virtual epidemiology: an agent-based approach to the modeling of h5n1 propagation and persistence in north-vietnam. In: Pacific Rim International Conference on Multi-Agents. pp. 26–33. Springer (2008)
3. Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., et al.: The global distribution and burden of dengue. Nature **496**(7446), 504 (2013)
4. Burattini, M., Chen, M., Chow, A., Coutinho, F., Goh, K., Lopez, L., Ma, S., Massad, E.: Modelling the control strategies against dengue in singapore. Epidemiology & Infection **136**(3), 309–319 (2008)
5. Grignard, A., Taillandier, P., Gaudou, B., Vo, D.A., Huynh, N.Q., Drogoul, A.: Gama 1.6: Advancing the art of complex agent-based modeling and simulation. In: International Conference on Principles and Practice of Multi-Agent Systems. pp. 117–131. Springer (2013)
6. Philippon, D., Choisy, M., Drogoul, A., Gaudou, B., Marilleau, N., Taillandier, P., Truong, Q.C.: Exploring trade and health policies influence on dengue spread with an agent-based model. In: International Workshop on Multi-Agent Systems and Agent-Based Simulation. pp. 111–127. Springer (2016)
7. Pontes, R.J., Ruffino-Netto, A.: Dengue em localidade urbana da região sudeste do brasil: aspectos epidemiológicos. Revista de Saúde Pública **28**, 218–227 (1994)
8. Sichman, J.S.: Operationalizing complex systems. In: Bernardo Alves Furtado, P.A.M.S.e.M.H.T.e. (ed.) Modeling Complex Systems for Public Policies, pp. 85–123. IPEA, Brasilia (2015)
9. Silva, J.S., Scopel, I., et al.: A dengue no brasil e as políticas de combate ao aedes aegypti: da tentativa de erradicação às políticas de controle. Hygeia: Revista Brasileira de Geografia Meédica e da Saúde **4**(6) (2008)
10. Singhi, S., Kissoon, N., Bansal, A.: Dengue and dengue hemorrhagic fever: management issues in an intensive care unit. Jornal de Pediatria **83**(2), S22–S35 (2007)

Modeling Pedestrian Behavior Under Panic During a Fire Emergency

Juhi Singh, Atharva Deshpande, and Shrisha Rao

International Institute of Information Technology, Bangalore
{juhi.singh ,atharvaanant.deshpande}@iiitb.org
shrao@ieee.org

Abstract. This paper presents an integrated model that approximates pedestrian behavior in case of a fire emergency, and its consequences. We have modeled a confined fire with a variable spread rate, based on the existing literature pertaining to the field. Fire has both psychological and physical impacts on the state of the agents. The model also incorporates clustering behavior in agents, which slows down the evacuation. The model helps recognize bottlenecks and compares the evacuation efficiency by comparing casualties across different scenarios. Simulation results are given as illustrations, and give qualitative insights into the risks and likely problems in specific fire scenarios.

Keywords: pedestrian evacuation · compartment fire modeling · clustering · panic

1 Introduction

Large gatherings of people may need emergency evacuations due to sudden dangers. Unfortunately, such evacuations are quite difficult at the best of times, given the unpredictable nature of emergencies. Evacuations are even more complicated in case of fires, and such emergencies have caused mass casualties worldwide. For example, 117 people were killed in a fire at a garment factory in Bangladesh, in 2012 [1]. A fire in a nightclub killed 100 people in Rhode Island in the US in 2003 [3], and the fire in the residential Grenfell Tower in the UK killed over 70 in June 2017 [13, 12].

Since the conditions of such events are difficult to emulate, it is extremely difficult to obtain reliable data. This makes simulations indispensable for public safety. Agent-based modeling in particular has a role.

This work is based on the model of by Trivedi and Rao [19] who use the Boids model [15]. Their model incorporates the effects of panic on decision making of agents, for instance increase in panic clouds the ability of the agents to make rational decisions [14]. The panic experienced by agents is quantified based upon its distance from the exit, the velocity of neighbors heading toward the exit, count of nearby agents who have high degrees of physical discomfort, and the lag in velocity compared to its neighbors, etc [9]

The social force model is used to calculate the force on an agent from its neighboring agents [4].

We have improved upon the existing framework, by further integrating real-life crowd behavior, using clustering [9]. We consider a group of evacuees as a network, with the agents as the nodes. Weighted edges between any two nodes represent the strength of a relationship or attachment between those particular nodes. Using this, we are able to model clustering behavior of pedestrians in our simulation, based upon the visibility, and the strength of attachment between any two (or more) nodes. Therefore, nodes with high level of attachment, tend to cluster together instead of following the crowd path. This covers cases of families or friends clustering together, or people actively seeking out family members during emergency evacuations [4] [5] [6].

We further model a basic compartment fire, based on the fire model proposed by Bishop et al. [17]. We integrate the fire as a causative agent for the evacuation. We use MASON, a Java-based multiagent simulation library, for our simulations.

2 Fire Model

We add a radially spreading, spherical fire as the hazardous factor in our simulations.

2.1 Background

We consider a compartment fire (one confined to a compartment) to our model. The lifetime of a compartment fire is characterized by the following stages: [17]

- Pre-Flashover Stage: This is the growth stage of the fire. The fire grows rapidly and is mostly fuel controlled.
- Flashover Stage: This is the stage of sharp increase in the hot gas temperature and fire intensity. This stage presents a non-linear stage of growth that leads to imminent disaster.
- Fully Developed Fire: This stage marks the fire reaching its maximum potential and engulfing the entire room.
- Decay Stage: This stage involves the period of decay of the fire, where the fire intensity gradually decreases.

We use a modified version of the fire model presented by Bishop et al. [17]. The model uses a zonal formalism for the modeling of fire. The compartment is divided into two zones: the hot gas layer and the rest of the compartment.
There is constant exchange of heat that takes place between the fire, the hot gas layer and the walls of the compartment:

$$\frac{dE}{dt} = G(T,t) - L(T,t) \tag{1}$$

where, $G(T,t)$ is the gain in energy which is chiefly determined by the growth of the fire. $L(T,t)$ is the loss of energy through the walls and the vent/exit, T is

the temperature any time t.
The existence of a feedback loop between the fire and the hot gas layer also affects the growth of the fire. The fire is responsible for the increase in the hot gas layer, which in turn radiates heat toward the base of the fire, causing an increase in the rate of growth of the fire. The fire growth rate is positive at first, ultimately leveling off.

2.2 Modified Fire Model

We use the model and notations described by Bishop et al. [17]. The equation of the rate of change of temperature is derived from the energy equation of the hot gas layer:

$$\frac{dE}{dt} = G(T,R) - L(T,R) \tag{2}$$

where $G(T,R)$ and $L(T,R)$ represent the gain and loss in energy respectively, at radius R of fire, while T is the temperature of fire. We use the equation of rate of change of fire radius, also as given by Bishop et al.[17]:

$$\frac{dR}{dt} = V_f[1 - exp(\frac{R - R_{max}}{R_{edge}})] \tag{3}$$

where, V_f is the flame spread rate in m/s, R_{max} is the maximum radius and R_{edge} is the maximum distance over which the effect of the fuel is felt. We use a non-dimensional form(in reference to units) of the equation, presented above, to make them easier to deal with.
We use Takeda's expression [7] for the rate of flame spread in our model. It is dependent upon the air supply to the compartment. We assume a constant air supply, thus making the fire-spread rate independent of temperature excess and the non-dimensional radius. Therefore,

$$V_f = s_1 m_a \tag{4}$$

where, s_1 is the spread rate, $m_a = \frac{w\xi}{\sqrt{F_r}}$, where ξ is the ventilation parameter, w is the non-dimensional width of the exit and F_r is Froude number [17].
Also, temperature is assumed to be the "fast" variable that reaches a stage of quasi-stabilization quicker whereas radius is considered a "slow" variable. This makes the spread rate independent of the temperature [2, 17].
Since maximum evacuation takes place during pre-flashover stage, we limit our simulation to it. We can vary two factors to achieve our goal, which being:

- Wall temperature parameter(u)
- Non-dimensional width of the exit.

We fix u to 0 which makes the walls perfectly conducting. This increases the loss in heat through the walls and makes heat loss a dominant factor, preventing flashover from taking place [18, 17].

3 Agent Model

We use the model proposed by Trivedi and Rao [19], which makes use of the Boids model [10] to govern the macro movements of the agents. We have also retained the model, proposed by the Trivedi and Rao [19], to calculate physical discomfort. The model uses the forces acting on the agent to determine the discomfort experienced by them. We have incorporated clustering behaviour in the existing model to account for observed real-life behaviour by pedestrians in similar situations. Additionally, we have made changes to the existing model to incorporate the fire in the model. The model makes use of changing priority orders based on the distance from the fire and the exit to calculate the forces acting on the agent.

3.1 Environment Description

The agents exist in an environment which is defined by the following set of properties [11][22]:

- Accessible: The environment is accessible, since the agents have access to all the parameter values that are required by them to decide the course of action. This entails that we have allowed the agents access to parameters, such as their distance from fire, direction to the exit door, direction of the crowd movement etc. Some additional parameters also become available to the agent depending upon its distance from additional exit doors, related agents etc.
- Non-Deterministic: The simulation model is non-deterministic. A multi-agent system is deterministic when the any action in the environment has a guaranteed effect. However, incorporation of panic model, adds uncertainty to the decision making process of the agents.
- Dynamic: Our environment incorporates the fire model, which makes the environment dynamic. The fire adds an external factor that modifies the environment, irrespective of the actions of the agents.
- Continuous: Our environment allows infinite number of actions and scenarios with each simulation, thus, making it continuous. Also,the model exists in continuous 2D space.

3.2 Additional Attributes and Behaviors

The agents themselves are defined by a 7-tuple $\langle r_i, w_i, b_i, p_i, v_i, l_i.\gamma_i \rangle$, [19][16] where the last attribute is the addition made by us to incorporate clustering.

- r_i: It is the radius associated with an agent.
- w_i: It is the weight of the agent.
- p_i: It is the position vector associated with the agent at any instance. This is used to determine the agents position with respect to the exit as well as with the hazardous entity (in our case fire)

- v_i: It is the instantaneous velocity associated with the agent at any instance.
- l_i: It is the ease distance associated with an agent.
- γ_i: It is the panic associated with any agent at an instance of time.
- b_i: It is the number of "buddies" associated with any agent. Agents that are buddies, share some relationship, which influences their movements during the evacuation. For instance, family members share a buddy relationship.

There are radial distance associated with each agent that defines their range of response i.e. within these distances, the agent reacts to the impetus. They determine the region within which the agent exhibits cohesive, alignment behaviors (d_i^c and d_i^l respectively). d_i^a is the radial distance which determines the comfort zone of the agent. Therefore, it is the distance that agent a_i tries to maintain from any agent a_j, where $j \neq i$. We added two additional distances associated with each agent:

- d_i^v: This radial distance determines the region of visibility of agent a_i. Within d_i^v, the agent is visible to its buddies, for $b_i > 0$.
- d_i^f: This radial distance determines the region, within which, the presence of fire causes a sharp increase in panic of the agent, along with a sudden increase in the velocity component in the direction opposite to the fire.

There are also "refinement factors" [19] associated with the agents. These refinement factors determine the influence, that each of the factors mentioned above, will have on the decision made by the agent at any instance of time. The five original multipliers used by Trivedi and Rao [19] are as follows:

- m_i^g: "Goal Multiplier"
- m_i^c: "Cohesion Multiplier"
- m_i^a: "Alignment Multiplier"
- m_i^s: "Separation Multiplier"—determines the factor associated with repulsive force between agents.
- m_i^o: "Obstacle Multiplier"—determines the factor associated with obstacle avoidance.

We add the following refinement factors to our model:

- m_i^f: Fire Multiplier: determines the intensity with which the agent avoids fire. This takes precedence over the goal multiplier when the agent is within close proximity to the fire i.e. $p_{fire} - p_{agent} \leq d_i^f$.
- m_i^b: Buddy Multiplier: determines the factor of pursuit of related agents. This factor is non-zero only when the buddy agents are visible to our agent, i,e, when the number of buddies within $d_i^v \neq 0$. It also takes precedence over the Goal multiplier, but is always lesser than Fire Multiplier.

We use the panic model as proposed by Trivedi and Rao [19], wherein the panic experienced by the agents is quantified. The panic is dependent upon the distance from the exit, physical discomfort etc. [9]

We include an additional component of panic to account for the panic caused due tp proximity to fire:

$$p_5 = \frac{k}{(f_i - p_i)^2} \tag{5}$$

where, f_i is the center of the fire and K is any constant.
Thus, the modified equation for calculating panic is:

$$\zeta_t = \frac{1}{5}\sum_{k=1}^{5} P_k \tag{6}$$

where,ζ_t is the sum of all factors of panic at time t.

$$\gamma_{i,t} = (\gamma_{i,(t-1)} + \zeta_t)/2 \tag{7}$$

where, $\gamma_{i,t}$ is the panic level of ith agent at time $'t'$

3.3 Clustering Model

We create a network, wherein the agents belonging to the network correspond to nodes in the network, while edges between them denote a pre-existing relationship. The edges are weighted and their weights determine the strength of the relationship. For instance, close family ties have higher weights associated with their edges. The weights vary from 0 to 1 i.e. $0 < w_e \leq 1$ These weights determine the m_i^b associated with any relationship. In our model, we have randomly created this network, while putting a modifiable upper limit to the number and size of the clusters. This model's affect is dependent upon the visibility of the related agents. That is an agent starts moving towards a related agent, only when they become visible to them. The visibility range is decided as follows:

$$d_i^v = c \cdot d_i^a \tag{8}$$

where, c is any constant.
In the presence of visible related agents, our agent calculated the velocity vector through the following formula:

$$v_i^b = c \cdot m_i^b \tag{9}$$

where, c is any constant and

$$m_i^b = \begin{cases} k, & \text{if } n > 0 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

Where k is a constant $k \in R$. Also,

$$\begin{cases} k > m_i^g, & \text{if } w_e > T \\ k < m_i^g, & \text{otherwise} \end{cases} \tag{11}$$

where, n is the number of visible related agents and T is Threshold value. Therefore, in cases where the relationship between agents is stronger than a certain threshold, reaching the related agent takes a higher priority than exiting. Instances of the above, have been found to frequently occur in real evacuation situations and are shown to slow down the evacuation process considerably.

Table 1. Properties of Fire in Simulation [19, 17]

Attribute	Symbol	Value
length of the room(metres)	l_r	26
maximum radius	m_f	$120/l_r$
effect of radius	e_f	$80/l_r$
spread rate	s_f	0.026
width of ventilation	w_f	$1/l_r$
ventilation factor	η	0.00792
Froude number	f_r	2.652

4 Integration of the Fire Model

We integrate the fire model, described above, in our simulation by adding the following attribute to fire:

- r_c: This refers to the radial distance withing which the fire results in casualties of all the agents.Therefore, all agents within this range are mortally injured.

As the agents reach with d_i^f distance of the fire, there goal changes from reaching the exit, to escaping the fire at any cost. This makes m_i^f the dominating refinement factor.
Proximity to the fire also increases the panic of the agent, thus kicking in self-preservative instincts and the agents tries to create as much distance as possible from the fire.
The velocity, after fire integration, is calculated by the following algorithm: In the algorithm, we make use of the function MOVE as defined by Trivedi and Rao [19].

5 Result

We simulate the model, in the following manner:

Algorithm 1 Goal Velocity Calculation

Calculate and assign velocity
$\boldsymbol{x_0}, \boldsymbol{v_0} \leftarrow 0$
$\boldsymbol{x_0} \leftarrow \boldsymbol{p_e} - \boldsymbol{p_i}$
$\boldsymbol{v_0} \leftarrow \mathrm{MOVE}(i)$ ▷ velocity component away from fire
if human within d_i^f **then**
    $\boldsymbol{x_1} \leftarrow (\boldsymbol{p_i} - \boldsymbol{p_f})$
end if
▷ velocity component towards buddies
for all buddies of agent i **do**
    if buddy j within d_i^v **then**
        $\boldsymbol{x_2} \leftarrow \boldsymbol{x_2} + (\boldsymbol{p_j} - \boldsymbol{p_i})$
    end if
    $\boldsymbol{x_2} \leftarrow \boldsymbol{x_2}/(N_b - 1)$
end for ▷ Velocity Calculation at time 't'
$\boldsymbol{v_t} \leftarrow \boldsymbol{v_0} + \boldsymbol{x_1} \cdot m_i^f + \boldsymbol{x_2} \cdot m_i^b$
$\boldsymbol{v_t} \leftarrow \boldsymbol{v_{t-1}} \cdot \gamma_{i,t} + \boldsymbol{v_t} \cdot (1 - \gamma_{i,t})$ ▷ Final velocity under the influence of panic

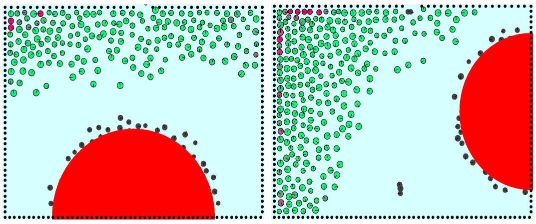
1. We first create the environment. This involves setting up the room, deciding the number of agents, position of fire etc.
2. We simulate fire, wherein we calculate the spread rate, etc.
3. We then calculate the refinement factors for each agent, depending upon its location and attributes.
4. We then calculate and set the refinement factors of an agent.
5. Panic calculation is done next, which ultimately along with the refinement factors, plays a part in calculating the final velocity.
6. We then Calculate the physical discomfort experienced by the agent.
7. Using the velocity calculations, we update the location of the agent. We also update the radius of the fire and its spread rate.

We consider different scenarios to demonstrate our results. All simulations are conducted using similar parameters in an effort to obtain better evaluations and comparisons. We do not claim that all the real scenarios are precisely described by these settings, but parameters are taken from prior sources to best replicate the real world scenarios.

The properties of agents in simulation is taken from previous studies [19, 4, 21, 5, 20]. The properties of fire are taken from Bishop et al. [17] and scaled to fit our simulations. These are described in Table 1 [8, 17].

In these simulations, we aim to qualitatively analyze evacuation plans to get better understanding of bottlenecks and critical paths formed during pedestrian evacuation.

The dimensions of the room and initial environment are the same as the compartment size in Trivedi and Rao [19]. We consider different sample cases by simulating singular vs dual exits in the room, and also by changing the location of the fire in the room. The cases are as follows:



(a) opposite (b) adjacent

Fig. 1. Case I with one exit door

1. **The position of fire with respect to the exit determines the number of casualties (Case** I**)** In this case, there is just one exit door for 400 agents. There are two sub-cases in these based on the position of fire in the compartment.
   (a) Case $I(a)$: The fire is positioned exactly on the opposite wall to exit door, as in Fig. 1(a)
   (b) Case $I(b)$: The fire is positioned in the middle of the one of the wall adjacent to wall of exit door as in Fig. 1(b)

   We compared the casualties of these cases and observed that, there was an increase of 12.048% in casualties from case $I(a)$ to $I(b)$. This is justified, since when the fire is opposite to the exit door, agents experience a repulsive force (due to the fire) in the direction of the goal (exit door). In the latter instance, the fire exerts a radial force, a component of which is in the direction opposite to the attractive force (toward the door).

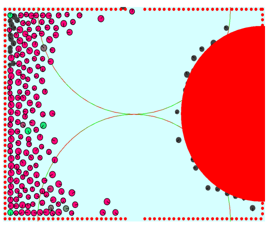
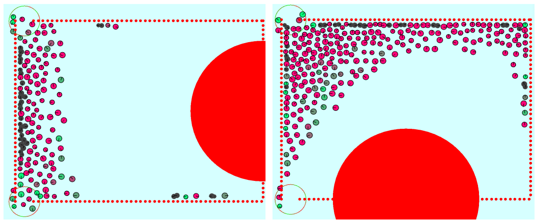


Fig. 2. Case II

2. **There is a decrease in casualties with multiple exits (Case** II**):** There are two exit doors in the middle of two opposite walls and fire is placed on the adjacent walls of exit doors as in Fig 2. Both the doors have equal visibility. There was a decrease of about 22.82% in casualties than in case $I(b)$ since there are two doors for exiting in case II even though the forces from fire act radially outwards and not toward the exit.



(a) opposite (b) adjacent

Fig. 3. Case III

3. **The relative position of the gates in multi-exit determines the rate of evacuation (Case III)** In this case, there are two doors on the left wall instead of the middle. The circle around the exits represents the visibility of the doors. The cases are as follows:
   (a) Case $III(a)$: In this, the fire is opposite to the two exits, as shown in Fig. 3(a). In this case, the evacuation was faster and with fewer casualties.
   (b) Case $III(b)$: In this case, fire position was adjacent to one of the exit doors, in Fig. 3(b). Here the evacuation was slower as one of the exit was too close to fire and was thus inefficient as an exit door.
   (c) Case $III(c)$: In this case, there are two exits on the left but there is no fire. This case was taken as the base case for comparison.

Table 2. Comparison of evacuation times in Case III

Sub-Cases	Evacuation time (seconds)
Case $III(a)$	34.86
Case $III(b)$	56.06
Case $III(c)$	48.31

The comparative analysis of these cases is shown Table. 2 (other cases can be compared similarly). We can see that, evacuation in case $III(a)$ was observed to be faster than case $III(c)$ as the repulsive force of fire acting on the agent aligns with the goal of exiting of the door and thus it is faster than case $III(c)$. In case $III(b)$, evacuation was slowest as one of the exits was too close to the fire and the agents were left with only one exit, at the corner, for evacuation.
We can also observe that placing exits at opposite corners, will be the most effective for evacuation, as at least one exit will remain fully functional in any case (i.e. any position of the fire)

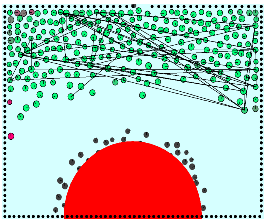


Fig. 4. Case IV

4. **Effect of Cluster (Case 4):** In this case, we have introduced clusters in case $I(a)$ as shown in Fig. 4. The lines indicate the agents forming the cluster. Maximum number of clusters formed are 10 and agents forming the cluster can be upto 6 in any simulation. This was done in an effort to best replicate real world scenarios. Even with constraints such as visibility and random cluster sizes, clustering slowed down the evacuation process. We averaged out the time taken to evacuate over 20 samples, to account for the randomness in the cluster size and number. Evacuation time without clusters was $46.8s$. With the inclusion of cluster, the time increased by 27.9% and reached $59.9s$.

6 Conclusion

Our model improves upon the existing simulations of pedestrian behavior. We were able to show how clustering can slow down the evacuation process. There also seems to be a correlation between the position of the door with respect to potential fires and the evacuation process. We also looked at how the visibility and the number of doors affect the evacuation process. The use of our model is illustrated on a small set of scenarios. It can be applied to other cases, including realistic settings, by choosing parameters suitably.

References

1. BAJAJ, V.: Fatal fire in bangladesh highlights the dangers facing garment workers (2015), https://www.nytimes.com/2012/11/26/world/asia/bangladesh-fire-kills-more-than-100-and-injures-many.html
2. Drysdale, D.: An Introduction to Fire Dynamics. Wiley (1985)
3. Grosshandler, W.L., Bryner, N.P., Madrzykowski, D.M., Kuntz, K.: Report of the Technical Investigation of The Station Nightclub Fire (NIST NCSTAR 2), Volume 1. Tech. rep., National Institute of Standards and Technology (NIST) (Jun 2005)
4. Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. Nature **407**(6803), 487–490 (2000)
5. Helbing, D., Farkas, I.J., Molnar, P., Vicsek, T.: Simulation of pedestrian crowds in normal and evacuation situations. Pedestrian and Evacuation Dynamics **21**, 21–58 (2002)

6. Helbing, D., Johansson, A., Al-Abideen, H.Z.: Dynamics of crowd disasters: An empirical study. Physical Review E **75**(4), 046109 (2007)
7. Hisahiro Takeda, K.A.: New modeling of liquid or thermoplastic pool fires in compartment. Symposium (International) on Combustion **19**(1), 897–904 (1982)
8. Hgglund, B.: A room fire simulation model. Fire and Materials **8**(2), 105–111 (1984). https://doi.org/10.1002/fam.810080208, https://onlinelibrary.wiley.com/doi/abs/10.1002/fam.810080208
9. Lu, X., Luh, P., Tucker, A., Gifford, T., S. Astur, R., Olderman, N.: Impacts of anxiety in building fire and smoke evacuation: Modeling and validation. IEEE Robot. Autom. Lett. **2**(1), 255–260 (Jan 2017). https://doi.org/10.1109/LRA.2016.2579744
10. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: MASON: A multiagent simulation environment. Simulation **81**(7), 517–527 (2005)
11. Macal, C.M., North, M.J.: Tutorial on agent-based modelling and simulation. Journal of Simulation **4**(3), 151–162 (2010). https://doi.org/10.1057/jos.2010.3, https://doi.org/10.1057/jos.2010.3
12. Moore-Bick, M.: Grenfell Tower Inquiry. https://www.grenfelltowerinquiry.org.uk/ (2018)
13. O'Hagan, A.: The Tower. London Review of Books **40**(11) (2018)
14. Pires, T.T.: An approach for modeling human cognitive behavior in evacuation models. Fire Safety Journal **40**(2), 177–189 (2005)
15. Reynolds, C.W.: Flocks, herds and schools: A distributed behavioral model. In: ACM SIGGRAPH Computer Graphics. vol. 21, pp. 25–34. ACM (1987)
16. Rockenbach, G., Boeira, C., Schaffer, D., Antonitsch, A., Musse, S.: Simulating crowd evacuation: From comfort to panic situations. pp. 295–300 (11 2018). https://doi.org/10.1145/3267851.3267872
17. S.R.Bishop, P.G.Holborn, A., D.D.Drysdale: Dynamic modelling of building fires. Applied Mathematical Modelling **17**(4), 170–183 (1993). https://doi.org/10.1016/0307-904X(93)90105-P
18. Thomas, P., Bullen, M., Quintiere, J., McCaffrey, B.: Flashover and instabilities in fire behavior. Combustion and Flame **38**, 159 – 171 (1980). https://doi.org/https://doi.org/10.1016/0010-2180(80)90048-6, http://www.sciencedirect.com/science/article/pii/0010218080900486
19. Trivedi, A., Rao, S.: Agent-based modeling of emergency evacuations considering human panic behavior **5**(1), 277–288 (Mar 2018). https://doi.org/10.1109/TCSS.2017.2783332
20. Viswanathan, V., Lee, C.E., Lees, M.H., Cheong, S.A., Sloot, P.M.: Quantitative comparison between crowd models for evacuation planning and evaluation. The European Physical Journal B **87**(2), 1–11 (2014)
21. Viswanathan, V., Lees, M.: An information processing based model of preevacuation behavior for agent based egress simulation. In: Pedestrian and Evacuation Dynamics 2012, pp. 125–133. Springer (2014)
22. Wooldridge, M.: An Inroduction to Multiagent Systems. John Wiley Sons, Ltd (2002)

An opinion diffusion model with deliberation

George Butler, Gabriella Pigozzi, and Juliette Rouchier

Université Paris-Dauphine, PSL Research University, CNRS, UMR 7245, Paris 75016, France
{name,surname}@lamsade.dauphine.fr

Abstract. In this article, we propose an agent-based model of opinion diffusion and voting where agents influence each other through deliberation. The model is inspired from social modeling as it describes a process of collective decision-making that iterates on a series of dyadic inter-individual influence steps and collective argumentative deliberation procedures. We study the evolution of opinions and the correctness of decisions taken within a group. We also aim at founding a comprehensive model to describe collective decision-making as a combination of two different paradigms: argumentation theory and agent-based influence models, which are not obvious to link since a formal translation and interpretation of their relationship is required. We find that deliberation, modeled as an exchange of arguments, reduces the variance of opinions and the number of extremists as long as not too much deliberation takes place during the decision-making process. Insofar as we define "correct" decisions as those whose supporting arguments survive deliberation, promoting deliberative discussion favors convergence towards correct decisions.

Keywords: Opinion diffusion · abstract argumentation · agent-based modeling · deliberation

1 Introduction

In a group, opinions are formed over affinities and conflicts among the individuals that compose it. Axelrod [3], a pioneer in opinion dynamics, shed light on two key factors required to model the processes of opinion diffusion, namely, social influence (i.e., individuals become more similar when they interact) and homophily (i.e., individuals interact preferentially with similar others). He showed that interactions through those factors lead to emergent collective opinions of which the individual had poor control. Since, a growing body of research has endeavored to identify the conditions under which social influence, at the micro (dyadic) level, translates into macro patterns of diffusion through repeated iterations [26]. Two types of models appear in the literature: on the one hand, the Ising-type models where opinions take discrete values [3, 15]; on the other, the continuous opinion models where opinions are represented by real numbers [10, 18, 19, 25, 30].

The question of group deliberation, defined as an exchange of arguments, is not explicitly taken into account in opinion diffusion. Opinion dynamics seem to miss the intuition that individual behavior may be determined by factors related to non-dyadic channels of interaction, such as deliberation arenas, and to the structure and size of the channels of communication themselves. When a group engages in a discussion, group size, what arguments are advanced, how discussion is organized over time, and the acceptability criteria for proposals may lead to a transformation of preferences [17] and play a crucial role in consensus formation [13,20,28]. Moscovici and Doise [20] explain that there are two types of "discussions" in deliberation, informal or *warm* and formal or *cold*, that potentially lead to consensus. They show that when a group is asked to reach an agreement through informal or non-procedural deliberation, the obtained consensus is more likely to be extreme compared to the average of the pre-consensus individual opinions. When deliberation is procedural, the obtained consensus tends to be milder and opinions less polarized. Opinion diffusion has also been used to track convergence towards "correct opinions". For example, authors in [16, 22] study network effects and signaling in diffusion; in [25], the authors explore how dyadic interactions diffuse true information on an exogenous true state of the world. In a deliberative context, a correct decision corresponds to one derived from a state of the world in which all arguments for and against the decision are taken into account [8]. Deliberation reveals such arguments. Hence, it may help a group converge towards correct decisions. For this reason, decision-making processes with deliberation should be explored.

The aim of our model is to breach the gap between deliberation and opinion diffusion. Drawing from [20], we model warm discussion using an opinion diffusion model based on social judgment theory [18, 27], and cold discussion using abstract argumentation theory [7, 12]. We engineer a decision-making process with voting that terminates according to deliberated decisions, as we draw inspiration from the literature in deliberative democracy [8, 13, 28]. We describe the effects of deliberation on opinions and on the *correctness*, a group's ability to correctly judge propositions, and *coherence*, a group's ability to accept deliberated proposals, of collective decisions by modeling decision-making processes as a sequence of deliberative and dyadic interactions among agents. In particular, we study how the frequency, size (number of agents), and voting rules of deliberative interactions impact opinions and the correctness and coherence of collective decisions.

Our model shows that deliberation has a significant overall impact on the distribution of opinions (variance) and on the overall shifts of opinion. We provide evidence of Moscovici and Doise's [20] results on consensus: when specifying opinion dynamics as only deliberative, the proportion of extremists and the variance of opinions are lower than in a non-deliberative specification of the dynamics. However, as observed in [28], if deliberation is mandatory in decision-making processes, more deliberation translates into an increase in the variance of opinions and of the proportion of extremists. The model also explains that the frequency of deliberative interactions as well as the number of agents that participate in deliberation increase judgment accuracy in a marginally decreasing fashion, but have no significant effect on the coherence of collective decisions. Last, we point out that results are strongly conditioned to the voting majority quota rule and to how agents advance arguments during deliberation.

The remainder of this paper goes as follows: in section 2, we present the model, provide the necessary basics to understand its implementation, and we introduce the metrics of interest. In section 3, we report and discuss our results; sections 4 and 5 are dedicated to related works and to the conclusion of the article.

2 A model for collective decision-making with deliberation

Let N be a group composed of $|N| = n$ agents. The group faces the question of whether to accept or reject a proposal $\mathcal{P}$ justified by an argument I. I, or *proposal argument*, is judged on how well it supports a principle $\mathbb{P}$ or its opposite $\neg\,\mathbb{P}$. Agents discuss the proposal on the basis of their adherence to the principle $\mathbb{P}$. When agents discuss informally, they are subject to random pair-wise influence; when they argue formally, they are impelled by the results obtained in the decision-making process. A *decision-making process* $\mathcal{D}(\mathcal{P}, I)$ on a proposal $\mathcal{P}$ is a sequence of formal and informal discussions that leads to a decision on the acceptance of $\mathcal{P}$ (see Fig. 1). A proposal $\mathcal{P}$ is accepted if the argument I that justifies it is accepted in deliberation and/or voted favorably by a majority of agents.

$$\mathcal{D}(\mathcal{P}, I) : \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & d & \overline{d} & \overline{d} & \overline{d} & V \\ \hline \end{array}$$

Fig. 1: A decision-making process $\mathcal{D}(\mathcal{P}, I)$. $\overline{d}$ stand for informal discussion steps, d for deliberative interaction, and V to a vote over I. Agents update their opinions according to the obtained result.

2.1 Deliberative agents and opinion dynamics with deliberation

Every agent i has an opinion, a relative position or degree of adherence $o_i \in [-1, 1]$ to the principle $\mathbb{P}$ and a couple $(T_i, U_i) \in [0, 2] \times [0, 2]$ $(U_i \leq T_i)$ of latitudes of rejection and acceptance, respectively, of informational cues. The idea is that there exist levels of relative tolerance from which informational cues have either an attractive or a repulsive effect on the individual [27]. An o_i close to 1 implies that agent i fully supports the principle $\mathbb{P}$, close to -1 that she rejects principle $\mathbb{P}$ or, equivalently, fully supports $\neg\,\mathbb{P}$.

Let $\mathcal{A}$ be a finite set of arguments, seen through the principle $\mathbb{P}$, that agents may hold in a debate over a proposal $\mathcal{P}$. Each agent i has a sack of arguments $\mathcal{A}_i \subset \mathcal{A}$ whose content reflects her relative position, o_i, on $\mathbb{P}$. Thus, agents possess partial knowledge on the relationships between the arguments in $\mathcal{A}$. If $a \in \mathcal{A}_i$, then agent i knows which arguments are in conflict with a. Each argument $a \in \mathcal{A}$ is given a real number $v_a \in [-1, 1]$ that stands for how much a respects or supports the principle $\mathbb{P}$. $v_a = 1$ means that argument a is totally coherent with the principle $\mathbb{P}$, whereas $v_a = -1$ reads "argument a is totally incoherent with the principle $\mathbb{P}$". Agents have an incentive to deliberate because they know that deliberation is an opportunity to either support or undermine a proposal that opposes their position on $\mathbb{P}$. They may present two types of behavior, *naive* and *focused*. Naive agents will only use deliberation to voice their opinions on the principle. Focused agents strategically argue in favor of proposal arguments that support the principle they favor, thus using all the information they have on the relationship between arguments. All agents (1) are able to assess the degree of support for $\mathbb{P}$ of all arguments, (2) agree on the existence of a conflict between any two arguments if such is announced during deliberation, and (3) are sincere when communicating their positions to each other. At time t, an agent i votes favorably for a proposal $\mathcal{P}$ of justification argument I if and only if $v_I(t) \times o_i(t) \geq 0$.

A dynamics for informal discussion At each informal discussion time step t, every agent i randomly meets one other agent j and updates her opinion according to the following dynamic equation:

$$o_i(t+1) = \begin{cases} o_i(t) + \mu(o_j(t) - o_i(t)) & if \quad |o_i(t) - o_j(t)| < U_i \\ o_i(t) + \mu(o_i(t) - o_j(t)) & if \quad |o_i(t) - o_j(t)| > T_i \\ o_i(t) & if \quad U_i \leq |o_i(t) - o_j(t)| \leq T_i \end{cases} \tag{1}$$

where the parameter $\mu \in [0, \frac{1}{2}]$ controls for the strength of attraction and repulsion in social influence and (T_i, U_i) for i's couple of latitudes of rejection and acceptance, respectively, for informational cues. The meeting and updating of opinions in this situation are loosely associated to Moscovici and Doise's warm discussion [20] and will be denominated the *warm discussion* model.

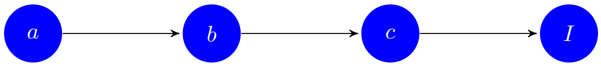


Fig. 2: Argumentation framework $AF = (\mathcal{A}, \mathcal{R})$ with $\mathcal{A} = \{a, b, c\}$ and $\mathcal{R} = \{(a, b), (b, c), (c, I)\}$. The labeling $\{\{c\}, \{I\}, \{a, b\}\}$ is conflict-free, a and b are undecided, c is accepted and I is rejected. $\{\{a, c\}, \{b, I\}, \emptyset\}$ is the only complete labeling obtained from the framework.

2.2 Abstract argumentation and deliberative models for collective decision-making

Deliberation, defined as an exchange of arguments, may be modeled by confronting, eventually contending, arguments. Following Dung's abstract argumentation theory [12], let $\mathcal{A}$ be a finite set of arguments and $\mathcal{R}$ a subset of $\mathcal{A} \times \mathcal{A}$ called *attack relation.* $(a, b) \in \mathcal{R}$ stands for "argument a *attacks* argument b", meaning that argument a is in conflict with argument b. One says that an argument c *defends* an argument a if there exists b such that $(c, b) \in \mathcal{R}$ and $(b, a) \in \mathcal{R}$. One names *argumentation framework* (AF) the couple $(\mathcal{A}, \mathcal{R})$ composed of a set of arguments and their attack relation, which can be seen as a digraph in which the nodes are the arguments and the arcs are the attacks. A *label* $\mathcal{L}ab(a) \in \{\textbf{IN}, \textbf{OUT}, \textbf{UND}\}$ of an argument $a \in \mathcal{A}$ denotes the acceptability status of a in a deliberation process. Intuitively, an argument is labeled **IN** if it is acceptable, **OUT** if it is not and **UND**, if nor **IN** nor **OUT** labels are applicable. Moreover, one defines a *labeling* on an argumentation framework $AF = (\mathcal{A}, \mathcal{R})$ as a complete function $\mathcal{L} : \mathcal{A} \to \{\textbf{IN}, \textbf{OUT}, \textbf{UND}\}$, $a \mapsto \mathcal{L}ab(a)$ that assigns a label to each argument in AF. A labeling-based *semantics* is a set of criteria that yields acceptable labelings. For example, if an argument a attacks an argument b, then an acceptable labeling should not assign the label **IN** to both arguments. Basic semantics demand labelings to be *conflict-free*, meaning that no two arguments that attack each other are labeled **IN**, or *admissible*, implying that the labeling is conflict-free and that for any **IN** labeled argument a, there exists another **IN** labeled argument c such that c defends (or reinstates) a.

The family of admissibility-based labelings goes from *complete* labellings, which are admissible labelings for which all labels (including the undecided) are justified [5], to *preferred* and *grounded* labellings which are complete labellings obtained from, respectively, maximizing and minimizing the number of arguments that are labeled **IN**. They capture properties such as credulity and skepticism in argumentation. For a more extensive account of semantics and labellings, refer to [5].

We incur to abstract argumentation theory because it provides a comprehensive formalism that bypasses difficulties related to the nature and construction of arguments. The formalism also lends itself well to graph theory and to model (collective) reasoning in a clear, coherent and easy way [29]. Given an argumentation framework, Dung's extension-based [12] approach is only interested in the set of acceptable arguments (according to a certain semantics). The labelling approach [7] assigns a label to each argument in the framework. Hence, the approach is more expressive since it distinguishes arguments that are not accepted from those that are undecided. Such distinction is crucial since the existence of undecided arguments is one of the reasons why deliberation takes place and carries on over time. This is the primary justification for using labeling-based semantics in our model. Figure 2 provides an example of an argumentation framework that models one "step" of deliberation over a proposal justified by an argument I.

Deliberative collective decision-making protocol. Debates take place on a table in which a *central authority* (CA) [6] fixes the deliberation procedure. The CA chooses the percentage of agents (n_D) from the population that actively participate in the deliberation, the labeling-based semantics (σ) used to assess the label of the proposal argument, and the number ($\underline{m}$) of debates that ought to take place before a decision is deemed sufficiently discussed. Additionally, it also controls the maximum number ($\overline{m}$) of debates that can take place before abandoning deliberation, and the number (t_D) of informal discussion steps between debates. The CA also decides which collective decision rules to apply during the process (e.g. whether there is voting on proposals) and the proportion (α) of favorable votes in the population necessary to accept a proposal. Given a proposal $\mathcal{P}$, the deliberation or debate protocol goes as follows:

1. The CA generates and makes public a central argument or proposal argument $I \notin \mathcal{A}$;
2. The CA randomly draws two sets of $\frac{n_D}{2} \times n$ agents with divergent views on $\mathcal{P}$, namely $\mathbb{P}$;
3. Each agent advances an argument from her sack $\mathcal{A}_i$. The CA makes sure that there are no repeated arguments with respect to previous debates on the same proposal (tables have memory);
4. The CA builds the debate's argumentation framework on the previously held debates over the proposal. It computes a labeling for the arguments using the semantics σ;
5. If the obtained label for I is undecided ($\mathcal{L}ab_d(I) = \mathbf{UND}$) or the number of debates steps held in the decision process is inferior or equal to $\underline{m}$ at time t, then the CA stops the debate and resumes it at the $(t + t_D + 1)$'th time step by repeating 1, 2, 3, 4 and 5;
6. Let $\mathcal{L}ab(I)$ be the final label given to the proposal argument I. If voting is not part of the process ($\alpha = 0$), $\mathcal{L}ab_d(I) = \mathcal{L}ab(I)$; otherwise if more than $\alpha \times n$ agents agree with I, I is accepted ($\mathcal{L}ab(I) = \mathbf{IN}$), refused if strictly less than $\alpha \times n$ agents agree with it ($\mathcal{L}ab(I) = \mathbf{OUT}$). If there is a tie $\mathcal{L}ab_d(I) = \mathbf{UND} \Rightarrow \mathcal{L}ab(I) = \mathbf{OUT}$ and $\mathcal{L}ab_d(I) = \mathbf{IN} \Rightarrow \mathcal{L}ab(I) = \mathbf{IN}$.

Notice that deliberation always ends: either agents debate and agree on the proposal's acceptability through procedural argumentation or, after $\overline{m}$ debate steps, they directly vote on it. Also, observe that voting for a proposal is the same as voting for the argument that justifies it.

2.3 Linking deliberation and informal discussion through opinions

Let $\mathcal{P}$ be a proposal, $v_I(t)$ the proposal argument I's level of support for a principle $\mathbb{P}$ and $o_i(t)$ an agent i's opinion at time t. Then, given the distance $\delta_i(t) = \frac{1}{2}|v_I(t) - o_i(t)|$ and the acceptability status $\mathcal{L}ab(I)$ of I at the end of a decision process over $\mathcal{P}$, agent i updates her opinion as follows:

$$o_i(t+1) = \begin{cases} o_i(t) + \gamma(v_I(t) - o_i(t)) & if\ \ \mathcal{L}ab(I) = \mathbf{IN},\ with\ probability\ p_a^{\delta_i(t)} \\ o_i(t) + \gamma(o_i(t) - v_I(t)) & if\ \ \mathcal{L}ab(I) = \mathbf{IN},\ with\ probability\ p_r^{\frac{1}{\delta_i(t)}} \\ o_i(t) & if\ \ \mathcal{L}ab(I) \neq \mathbf{IN},\ with\ probability\ 1 - p_r^{\frac{1}{\delta_i(t)}} - p_a^{\delta_i(t)} \end{cases} \quad (2)$$

where $\gamma \in [0, \frac{1}{2}]$ is the strength of repulsion and attraction in the dynamics. p_a and p_r are probability parameters that control for the possibility that an agent is attracted to and repulsed from agreements reached during debates. The equation combines the probabilistic nature of the effect of deliberation based on a principle similar to the one in social judgment theory [27], be it a moderating [13, 20] or polarizing [28] one. It follows that deliberated informational cues may potentially influence any agent in the group. We call the model in which agents only update their opinions by Equation 2 the *cold discussion* model, as we associate it to Moscovici and Doise's [20] cold discussion. We call the *mixed discussion* model the model defined by Equations 1, 2 and the decision-making protocol.

2.4 Simulations

A time step in the model corresponds to either a debate, a step of dyadic social influence or a vote that makes agents update their opinions[1]. Simulations stop once 100 decision-making processes over 100 randomly generated proposals terminate. We observe how deliberation affects opinion distributions, coherence between majority voting and deliberative results, and judgment accuracy taking as reference the warm and cold discussion models.

Observations. At the end of each simulation ($t = S$), we observe the following metrics:

- **Variance of opinions ($Var(o)$)**: the variance of opinions at time S. The higher the variance of the distribution, the more "diverse" opinions are in the opinion pool;
- **Shift in opinions (Sh)** [9]: statistic that measures the aggregated change in individual opinion at time S with respect to time 0, $Sh = \frac{2\sum_{i \in N} |o_i(0) - o_i(S)|}{max_{i \in N} o_i(0) - min_{i \in N} o_i(0)}$;
- **Proportion of extremists in the population ($prop_{ex}$)**: percentage (%) or proportion of agents in the population with non-moderate opinions (i.e. $|o_i(S)| \geq 0.75$);
- **Judgment or consensual inaccuracy (ec)**: it consists of a statistic measuring a group's ability to infer correct labels for proposal arguments. It measures the correctness of the decisions taken by the group. Correct labels are obtained from the argumentation framework AF_ε that contains all arguments and their attacks. Let $\mathcal{I}$ be the set of all discussed proposal arguments up to S and $\mathcal{L}ab_\varepsilon(I)$ the label given to I in AF_ε. We use a Hamming-based distance on labellings as introduced in [2] to explicitly define the statistic: $ec = \frac{1}{|\mathcal{I}|}\sum_{I \in \mathcal{I}} a_I |\mathcal{L}ab_\varepsilon(I) \neq \mathcal{L}ab(I)|$, where $a_I = \frac{1}{2}$ if $\mathcal{L}ab_\varepsilon(I) = \mathbf{UND}$ or $\mathcal{L}ab(I) = \mathbf{UND}$ and $a_I = 1$, otherwise;
- **Coherence (ir)**: let $\mathcal{L}ab_d(I)$ be the label obtained for I from the deliberation process without voting. The coherence statistic measures how well voting results adjust to results obtained during deliberation: $ir = \frac{|\{I \in \mathcal{I} \mid \mathcal{L}ab_d(I) = \mathbf{IN}, \mathcal{L}ab(I) = \mathbf{IN}\}|}{|\{I \in \mathcal{I} \mid \mathcal{L}ab_d(I) = \mathbf{IN}\}|}$.

Initialization. All agents start off with an opinion o_i drawn from a uniform distribution $\mathcal{U}(-1, 1)$. For all agent i, we set $(U_i, T_i) = (U, T)$ for some $(U, T) \in]0, 2[\times]0, 2[$, μ to 0.1 and p_r to 0.05. Given o_i, agents randomly draw a set $\mathcal{A}_i$ ($|\mathcal{A}_i| = k$) of arguments from a balanced[2] argument pool $\mathcal{A}$ of $m = 600$ non-neutral arguments on the basis of o_i. Each argument $a \in \mathcal{A}$ is given a level of support for the principle $\mathbb{P}$, v_a, obtained from a uniform distribution $\mathcal{U}(-1, 1)$. The attack relation $\mathcal{R}$ that gives birth to the consensual argumentation framework AF_ε is established according to the v_as and is given a permanent labeling $\mathcal{L}_\varepsilon^\sigma$ computed using σ = grounded semantics. On the proposal side, we create an argument $I \notin \mathcal{A}$ whose support for $\mathbb{P}$ is also drawn from a uniform distribution $\mathcal{U}(-1, 1)$, and is given the label $\mathcal{L}ab(I) = \mathbf{UND}$. We allow I to attack no argument, yet allow other arguments to randomly attack it. Finally, we set the maximum number of debates to $\overline{m} = 7$ and following [18], we set the number of agents in the model to 400.

3 Simulation results

We obtain two kinds of results. The first is global and answers the question on how deliberation affects opinion formation. It consists of the comparison between the warm (Eq. 1), cold (Eq. 2),

[1] In warm discussion, agents vote for the proposal arguments, but do not update their opinions.
[2] By balanced we mean with as many arguments with $v_a < 0$ as with $v_b > 0$.

Length of $\mathcal{D}_I$		**Decisional in D_I**		**Social influence**		**Deliberation influence**		**Cognitive**	
t_D	$\underline{m}$	n_D	α	T	U	p_a	γ	k	$focused$
{1,3,6}	{1,3,6}	{0.01,0.02,0.05}	$\{0, \frac{1}{2}, \frac{2}{3}\}$	{1.4,1.8}	{0.2,0.6}	{0.1,0.3,0.5}	{0.05,0.1,0.2}	{4,8,16}	$\{True, False\}$

Table 1: Multimodal parameter domains used to compare warm, mixed, and cold discussion.

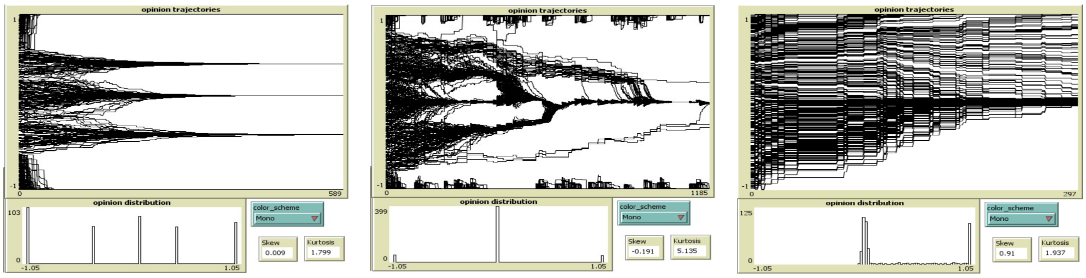


Fig. 3: From left to right, opinion trajectories and distributions for warm, mixed, and cold discussion.

Model \ Metric	$Var(o)$	$prop_{ex}$	Sh	ec	ir
Warm vs. Cold discussion	[0.191,0.216]	[0.170,0.197]	[-29.48,-28.28]	[0.074,0.082]	[0.165,0.172]
Mixed vs. Cold discussion	[0.142,0.147]	[0.150,0.155]	[-20.85,-19.76]	[0.003, 0.002]	[-0.038,-0.030]
Mixed vs. Warm discussion	[-0.072,-0.046]	[-0.045,-0.018]	[8.332, 8.859]	[-0.079,-0.071]	[-0.205,-0.201]

Table 2: Mean difference 0.95 confidence intervals for metrics by model comparison.

and mixed (Eq. 1, Eq. 2 w. deliberation protocol) discussion models. We simulate from 10 to 30 runs for each model and scenarii on the parameter space induced by the initialization and Table 1.

The second kind of results consists of a sensitivity analysis. It addresses the questions regarding the importance of procedural deliberation parameters, namely n_D, t_D, α, and $\underline{m}$, and agent behavior on our metrics in the mixed discussion model. We span their domain as described in Table 3, and generate 36,000 observations. Simulations and analyses are performed in Netlogo 6.0.4. and R 3.2.3.

Comparing the different models From the simulations, we observe that the variance of opinions and the proportion of extremists are strongly correlated ($\rho \approx 0.95$, $p < 0.001$). We infer that cold discussion favors judgment accuracy, reduces the variance of opinions and the proportion of extremists. Although there is opinion polarization, only one group of extremists forms, probably the one in favor of the first deliberated results (see Fig. 3). Otherwise, a moderate consensus around neutrality forms. Warm discussion, on the other hand, is responsible for an increase in the variance of opinions and in the number of extremists. Coherence is maximal since agents only vote for proposals. Interestingly, we see that the mixed model is a compromise of the warm and cold discussion models (see Table 2). Deliberation not only contributes to obtaining correct answers but also to a slight decrease in the variance of opinion and in the proportion of extremists. However, it does not do better than the cold discussion model on coherence and produces less shifts of opinion.

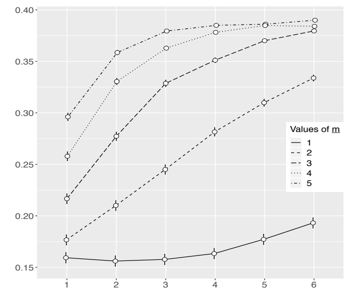
Procedural parameters of interest					Other parameters						
t_D	$\underline{m}$	n_D	α	$focused$	$\overline{m}$	T	U	p_a	p_r	γ	k
{1,2,...,6}	{1,2,...,5}	{0.01,0.02,...,0.05}	$\{0,\frac{1}{2},\frac{2}{3}\}$	$\{True, False\}$	{7}	{1.6}	{0.2}	{0.3}	{0.05}	{0.2}	{12}

Table 3: Domains and types for procedural and behavior parameters in sensitivity analysis.

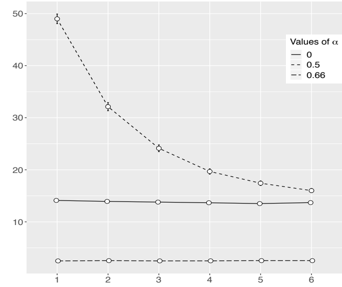
A first reading of the result says that there is a trade-off between judgment accuracy and variance in opinion. More accuracy is related to slightly less extremism, which points to the fact that extremism may not contribute to successful deliberation.

Minimum number of debates ($\underline{m}$). We observe that minimum number of debates has a significant, well-observed effect on all of our metrics excluding *coherence* (ir). Taking variance of opinions, we notice that the more debates there are, the bigger the value of the metric is, and the higher n_D and t_D are, the weaker is the overall effect (Fig. 4d and Fig. 4a). Moreover, the marginal increase of the minimal number of debates on the variance of opinions is decreasing. In contrast, shifts in opinion are less and less likely as $\underline{m}$ grows and this independently of other parameters. Again, the effect is marginally decreasing and is only truly significant when $\alpha = \frac{1}{2}$. An explanation of these effects may be linked to the design of the system. First, the variance of opinions is higher when deliberation is asked for because the more deliberation steps there are the higher the chances are that the proposal argument is deemed unacceptable. Mechanically speaking, increasing the minimal amount of debates implies that, whenever a decision is to be taken, at least $\underline{m} \times t_D$ time steps have to take place, and, if an argument is considered undecided, t_D time steps are added to the process. So, unless the debate yields decisive labels for proposal arguments (less likely considering that σ = grounded semantics), more non-deliberation steps take place in the decision process and the higher the variance of opinions is. Concerning shifts, when $\alpha \neq \frac{1}{2}$, either the system is too stiff to accept any proposal argument, and opinions do not change much, or the effects of pair-wise discussion and deliberation cancel out (Fig. 4b). On the side of labeling-based metrics, the more debates are asked for, the more accurate a group is in its judgment—the effect being smaller as $\underline{m}$ grows. When agents are naive, the effect is more linear; when they are focused, the strongest effects of adding more deliberation are found when levels of deliberation are already low (Fig. 4f). This can be explained by the fact that the more debates there are in the decision process, the closer one gets to the consensual argumentation framework. The effect is stronger for the focused agents because, when reconstructing the framework, they take into account the deliberated proposal and advance the most pertinent arguments they posses relative to the attack relation $\mathcal{R}$.

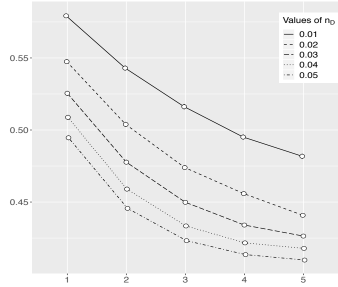
Proportion of the population in deliberation steps (n_D). Like with $\underline{m}$, n_D has a significant effect on the proportion of extremists and on the variance and shift of opinions (Fig. 4e). This may result from the fact that being able to put more arguments in play at the same debate step can increase the chances of revealing the cycles around the proposal argument. Given that we use grounded semantics, the arguments in the cycles are labeled **UND** thus postponing debates more often than if n_D was lower. Postponing debates, in turn, increases the number of informal interactions in the decision process, which increases the variance in opinion and limits the effect of deliberation. Moreover, the effect of this parameter is very dependent on the value of α (Fig. 4e). For the shift metric, for instance, $\alpha = \frac{1}{2}$ makes the effect of n_D negative, while $\alpha = 0$ makes it positive



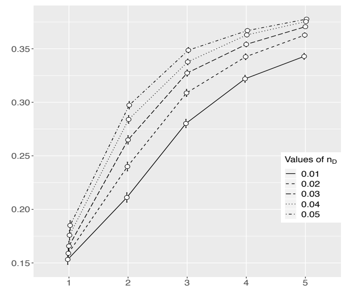
(a) ($t_D \times prop_{ex}$) by $\underline{m}$.



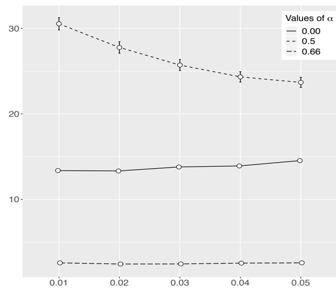
(b) ($t_D \times Sh$) by α.



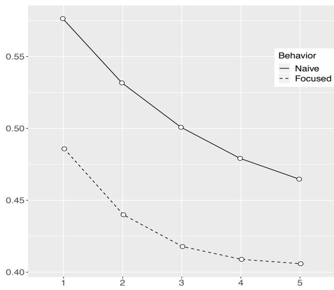
(c) ($\underline{m} \times ec$) by n_D.



(d) ($\underline{m} \times prop_{ex}$) by n_D.



(e) ($n_D \times Sh$) by α.



(f) ($\underline{m} \times ec$) by $focused$.

Fig. 4: Curves of mean observations for 36,000 runs on metrics (0.95 confidence intervals).

to a lesser absolute degree. For higher requirements for deliberation ($\underline{m}$), adding more individuals to the deliberation process has weaker effects on the variance of opinions and on the other metrics. It is also interesting to notice that it has no significantly different effect on the metrics on whether agents are naive or focused. This is a surprising result as one would have expected more focused agents in an arena to heavily impact the proportion of extremists. They play to knock out opposing proposal arguments and thereof hinder the opinion-moderating effects of deliberation. Similar to $\underline{m}$, adding more people into the deliberation process increases judgment accuracy (Fig. 4c).

Steps between deliberation steps (t_D). In all configurations, t_D increases the proportion of extremists (variance) and decrease the shifts in opinion. The shifts and the effect on the variance of opinions are only observable for $\alpha = \frac{1}{2}$ (Fig. 4b). t_D is highly linked to $\underline{m}$ by construction. When $\underline{m} = 1$, the curve linking the variance of opinions and t_D is convex. As $\underline{m}$ increases, the curve becomes more and more concave, which means that t_D has a more important effect on the opinion distribution as collective decision-making processes are longer. This seems counter-intuitive yet it reflects the multiplicative relationship between deliberation and pair-wise interactions. If $\underline{m}$ is low, and t_D high, the effective number of pair-wise interactions are, on average, fewer in the deliberation process, which bounds the increase in opinion variability. Additionally, since the grounded semantics yields few **IN** arguments w.r.t. other admissibility-based semantics, getting closer to the consensual argumentation framework may lessen the number of opinion updates due to deliberation. Last, a lower $\underline{m}$ makes deliberation more influential on opinions.

Acceptability voting quota (α). By far, the most influential parameter in our study. It changes the direction and the intensity of the effect of other procedural parameters and, by construction, heavily constrains the road to accepting a proposal. In few words, α determines which model of discussion, the warm or the cold, dominates the dynamics. One either gives too much weight to deliberated results ($\alpha = 0$) and the effect of pair-wise interactions becomes negligible, or too much weight to pair-wise interaction ($\alpha = \frac{2}{3}$). It follows that updates due to deliberation happen rarely and opinions do not moderate. Concerning labeling-based metrics, α entirely determines the coherence statistic. For $\alpha \neq \frac{1}{2}$, coherence is maximal when $\alpha = 0$ and when $\alpha = \frac{2}{3}$. For the former, maximality is trivial: agents accept any deliberated argument; for the latter, since the latitudes of acceptance (rejection) are too low (high), pair-wise interactions alone do not unevenly polarize the population in such way that deliberated arguments are accepted by a $\frac{2}{3}$-majority.

4 Related Work

We see our model as a contribution to the influence and opinion dynamics field in agent-based modeling (ABM) and a pragmatic application of abstract argumentation theory. To our knowledge, we are unaware of existing literature on ABM that explicitly relates collective decision-making, deliberation by abstract argumentation, and opinion diffusion as we have done it. This said, many models in the literature of opinion diffusion are interested in opinions because they influence collective decisions and can be used to reveal certain types of social phenomena. For instance, in [15] the authors are interested in consensus and in how a group collectively decides on an action when it is given two alternatives. In other models, authors are interested in the emergence of extremism [19] and on the distribution of opinions when extremists are introduced in the population [10], while other authors coin the notion of opinion polarization as an emergent property [19]. They show, using models of "bounded confidence" and opinion diffusion with trust, that three different kinds of steady states (unipolar, bipolar and central) were possible depending on whether agents were sufficiently uncertain about their opinions, sufficiently connected, and/or a certain proportion of individuals were already extreme. Similarly, work on collective cognitive convergence and opinion sharing [22] show that consensus towards a certain opinion or cognitive state is always possible yet dependent on noise, variability and awareness of agents. Closer to opinion formation and argumentation, authors in [14] define an agent's opinion as a function of the arguments she holds and their relationship (logical). They device a peer-to-peer dialog system (NetArg) that uses only abstract argumentation to study opinion polarization and opinion dynamics. When it comes to abstract argumentation theory, we take an approach that wires two type of dialogues that are well-studied in the literature: persuasion dialogues [21] and deliberation dialogues [1]. The line of work that might be closest to ours is the one on mechanism design [23], or the problem of devising an argumentation protocol where strategic argumentation has no benefit. We tackle mechanism design in a different way. Instead of considering strategy-proofness, we are interested in how differences in protocol can result in "better" collective choices and guarantee that opinion distributions are favorable for deliberation ("reasonable" level of variance). For a survey on persuasion dialogue, see [21].

Work on agent-based argumentation usually assumes that the semantic relationship between arguments is fixed [23, 24]. Other models which do not make this restrictive assumption can also be found in the literature and derive from the class or family of opponent models [6] in which two opposing sides attempt to win the dialogue. Our model is in the intersection of these, but the framework that combines opinion diffusion of the kind and argumentation seems original.

The idea of mixing interpersonal influence and vertical communication is not new. For instance, [10] and [11] describe and implement such ideas in innovation diffusion. In both cases, vertical communication is modeled as exogenous information. The originality of our work is in that information emitted as vertical communication is endogenous. It is issued from a deliberation model that agents shape on the basis of their opinions, arguments, and behavior. In the spirit of [4], where the authors control for the design of vertical communication, we control for the process generating vertical information.

5 Conclusion

The main objective of this article was to build a bridge between decision-making, argumentation, and opinion diffusion in an agent-based paradigm. We proposed a model that combined abstract argumentation theory and a bounded confidence opinion diffusion model and showed to what extent it could explain variability in opinion and correctness of collective decisions. The model revealed that (1) to ask for more deliberation, (2) to allow for more agents to participate in deliberative instances, and (3) to make deliberative interactions less frequent in time guaranteed an increase in the variance of opinion and in the proportion of extremists in a group. These results are consistent with findings in [18] and in [17,28], which stress that deliberation may polarize groups and may have a meager effect on shifts in opinion; and inconsistent with [13] where it is argued that deliberation moderates opinions. Deliberation alone did moderate opinion as noted in [20] yet, when integrated into a complex system in which individuals were allowed to interact with one another, its influence was overshadowed by other individual-prone opinion dynamics. Undeniably, the skeptical way of reasoning over arguments during deliberation used in the model played an important role in the weakness of the effect of deliberation. Nevertheless, deliberation still increased judgment accuracy yet in a marginally decreasing fashion. We also showed that voting within the deliberation protocol not only increased the proportion of extremists and the variance of opinions but also determined how coherent deliberation and voting are with one another. The voting quota for proposal acceptability determined which part of the mixed model (deliberation if small, pair-wise influence if big) dominated the dynamics.

Extensions of this model include better-thought deliberation protocols where deliberation only affects agents that actually debate. The effect of deliberated results and its spread within a group should be observable through the entanglement of pair-wise interactions and debates. Eventually, a protocol of argument exchange would be necessary. The introduction of trust, networks, multi-dimensionality of opinions, and learning are ways to extend the model and relax unrealistic assumptions. To conclude, exploring different argumentation ontologies and opinion dynamics and finding case studies for the model are essential points to build from in future work.

References

1. K. Atkinson, T. Bench-Capon, and P. Mcburney. A dialogue game protocol for multi-agent argument over proposals for action. *Autonomous Agents and Multi-Agent Systems*, 11(2):153–171, 2005.
2. E. Awad, M. Caminada, G. Pigozzi, M. Podlaszewski, and I. Rahwan. Pareto optimality and strategy-proofness in group argument evaluation. *Journal of Logic and Computation*, 27(8):2581–2609, 2017.
3. R. Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203–226, 1997.
4. V. Barbet and J. Rouchier. Tension between stability and representativeness in democracy. 2018.

5. P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. *The knowledge engineering review*, 26(4):365–410, 2011.
6. E. Bonzon and N. Maudet. On the outcomes of multiparty persuasion. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Vol.1*, pages 47–54. International Foundation for Autonomous Agents and Multiagent Systems, 2011.
7. M. Caminada. On the issue of reinstatement in argumentation. In *European Workshop on Logics in Artificial Intelligence*, pages 111–123. Springer, 2006.
8. S. Chambers. Deliberative democratic theory. *Annual review of political science*, 6(1):307–326, 2003.
9. G. Deffuant, F. Amblard, G. Weisbuch, and T. Faure. How can extremism prevail? a study based on the relative agreement interaction model. *Journal of artificial societies and social simulation*, 5(4), 2002.
10. G. Deffuant, S. Huet, and F. Amblard. An individual-based model of innovation diffusion mixing social value and individual benefit. *American Journal of Sociology*, 110(4):1041–1069, 2005.
11. S. Delre, W. Jager, T. Bijmolt, and M. Janssen. Targeting and timing promotional activities: An agent-based model for the takeoff of new products. *Journal of business research*, 60(8):826–835, 2007.
12. P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.
13. J. Fishkin and R. Luskin. Experimenting with a democratic ideal: Deliberative polling and public opinion. *Acta politica*, 40(3):284–298, 2005.
14. S. Gabbriellini and Paolo Torroni. A new framework for ABMs based on argumentative reasoning. In *Advances in Social Simulation*, volume 229, pages 25–36. Springer Berlin Heidelberg, 2014.
15. S. Galam and S. Moscovici. Towards a theory of collective phenomena: Consensus and attitude changes in groups. 21(1):49–74, 1991.
16. R. Glinton, O. Scerri, and K. Sycara. An investigation of the vulnerabilities of scale invariant dynamics in large teams. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Vol.2*, pages 677–684. International Foundation for Autonomous Agents and Multiagent Systems, 2011.
17. K. Hansen. *Deliberative democracy and opinion formation*. University Press of Denmark Odense, 2004.
18. W. Jager and F. Amblard. Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. *Computational & Mathematical Organization Theory*, 10(4):295–303, 2005.
19. M. Meadows and D. Cliff. The relative disagreement model of opinion dynamics: Where do extremists come from? In *International Workshop on Self-Organizing Systems*, pages 66–77. Springer, 2013.
20. S. Moscovici and W. Doise. *Dissensions et consensus: une théorie générale des décisions collectives*. Presses Universitaires de France-PUF, 1992.
21. H. Prakken. Systems for persuasion dialogue. *The knowledge engineering review*, 21(2):163–188, 2006.
22. O. Pryymak, A. Rogers, and N. Jennings. Efficient opinion sharing in large decentralised teams. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Vol.1*, pages 543–550. International Foundation for Autonomous Agents and Multiagent Systems, 2012.
23. I. Rahwan and K. Larson. Mechanism design for abstract argumentation. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Vol.2*, pages 1031–1038. International Foundation for Autonomous Agents and Multiagent Systems, 2008.
24. T. Rienstra, M. Thimm, and N. Oren. Opponent models with uncertainty for strategic argumentation. In *IJCAI*, pages 332–338, 2013.
25. J. Rouchier and E. Tanimura. Learning with communication barriers due to overconfidence. *Journal of artificial societies and social simulation*, 19(2), 2016.
26. D. Rousseau. Reinforcing the micro/macro bridge:organizational thinking and pluralistic vehicles. 2011.
27. M. Sherif and C. Hovland. Social judgment: Assimilation and contrast effects in communication and attitude change. 1961.
28. C. Sunstein. The law of group polarization. *Journal of political philosophy*, 10(2):175–195, 2002.
29. G. Vreeswijk. An algorithm to compute minimally grounded and admissible defence sets in argument systems. In *COMMA*, pages 109–120, 2006.
30. Y. Zhang, H. Duan, and Z. Geng. Evolutionary mechanism of frangibility in social consensus system based on negative emotions spread. *Complexity*, 2017, 2017.

Agents with Dynamic Social Norms

Samaneh Heidari[1], Nanda Wijermans[2], and Frank Dignum[1,3]

[1] Utrecht University, the Netherlands
s.heidari@uu.nl
[2] Stockholm University, Sweden
nanda.wijermans@su.se
https://orcid.org/0000-0003-4636-315X
[3] Umeå University, Sweden
frank.dignum@umu.se, f.p.m.dignum@uu.nl

Abstract. Social norms are important as societal agreements of acceptable behavior. They can be seen as flexible, but stable constraints on individual behavior. However, social norms themselves are not completely static. Norms emerge from dynamic environments and changing agent populations. They adapt and in the end also get abrogated. Although norm emergence has received attention in the literature, its focus is mainly describing the rise of new norms based on individual preferences and punishments on violations. This explanation works for environments where personal preferences are stable and known. In this paper, we argue that values are the stable concepts that allow for explaining norm change in situations where agents can move between social groups in a dynamic environment (as is the case in most realistic social simulations for policy support). Values thus reflect the stable concept that those are shared between the agents of a group and can direct norm emergence, adaptation, and abrogation. We present the norm framework that enables describing and modeling value and situation based norm change and demonstrate its potential application using a simple example.

Keywords: Social Norm · Social Norm Dynamics · Norm Framework · Value based Norms · Personal Values.

1 Introduction

Social phenomena are part of our thinking [11]. Therefore, it is mandatory to consider social aspects to study decision making and system behavior. Especially, if the purpose of the study is to explore the mutual effects of micro-level decisions and macro-level behavior of a system. Among different social aspects, we are interested in studying social norms, as norms play an important role in guiding all human societies [6]. Social norms are more important to study and consider in the absence of a central monitor/control [14].

Considerable research effort has been dedicated to developing models, architectures, and theories that concern social norms in making decisions. However, there are some points that have been omitted in the research efforts in two main issues: putting the focus on norm reactivity to environmental changes without regard for factors that drive norm stability, and favoring implicit, rather than explicit, representations of norms.

Studying the reactivity and stability of social norms cannot be effective without considering values, an element which is lacking in the previous works. In the absence of any stabilizing factors, modelled norms might quickly react to any change. However, many real norms remain rather stable over long periods of time due to their connection to fundamental values, which are, by their nature shared between groups of people and very stable over a persons lifespan. As for the issue of norm representation, researchers assume that social norms are explicitly defined in advance and use norms as constraints. Such an assumption is useful for simplifying the study of the effects of specific norms in a given scenario, but takes away the possibility of studying norm dynamics (such as norm emergence) and norm recognition [13]. Social norms are distributed concepts rather than central. Each person might have his own interpretation of a social norm.

The simulation of social norms and their effects on decision making and on the behavior of the system has gained much interest in the field of social simulation. Therefore, we believe that a framework that deals with values and norm dynamics is relevant for many social simulations. We introduce a normative framework that covers key dynamics of social norms, their effect on micro-level and macro-level, and their relation with values. The social norms are dynamic in our normative framework. In other words, norms might undergo changes due to changes in the environment including change in the group members, economy, and ecology.

In this paper, we start with some background information and introduce the concept of values as we use it (section 3), and how they relate to social norms (section 4). We introduce the framework (section 4). Then, we discuss alternative representations and dynamics of norms in a normative decision model, and how our framework covers the dynamics of norms(section 5). We summarize the paper in section 6.

2 Related work

The first question that need to be answered to make a normative framework is: what is the definition of social norms.

Bicchieri defines norms as: "the language a society speaks, the embodiment of its values and collective desires". She specifies norms as behavioral rules that will be triggered in certain social roles or situations [3]. Interesting enough she also mentions that norms are embodiments of values. This is in line with Schwartz, who also argues that specific norms for concrete situations are connected to a set of abstract values [15].Thus when we use norms we should also model the values from which they are the embodiment. Somehow this aspect is hardly ever used, but we will show its importance in this paper. Bicchieri also mentions that sociologists have not agreed upon a common definition [6].

However, Gibbs discusses different viewpoints of sociologists on social norms [9]. We used his discussion, to extract the points that his discussion emphasized and we cover them in our framework. According to his discussion: norms are agreements of group members; norms regulate behaviour; norms are group expectations in certain circumstances about what should and what should not be done; norms are based on cultural values; norms are abstract patterns of behaviour; and norms are alternative ways to achieve goals.

These points together will cause dynamics of norms in a group. As norms are agreements of the members, these agreements can change if the members change their mind. So, we want our normative framework to have the possibility of covering emerging norms, changing norms, and preserving norms.

Of course we are not the first ones to describe normative frameworks. Some groundwork was done in [2, 4, 8, 17]. Neumann compared some architectures that covers social norms [13].

EMIL-A-1 [17] is a normative architecture based on the EMIL-A architecture which has an explicit norm emergence possibility. However, in this framework there is no connection with values yet. Segura introduces a normative architecture that includes sanctions and punishments to increase cooperation in social systems [17]. He mentions that norm emergence and making a self-policy system can be achieved by means of different punishment technologies. However, he mentions that social norms are social cues that guide behaviour even in the absence of explicit punishment systems. From this we take that our framework should not exclusively rely on a punishment system. Most norm abidance comes from the wish to group conformance. Thus, indirectly the group determines the abidance of the norm. We will incorporate this element by letting agents abide by a norm dependent on the visibility of the norm. It does not mean that punishment does not play a role, but rather that it is not the main driver of norm emergence and norm compliance.

In our framework, we define norms as social behavior that might involve punishment or not. In other words, some norms will be followed because people need to satisfy their conformity value[4] and be a good member of their group [7].

As said above, values are the main source of norms as values are "ideals worth pursuing"[7]. Therefore, values can be seen as one of the main ultimate motives of deliberated actions. Norms and values are evaluation scales. However, norms are more concrete embodiments of values. Norms refer to certain behavioral choices in particular contexts; values are criteria to prioritize particular types of actions and situations[12]. For example, a person who highly values unisersalism would like to give away some money for altruistic reasons. However, there might be some social norms that determine how much to donate, when to donate, etc.

As the basis of our framework we use the value system as developed by Schwartz. Schwartz represents a universal theory on value system that is widely known and accepted by researchers [1, 16]. We will explain this value system in more detail in the next section (section 3). Also, we will explain our previous work on representing a value framework based on Schwartz's value theory in section 3.

[4] Conformity is one of 10 abstract values that Schwartz presents in [16]. Conformity drives obedience to rules and social expectations or norms.

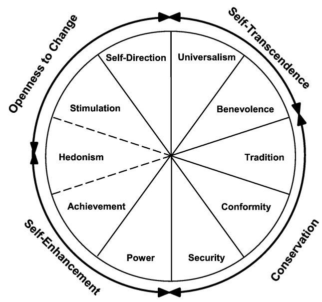


Fig. 1. Schwartz value circle, categorization and dynamicity of abstract personal values [16]

3 Summary of the Value Framework

Our norm framework is a based on our previous work on the value framework[10]. In this section, we briefly review the value framework [5] and discuss how this framework is used as a basis for our norm framework.

Considering Schwartz's value circle, we introduce a framework for decision making based on (personal) values. Schwartz introduces 10 abstract values that are supposed to be universal (figure 1). However, the importance and priorities of the values differ. The importance of a value is a degree that shows the salience of a value in a certain situation and time. A value like universalism is less important after just having spend a day doing community work. At that time it might be allowed to relax and enjoy some nice dinner with friends. The priorities between values indicate a base preference between the values. I.e. whether universalism is more important than conservatism in cases where both values are salient for choosing a course of action. Using the visualization of the Schwartz value circle, there are some relations between the priorities of these values. The closer to each other the values are in the circle, the closer is their priority. [16].

Similar to the Schwartz circle, in our framework each value has a degree of importance. We defined mathematical equations that maintain the circular relation of the importance of values. To reflect the heterogeneity of agents, agents can have different value importances. In other words, they can assign different degrees of importance to their abstract values. Therefore, two agents with different importance distributions might take different decisions under the same external condition.

In our value framework [10] agents make a deliberate value-based decision. We operationalised the framework using an agent-based model (ABM). For the ABM we defined value trees to connect Schwartz abstract values to actions. The root of these trees are the Schwartz abstract values and the leaves of the trees are actions that agents can perform. Nodes that are closer to the leaves are more concrete. Figure 2 depicts a possible tree of power value for job selection in a simulation

[5] For further details and implemented version of the framework see[10].

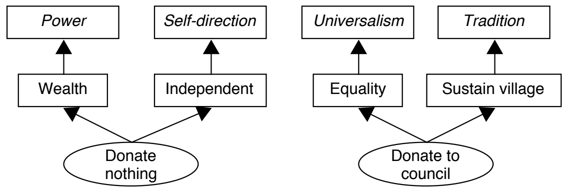


Fig. 2. A sample of value trees related to donation action

based on our value framework. It should be noted that the arrows in figure 2 is showing the direction of value satisfaction. In other words, if an agent performs an action, he will sweep the related value tree up to the root. Then, the assigned water tank to the root will be filled.

We use water tank model to represent value satisfaction and thus salience of values. We assigned one water tank to each value tree. Each water tank has a threshold level (which is the importance of its value) and its water drains over time. Every time that an action is taken, some water will be poured into the related water tank. Each agent decides what to do based on the difference between water level and threshold of his water tanks. A positive difference means that the value is satisfied; consequently, a negative difference means that the agent did not satisfy the value enough times.

In the next section we will extend this framework with norms. The norms are placed in between the values and the actions. Thus norms can be seen as concrete rules for deciding on actions that will promote a certain value. Thus, instead of having to reason with whole value trees we can use the norms as concrete representations of them. However, by placing the norms in the context of the value trees the agent can also reason about violating a norm in a concrete case, of adopting a norm or adapting it and even abrogating it.

4 The Norm Framework

In this section, we introduce a norm framework for building normative agent-based models and agent-based simulations. In this framework, agents deliberate based on their individual values and the social norms of the groups they are part of. Social norms are formed based on individual values. Agents participate in the dynamics of social norms by following, violating, or even by performing actions that slightly deviate from social norms, thus making social norms dynamic in this framework. In other words, norms might undergo changes due to any change in the environment including group structure, economy, and ecology. On the other hand, because they are tied to values, these changes are also opposed, constrained and directed in a controlled manner. We make use of a preliminary simulation in section 5 to show how these norm dynamics can have profound influences on the behavior of the agents, as well as the structure and behavior of groups.

4.1 Norm definition

As mentioned earlier, we use the following aspects of norms as used by Gibbs [9] as a basis for the model of social norms in our framework:

1. norms are agreements of group members,
2. norms regulate behaviour,
3. norms are group expectations in certain circumstances about what should and what should not be done,
4. norms are based on values,
5. norms are abstract patterns of behaviour, and
6. norms are alternative ways to achieve goals.

The above points do not mention sanctions explicitly. We follow Gibbs and Bicchieri who mentions that social norms may or may not be supported by sanctions [3]. Thus we do not take sanctions as the main drivers of norm emergence and compliance and they are not part of our core norm model, although they can be added to it to strengthen the effect of norms in certain contexts.

Expanding on point 4 (norms are based on values), provided by Gibbs, we use Bicchieri's research on social norms to connect norms and values. She mentions that norms are embodiments of values [3]. This point of view is supported by other research that illustrates that norms are connected to a set of abstract values with the *aim* of achieving those values [7, 15].

Bardi and Schwartz believe that values do not play a role in making behavioral choices directly and consciously for most people. However, people act mostly according to their value system, which is mostly unconscious [1]. In other words, most people have a certain value system, but they do not refer to it for every single decision. Our interpretation of their work is that a person should live a normal life even without deliberating about all his actions through his values. This can be realized by assuming that social norms cover most of the actions that are needed for interactions with other people in daily life.

4.2 Norm type, structure and relation to values

Considering the arguments in the previous section, we explain how we formulate social norms and how we formulate norms as embodiments of values. We formulate social norms as actions that agents consider to do or not to do in certain conditions.

Therefore, we define a norm n as follows : $n =< v, c, t, a, pe, ne >$ in which n as a social norm guides the agents to satisfy value v by performing action a, under condition c. Depending on the norm type t the agent might get positive consequence pe by following n, get a punishment ne by violating n, or there is no positive social consequence nor any negative social consequence by following or violating n.

Taking the provided definition of social norms, we see that norms are not a necessary completion of values but rather norms and values are complementary. A norm is an edge in the value tree that connects two nodes of the tree whose distance is at least 2. It means that there is a path between these two nodes with length of at least 2 . If one of the nodes is an action (leaf of the value tree), the norm is a specific norm; otherwise, it is an abstract norm. Thus the norm can be seen as a shortcut for a value. Reasoning from an action upwards an agent can stop at a node where a norm is connected. Following the norm guarantees promoting the value. Thus if most actions are connected with concrete norms to the value tree above, very little (expensive) reasoning about values has to be done. However, this construction also allows the comparison of two norms by checking which values they promote and which of those values has higher priority and importance. This allows for reasoning about violation of a norm in case norms are inconsistent in specific situations or in case another value is more important than the value promoted by the norm. E.g. speeding in the highway in order to be in time for dinner.

The importance of following a norm differs depending on the importance degree of its supported values and the Norm type. Therefore, consequences of violating and following norms differ. We consider four types of norms: should follow (to represents soft social norms), have to follow (to represents strict social norms), and must follow (to represents laws).

Also, personal characteristics of people play a role on how much they might consider social norms in their decisions, especially if the social norms are in conflict with their personal values. For example, if a person values universalism a lot, he will internalize "donating to public benefits" norm as such a norm serves universalism value. Internalizing a norm raises the probability of considering it in decisions.

4.3 Decision Making

To recognize a normative behavior, an agent considers what most people do [3]. This is what Cialdini et al. define as descriptive norms [5]. However, agents consider the social standing of the person who is performing an action, $S(a)$. $S(a)$ represents the social status of agent a. If an agent has a good social standing, other agents consider his actions with a higher probability.

In our framework, each agent makes a decision about what action to perform considering their personal values and social norms. Each agent has its own value trees. These value trees are not necessarily complete from root to very concrete leaves. In other words, some of the agents might not have value trees explicitly. Either an agent has complete value trees or not, they have those shortcuts that they adopt from the society. Those shortcuts are norms. Norms can cover primary needs of people so that they do not need to reason upon their values to make a decision. Therefore, the agents that have complete value trees explicitly are the ones representing deep thinking people in the real world. In other words, these agents can deliberate about their actions explicitly.

Each agent can be a member of several groups. Therefore, each agent has a list of norms that he adopts from his groups. Such a list is dynamic for two reasons. First, social norms are not explicitly available, but rather individuals have their own understanding of norms. Second, norms are influenced by the environment. In other words, any change in the environment including changes in group members, economic situation, and ecological situation might lead to changes in the social norm. If changing the group members alters the collective values of the group, the group norm will change slightly; as the norms are connected to values. If there is any change in ecology or economy that makes following a norm not viable, a norm might abrogate slowly. For example, assume that there is a norm on donation in a community because people value equality a lot. If many new people who are self-oriented join the society, they can slightly change the norm to donate less frequently, or donate less. Or assume that economic inflation happens and people cannot earn enough money. The donation norm may change to alternative actions such as sharing food, donating cloth, etc.

We assume that each group has its own social norms. Each group might have a different norm on how to do one certain action. It should be noted that we do not consider explicit representations for norms. We do not consider a group as a central element that control and keep norms. But rather, agents perceive norms of a group by monitoring the behaviour of its members over time. To consider group membership and norms, each agent has a list $< N, g >$ in which N is a set of social norms that the agent assigns to group g.

To give an example of group norms, assume "turning trash into treasure to save the environment" as a norm that is serving the universalism value. Assume an agent is working in a company. His colleagues have the norm of "separating plastic bottle caps to donate to charity". The same agent is living in a neighborhood with a norm of "separating glass waste color-wise". Both norms serve universalism value, however they are valid in different contexts.

Each agent considers social norms in his decisions depending on how many times he observed a norm n has been followed by his group mates. An agent will increase the probability of following n, if he observes n has been repeated over time regularly. Normative action of a group for an agent is a weighted average action:

$$n = \frac{\sum_{a_i, a_i \neq a_j}^{a_i \in g} S(a_i) * (\text{performed action by } a_i)}{\sum_{a_i, a_i \neq a_j}^{a_i \in g} S(a_i)},$$

where, $S(a_i)$ is social standing of agent a_i.

So, each agent needs to keep how many times a norm is repeated. A norm has a chance of abrogation if agents stop following it for long enough time. Therefore, we need to keep a variable showing how many time steps a norm has not been repeated. So, each agent keeps a norm repetition as a set of $< n, r, nr >$ that shows norm n has been fulfilled r times and not been used nr times.

As mentioned earlier, an agent regards several factors to make a decision including personal preference, norms, motivations, culture, etc. In this paper, we consider norms and personal values as two factors that effectively regulate behavioral choices. An agent a_j considers both its personal preference and social norm of group g to make a decision in that group. Therefore, we formulate the normative decision according to the following equation:

$$decision = P_n(t) * n(a_j) + (1 - P_n(t)) * \text{personal preference}.$$

Where $n(a_j)$ is norm n that agent a_j considers in his decision. $P_n(t)$ is a probability function that depends on the history of norm n till time t. More explanation on $P_n(t)$ is provided in the section 4.4.

4.4 Norm life cycle

In this framework, we consider four phases for a norm, observation, adoption, internalization, and abrogation. Therefore, we define a function $P_n(t)$ (probability of following a norm n) for each agent as follows:

$$P_n(t) = \begin{cases} F_{observe}(t) & \text{if } t \in \text{observation phase} \\ F_{adopt}(t) & \text{if } t \in \text{adoption phase} \\ F_{internal}(t) & \text{if } t \in \text{internalization phase} \\ F_{abrogate}(t) & \text{if } t \in [0, nr] \end{cases}$$

Functions $F_{observe}(t)$, $F_{adopt}(t)$, $F_{internal}(t)$, and $F_{abrogate}(t)$ determine P_n when norm n is in observation, adoption, internalization, and abrogating phase respectively. The repetition times to enter to a new phase of a norm are relative and can be changed based on the particular domain. Despite the numbers assigned to norm phases, the agent increases r by 1 if he observes that most of his neighbors performed accordingly. Otherwise, he resets r and increases nr by 1. In the latter case, the agent will create a new potential norm for an action a. If a starts repeating he will update r; otherwise, he will remove the created norm. Also, when nr reaches the maximum time, the agent will remove the norm as well.

In order to make decisions on norms that might be in different phases of the life cycle we need to have the possibility of considering external and internal norms in our framework. By external norms we mean behaviors that an agent expresses/shows to public. Internal norms are the ones that are compatible with the personal values of an agent and he would like to follow whenever possible. Internal norms can be different from what other people can externally see. For example, an ungenerous person does not want to donate anything (internal norm), but will donate a small amount in order to keep up appearance of following the group norm of donating (external norm).

In the current simulation, the internal norm is represented by using a weighted sum of the values and the external norm in order to decide on a behavior. Thus an internal norm is kept implicit and not managed separately. However, in our framework, internalized norms are the norms that the agent will follow even after leaving a group. Those are the norms that has been repeated enough and are in line with the values of an agent. Therefore, internalized norms are stored as $< N, g >$, where $g = NULL$.

5 Discussion

This section illustrates one of the possible simulations that we developed based on the introduced norm framework. Using this simulation, we discuss some of the interesting simulation examples that explain the importance of a) value-based norms, b) norm dynamics and norm stability; and c) allowing for dynamic groups (agents can enter and leave groups).

We explain how our norm framework helps exploring our questions: how personal values of group members influence social norm of a group, how values make social norms more robust against small changes, how values cause the emergence of a new norm, how values guide the changes of existing norms, and how the social norm influences the individual behaviour of the members.

Simulation settings. We implemented a community in which we study behavior related to contributions to public good in the form of donations. The amount of donation is normative. So, there are norms going around on the normative amount of donation. Personal preference of the donation amount is connected to values, but it also serves the normative amount of the group which is served by group adherence.

Agents are heterogeneous in their values and organize into different groups. Agents considers social status of all of his group-mates are equal ($S(a_i) = 1$). Agents cannot choose some groups (family), and they can choose some groups(neighbors, colleagues, etc). An agent can also belong to more than one group at a time.

One possible setting of $P_n(t)$ that we used for our simulation is:

$$P_n(t) = \begin{cases} \alpha_1 * t & \text{if } 0 < t < 5 \\ e^(t - 10.35708268) - 0.00028536; & \text{if } 5 <= t < 10 \\ 1 - 1/t^0.5 & \text{if } 10 <= t < 20 \\ 1/(1 + 0.0078 * 0.5^{(25-t')}) & \text{if } t' >= 10 \end{cases}$$

in which t' is number of times that norm n has stopped repeating. According to this setting, the probability of following a norm does not increase mush as the agent is still not sure about the norm. However, P_n increases exponentially during adoption phase. As mentioned prior, a norm enters to the internalization phase if it has been repeated enough by other agents and if it is compatible with the personal value of an agent. Therefore, an internalized norm has a higher chance of being followed by an agents.

Assume group g_1 has 4 members, agents a_1, a_2, a_3, a_4, who value power a lot (with the importance of 80%). Therefore, norm of the group emerged as $n_1 =<$ power, "having more than enough money", should follow, donate $5\% - 10\%$, raise social status, null $>$. Consider agents a_5, a_6, a_7, a_8 highly value universalism (with the importance of 80%) and they used to donate about 50% on average (either because of their internalized norm or because of their other groups)

Scenario 1. robustness of norms Our simulation shows if a_5 joins g_1, he starts adopting norm n_1. He donates 10% mostly (according to external norm of the group). Agent g_5 seldom deviates from norm n_1 to keep his social image. But, he rarely donates 50% (according to his internalized norm) to satisfy his universalism value. However, his attitude does not change the norm. After he start adopting the norm, a_6 joins the group. The same will happen to a_6 and any other agents that joins the group with the same pattern. In this scenario, norm is stabled over time. Even though the social norm is different from the internalized norm for $a_5..a_8$.

An exceptional case can lead to changing the group norm. If a new universalist agent a_i join the group at time tick t. Assume agents $a_5..a_8$ join the group at time tick $t+1$ to $t+4$ respectively and donate 50%. Agent a_i observes that the average donation is 15%. With our simulation setting that observation time is 5 time ticks, he will start adopting norm of "donate about 15%" as the norm of g_1. If more agents similar to a_i join the group and the same story happens to them, n_1 will deviate a bit from its original amount.

Scenario 2. changing of norms We ran the simulation to check what will happen if a lot of new members enter a group at the same time. We let agents $a_5..a_8$ joins group g_1 together at time tick t. During the observation phase, agents $a_5..a_8$ donate 50% according to their internalized norm. Therefore, they observe that donation amount is about 27% on average. So, they adopt "donating about 27%" as the norm of g_1. However, the other agents $a_1..a_4$ start realizing that normal donation is changing from time tick $t + 1$. When they observe the new donation amount for more than 10 time ticks (which is the minimum time to abrogate a norm in our simulation setting), they abrogate their perceived norm and start observing the group behaviour again. From time $t + 5$ onward, the new members mostly donate the normative of 27%. From time $t + 6$, the new members will see that the average amount is different from what they start adopting (which is about 12% now). Therefore, they do not adopt normative amount 27%, but rather start observing whether 12% is a norm till time tick $t + 10$. Continuing this run, the normative donation amount of the group converges to 27%. The convergence happens because agents ignore some of the random deviation from norm.

The above simulation scenarios shows partly how individual values guide emergence, robustness, and changes of social norms. In these two scenarios, the same agents joined a group with different patterns. If new agents join gradually, they can hardly change values' balance of the group. Therefore, norm the group stays stable. But, if new agents join altogether a the same time, they can change existing norm if it is against their values.

6 Conclusion

In this paper, we introduce a norm framework. Such a framework considers social norms as non-static social elements. In our framework, norm dynamics arise from dynamic environments. Such a framework is not completely new in the field of social simulation. However, we connect norms to personal values and consider norms as embodiments of personal values. This connection makes the norms robust against small dynamics in the environment. In addition, it is more realistic as there is no need to have a central element to monitor and keep social norms. But rather, social norms are distributed between agents as their perception of social norms. We discuss how such as a framework can express the way values guide norms (emergence, changing, abrogation, and internalization). We explained it using a preliminary simulation scenarios.

Bibliography

[1] Bardi, A., Schwartz, S.H.: Values and behavior : Strength and structure of relations. Personality and Social Psychology Bulletin **29**(10), 1207–1220 (2003)

[2] Beheshti, R.: Modeling Social Norms in Real-World Agent-Based Simulations. Ph.D. thesis, University of Central Florida (2015)

[3] Bicchieri, C.: The Grammar of Society: The Nature and Dynamics of Social Norms, vol. 1. Cambridge University Press (2006)

[4] Broersen, J., Dastani, M., Hulstijn, J., van der Torre, L.: Goal generation in the boid architecture. Cognitive Science Quarterly **2**(3-4), 428–447 (2002)

[5] Cialdini, R.B., Kallgren, C.A., Reno, R.R.: A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. Advances in Experimental Social Psychology, vol. 24, pp. 201–234. Academic Press (1991)

[6] Conte, R., Andrighetto, G., Campenni, M.: Minding norms: mechanisms and dynamics of social order in agent societies. Oxford University Press (2014)

[7] Dechesne, F., Di Tosto, G., Dignum, V., Dignum, F.: No smoking here: Values, norms and culture in multi-agent systems. Artificial Intelligence and Law **21**(1), 79–107 (2013)

[8] Dignum, F., Dignum, V., Prada, R., M. Jonker, C.: A conceptual architecture for social deliberation in multi-agent organizations. Multiagent and Grid Systems **11**, 147–166 (2015)

[9] Gibbs, J.P.: Norms: The problem of definition and classification. American Journal of Sociology **70**(5), 586–594 (1965)

[10] Heidari, S., Jensen, M., Dignum, F.: Simulation with values. Social Simulation Conference (2018)

[11] Hofstede, G.J.: Grasp agents: social first, intelligent later. AI & Society : the journal of human and machine intelligence (2017)

[12] Marini, M.M.: Social values and norms. `https://www.encyclopedia.com/social-sciences/encyclopedias-almanacs-transcripts-and-maps/social-values-and-norms` (2018)

[13] Neumann, M., Mainz, J.G.u., Neumann, M.: The cognitive legacy of norm simulation. Artificial Intelligence and Law **20**, 339–357 (2012)

[14] Pasrav, C., Dignum, F.: Norms in social simulation: balancing between realism and scalability. Social Simulation Conference (2018)

[15] Schwartz, S.H.: Normative influences on altruism. Advances in Experimental Social Psychology **10**, 221–279 (1977)

[16] Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. Advances in Experimental Social Psychology, vol. 25, pp. 1–65. Academic Press (1992)

[17] Villatoro, D.: Social Norms for Self-policing Multi-agent Systems and Virtual Societies. Ph.D. thesis, Universitat Autònoma de Barcelona (2011)

A collective action simulation platform

Stephen Cranefield*, Hannah Clark-Younger†, and Geoff Hay†

Department of Information Science,
University of Otago,
PO Box 56, Dunedin, New Zealand
stephen.cranefield@otago.ac.nz

Abstract. In this paper, we discuss some types of expectation that contribute to the behaviour of social agents, and investigate the role that these social expectations can play in the resolution of collective action problems. We describe our Collective Action Simulation Platform (CASP), a framework that allows us to integrate the Java-based Repast Simphony platform with a Prolog-based event calculus interpreter. This allows us to run simulations of agents who make reference to social expectations when reasoning to make decisions about which actions to perform. We demonstrate the use of CASP in modelling a simple scenario involving agents in a collective action problem, showing that agents who are informed by social expectations can be led to cooperative behaviour that would otherwise be considered "non-rational".

1 Introduction

Collective action problems involve members of a community who must coordinate or collaborate in order to achieve a collective, rather than individual, benefit [10, 13, 16, 23, 33]. In the common case when the benefit is non-excludable (all can share in it, regardless of their contribution), this leads to the *free rider problem*: individuals can benefit from the collective action of others, while failing to make their own contribution [14]. Problems of this sort include coordinating access to a common resource pool (e.g. river water or a fishery) and collectively reducing carbon emissions from energy use.

Mathematical analysis of this problem has led to the conclusion that, in the absence of coercion or individual inducements, each community member's rational decision is to be a free rider [12, 23]. This is a *social dilemma* whereby individual rational reasoning leads to a sub-optimal outcome for the community. In the context of collective access to a shared resource, Hardin [11] argued that the individual payoff from increasing personal use of the resource will outweigh any reduced value of the resource due to overuse. As this reasoning is repeated

* Stephen Cranefield acknowledges funding from the Marsden Fund Council from Government funding, administered by the Royal Society of New Zealand.

† The second and third authors contributed to this paper while working at the University of Otago. Hannah Clark-Younger is now at Soul Machines, hannah.clark-younger@soulmachines.com.

by all resource users, the seemingly inevitable outcome is the ruination of the resource, referred to by Hardin as "the tragedy of the commons".

However, inspired by the observation that cooperative behaviour is observed in human and animal societies, researchers have proposed a wide range of mechanisms that allow the social dilemma to be broken. Holzinger [16] discusses mechanisms proposed across a range of disciplines (e.g. philosophy, sociology, economics and politics) and categorises them into (a) individual solutions based on internal motivations (e.g. altruism), (b) individual solutions based on rational expectations (e.g. the existence of social conventions), (c) collective solutions based on social choice mechanisms (e.g. voting), and (d) collective mechanisms based on enforcement mechanisms (e.g. rules and sanctions). Reuben [33] also discusses a range of proposed solutions, including the existence of private incentives for cooperation, changing the game to include (e.g.) repeated interactions, consideration of heterogeneous social preferences amongst the agents, the use of evolutionary and learning models to explain the emergence of cooperation, and the existence of non-uniform social network structures. After extensive fieldwork, Ostrom [24] identified eight principles that she found to be common to the governance rules of successfully managed real-world community resources.

Most prior work outlined above has made advances in the high level understanding of collective action by focusing on the very abstract mathematical models of *game theory* [18]. These models typically assume that participants select their actions simultaneously, and choose to maximise their immediate reward, with the reward structure defined by a payoff matrix. These models cannot be seen as sufficiently realistic models for human behaviour. In particular, there is a lack of consideration of the (bounded) reasoning processes that can lead community members to participate in collective action [33]. This research gap is significant given an increasing interest in the field of *social computing* [9, 34], which aims to develop software that assists members of a community to collaborate effectively.

Our aim is to investigate the collective action problem by drawing on prior work on computational models of social reasoning, using mechanisms such as social expectations [8] and norms [10, 15, 38]. There has been little prior work in this area other than the work of Pitt and colleagues, who have developed computational models of Ostrom's principles for self-governance of common-pool resources [24, 32] and on social capital [28–31]. The latter work investigated three forms of social capital identified by Ostrom and Ahn [25]: measures of trustworthiness, social network connections, and the observed conformance to the rules of one or more potential 'institutions' governing the interactions in the community. The results showed that choosing whether to cooperate based on a linear function of these social capital measures enhances collective action in settings where pure game-theoretic reasoning allows no polynomial-time algorithm for generating stable and socially optimal behaviour [27]. The three types of social capital studied in this work can be seen as means to coordinate the expectations of community members regarding the behaviour of others.

This paper elaborates on this intuition by proposing a novel model for collection action problems in which agents consider *social expectations* when deliberating about their action choices, and describes an agent-based simulation framework based on the model. An implemented group fishing simulation scenario is presented, and it is shown how cooperation can emerge when socially aware agents have knowledge of, and trust in, particular types of social expectation.

2 Expectations

Expectations drive our daily behaviour in many different ways. We hang our washing on the line in the morning because we expect that it will be dry by the afternoon, we turn up to work because we expect our employer to pay us, and we turn off our cellphones at the movies because we are expected to do so.

We can think of expectations as falling into two interesting categories—expectations about the consequences of our actions, and social expectations. Hanging out our washing falls into the first category, because our expectation is that the moisture will evaporate in the warm, dry air. Expecting our employer to pay us and turning off our cellphones at the movies are examples of social expectations, because they involve other people. In the first case, we have an expectation *of someone else*, and in the second case, other people have an expectation *of us*. We also expect our actions to have social consequences: I expect that if I use my cellphone during the movie, I will be glared at and possibly not invited on another social outing by my companions. Based on these observations, this paper groups observations into the following four types:

Type A Expectations we have about the physical consequences of our actions
Type B Expectations we have about other people
Type C Expectations other people have about us
Type D Expectations we have about the social consequences of our actions

These expectations encode the rationale for agents to make decisions based on their previous experience and social knowledge [8].

Expectations are related to predictions: like predictions, expectations concern a belief about the future. However, we follow Cristiano Castelfranchi in distinguishing expectations as those future-directed beliefs that are "evaluated against some concern, drive, motive, goal of the agent." [5, p. 264] That is, expectations are predictions where we *have a vested interest in the outcome*.

Expectations are also related to obligations. We could rephrase our examples above to say that we turn off our cellphone in the movie because we have a social obligation to do so. In fact, obligations of a certain kind can be identified with social expectations (types B and C), at least for our purposes. For a fuller discussion of these relationships, see [8].

3 Reasoning about expectations

Our research aim is to investigate the role of expectations in fostering collective action. Our approach is to build agent-based simulations [6, 21] in a number of different collective action scenarios and vary the type of expectation-based social reasoning used by agents to select their actions. In this paper we consider a team fishing scenario, which is described in Section 5. To facilitate our simulations we have developed a Java framework that extends and specialises Repast Simphony [22] to support expectation-based action selection. This is described in Section 4.

An important aspect of expectations is their dynamic behaviour. Domain-specific expectations are created (or activated) when specific conditions hold. Once expectations are active, they persist until they are fulfilled or violated. As they represent constraints on the future, and the future is gradually revealed as time goes by, these constraints may (in general) be partially evaluated and simplified over time (this is known as progression). If an expectation is reduced to *true*, it has been fulfilled. If it is reduced to *false* is has been violated. Rather than specifying these processes as algorithms to be implemented in our Repast application code, we choose to use a declarative mechanism: the event calculus (EC) [17, 35], which has been used by many researchers to model the dynamics of social and institutional constructs such as commitments [7, 37] and norms [1, 2].

The event calculus is a logic-based formalism for defining the effects of actions and reasoning about how actions change the state of the world. This can be extended to track the creation, fulfilment and violation of expectations expressed using linear temporal logic [8]. A significant advantage of the event calculus is that it is directly executable: it supports *temporal projection*: an inference process in which a trace of events is combined with a logical description of the effects of actions to extend the trace with inferred action effects (social ones, in our case).

The EC supports reasoning about *events* and *fluents*. The latter are used to represent any type of state that can change over time. A fluent is associated with a value by a term of the form $F = V$; however, in this paper we only use Boolean-value fluents where V is *true* or *false*. The EC includes an *inertia* principle—the value of a fluent remains unchanged unless an action occurs that changes it.

In this work, we use EC rules to define the effects of social actions, with fluents representing the social state of the system. We extended an implementation of the event calculus (RTEC [3]) to include features of the expectation event calculus (EEC) [8]. This extension treats fluents of the form $\mathit{exp\_rule}(\mathit{Condition}, \mathit{Expectation}) = \mathit{true}$ specially: when *Condition* comes true, a new expectation fluent is created for the next tick. Expectation fluents express constraints on future actions and fluent values, and may contain the linear temporal logic operators *next*[1], *until*, *eventually* and *always*. The EEC's temporal projection progresses expectation fluents from the previous state to the current one. This involves re-expressing them from the (temporal) viewpoint of the *next* state,

[1] As we are using discrete time simulations, there is always a unique next state—in effect we are using a version of the discrete event calculus [20].

e.g. $next(\phi)$ is transformed into ϕ. When expectations become fulfilled or violated, this is recorded using special fulfilment and violation fluents.

4 The Collective Action Simulation Platform (CASP)

To support simulation experiments with agents that can use knowledge about expectations when choosing actions, we developed the Collective Action Simulation Platform (CASP) as a framework that extends the Java-based Repast Simphony simulation platform [22]. The key aspects of CASP can be summarised as follows. CASP integrates Repast Simphony with an event calculus interpreter to represent social state and to enable the use of EC rules to specify the social effects of actions. Agent reasoning is performed in two stages. A rule engine is used to determine actions relevant to the current state (including social state stored in the EC interpreter), based on rules associated with the agent's current roles. Then the agent selects one of the actions to perform. This selection may also consider the social state (e.g. an agent may get greater utility from performing actions that another agent expects it to do[2]).

CASP provides Agent and ControllerAgent abstract classes. These include a reference to a façade class that encapsulates an event calculus (EC) [35] interpreter (an extension of RTEC [3] running in SWI Prolog [36]). The Agent class also provides access to a Maxant rule engine[3], to support rule-based generation of possible agent actions. There are also abstract Institution and Institution-Role classes. To develop a simulation using CASP, the programmer provides scenario-specific subclasses of these four types of abstract class.

A simulation using CASP executes as follows:

- The Repast scheduler is used to run the controller agent's `step()` method once for each simulation cycle, followed by the `step()` method for each of the other agents.
- The controller agent runs one step of the event calculus interpreter to perform *temporal projection*: given the values of fluents and the actions that agents performed during the previous simulation 'tick', it applies a supplied domain theory (a set of EC rules) to update the Prolog fact base with fluent values for the current tick. These fluents represent the social effects of actions[4], as defined by the EC rules, as well the creation of new expectations from expectation rules, and the progression of existing expectations from the previous state.
- The controller uses the EC façade object to query the fluents that currently hold, and caches these. Other agents can query these via the façade.

[2] This would apply especially to obligations, which are specialised types of expectations. Currently CASP supports only generic expectations that a programmer can choose to interpret as (e.g.) obligations or commitments within the EC rules provided.

[3] https://github.com/maxant/rules

[4] The programmer can also choose to model the effects of physical actions using the EC, or these can be modelled entirely within the Repast agents' Java code.

	CC	**CD**	**DD**
C	$r-c$	$r-c$	$-c$
D	r	0	0

Fig. 1. The payoff matrix for a 3-person threshold game with a threshold of 2

– For each non-controller agent, the `step()` method invokes an instance of the Maxant rule engine, which is dynamically configured with all the rules associated with some role that the agent is currently playing. The rules have a condition expressed in the Java-based MVEL expression language[5]. The rule outcome is the recommendation of an action considered to be relevant when the condition holds. The condition may include queries on the social state (in the EC engine) as well as arbitrary MVEL expressions. In general, multiple rules from different roles may be triggered, and the agent's `step()` method must choose one action to perform from this set of relevant actions. This is done using application-specific code, which may consider estimated action utilities and/or queries on the social state. Finally, the selected action is performed by calling a Java method associated with the action name. The action implementation may involve asserting an event occurrence to the EC interpreter and/or the agent adding or removing roles.

The programmer is free to add other logic to the controller and other agents, e.g. to look up references to other agents using the Repast Simphony API and to make method calls on them to pass on or request information (CASP does not attempt to provide an inter-agent messaging system).

5 Scenario

In this section, we investigate how expectation-based social reasoning can result in the achievement of collective action in the context of a simple group fishing scenario inspired and abstracted from a study of the culture of the Trobriand Islands [19]. In this culture, all men in coastal villages are fishermen, and are expected to participate in a fleet of fishing boats when a leader decrees that it is a fishing day. We assume that the success of a fishing expectation is dependent on the number of participants, but a minimum number of fishermen is required for the expedition to be successful at all (e.g. if the boats need to encircle a school of fish).

This is similar to an n-person *threshold* game [4, 26], in which there is a fixed cost c of cooperating, and the payoff is 0 if a cooperation threshold is not reached; otherwise both cooperators and defectors receive a reward of r. Figure 1 shows the payoff matrix for a three-person threshold game with a threshold of 2. The rows shows the payoff for an individual's action (cooperate or defect) dependent on the actions of the other two players.

[5] https://github.com/mvel/mvel

A fundamental question in game theoretic analyses is whether a game has one or more *Nash equilibria*. These are player action selection strategies that cannot be improved upon under the assumption that the other players will keep their strategies unchanged. Strategies can either be *pure*, which always choose the same specific action, or *mixed*, which choose amongst the available actions using a fixed set of probabilities of choosing these actions.

Evolutionary game theory (EGT) studies the setting of large populations of players, who continuously over time interact in cycles of game playing then replication. During game playing, players are randomly grouped to play a given game. Replication of players (and hence their strategies) is then done in proportion to the 'fitness' of their strategies (essentially the accrued payoffs). EGT analyses seek to find evolutionary stable strategies (ESSs), which essentially are those that cannot be successfully invaded by other strategies [4].

Bach et al. [4] analysed n-person threshold games using the methods of evolutionary game theory. For the case when all players receive the same award when the cooperation threshold is reached, they found that depending on the relative payoffs for different outcomes, there is either a single pure ESS (*always defect*) or an ESS that is a mixed strategy as well as an unstable mixed Nash equilibrium.

Assuming the second case holds, this suggests that we should expect fishermen to follow a mixed strategy when a fishing day is announced. However, this does not seem to depict real social behaviour where some individuals may always defect or follow a mixed strategy, but others appear to become committed to the collective action and always cooperate. We therefore investigate how this state of cooperation could be explained by social understanding of the expectations created by joining fishing teams.

Our setting differs from an n-person threshold game, in that we do not assume that defectors receive a reward from the group activity. Rather, in line with a two-person stag hunt game, we assume the fishermen have a choice between fishing alone (for a small reward) or cooperating by participating in the fleet.

We have modelled this scenario using CASP. In our simulation, there are villager agents who have 'energy', which is decremented each 'tick' of the simulation—but they can choose to perform actions that increase their energy. The villagers do some reasoning to decide which action to perform each day (represented as a tick). At this stage we have not performed extended simulation experiments to evaluate the dynamics of populations of agents with different types of reasoning rule. Here, we demonstrate the use of expectations when choosing possible actions, and selection of an action given different personality types that are distinguished by the value they place on the social expectation rules (this can be seen as their level of confidence or belief in the accuracy of these rules).

The actions available to villagers are the following:

Fish Alone: Go fishing alone (and gain one unit of energy).

Join Fishing Team: Join the fishing team (which introduces some obligations on the agent and on others toward the agent).

Fish With Team: Show up to fish cooperatively with the team. If nobody else shows up, gain nothing, but if at least one other person shows up, gain 10 units of energy.

As the simulation begins, the agents are initialised with a random (but low) starting amount of energy. They are also randomly allotted a personality type, which will affect how they make their decisions throughout the simulation. The personality types we used were 'loners,' who are averse to joining teams; 'shirkers,' who don't mind being a part of a team, but are happy to shirk any duties that come with this membership; and 'cooperators,' who will take the existence of a social expectation on them as an overwhelming reason to act accordingly (though only if there are consequences of not doing so). We, thus, essentially have only self-interested rational agents, but show that we can achieve cooperation on collective actions based on social expectations. Note that the logic we have implemented to select actions are examples of possible personality types; the strength of our approach is that it provides the ability to model and experiment with various individual reasoning rules that take account of expectations.

On every tick of the simulation, the agents reason based on a series of internal rules to determine the actions that are relevant to the current situation, and then select and perform one of these actions. As explained in Section 4, both these decisions can be influenced by the social state as generated by the event calculus interpreter, given a set of domain-specific EC rules. Figure 2 shows three EC rules that, at the start of the simulation, initiate fluents representing conditional expectation rules. The first states that if agents join the fishing team, they will expect to never be hungry. This is a strong expectation given that defectors could cause a fishing excursion to fail, but represents the personal motivation for joining a team[6]. This is an example of a type A expectation (one addressing a physical consequence of an action). The second rule expresses the obligation taken on when joining a team: one is then expected to always fish with the team. In our agent reasoning this is considered both as a type B and a type C expectation: from any agent's viewpoint it constrains both its own behaviour and also the behaviour of other team members. The third rule expresses social knowledge about the effects of a team member defecting from a fishing expedition—if the social expectation created when joining the team is violated, then it is expected that the defector will be sanctioned. This is a type D expectation.

Initially, the agents have the following rules to suggest relevant actions:

Fish Alone Rule: If I am hungry, I (prima facie) should go fishing alone.

Join Fishing Team Rule: If (a) I am very hungry, and (b) I expect that if I join the team, I will never be hungry again, and (c) if I know that anyone who joins the team will be expected to go fishing with the team, then I (prima facie) should join the fishing team.

[6] Further extensions to the expectation event calculus reasoner could allow more complex temporal expressions to be used, e.g. a given event should occur once within every occurrence of a recurring time period.

```
initially(
    exp_rule(happ(join(Agent, fishingteam, fishingteam_fishermanrole)),
             always(not(isHungry(Agent))))
    =true).

initially(
    exp_rule(happ(join(Agent, fishingteam, fishingteam_fishermanrole)),
             always(happ(fishWithTeam(Agent))))
    =true).

initially(
    exp_rule(viol(_,_,_,always(happ(fishWithTeam(Agent))))),
             eventually(happ(sanction(Agent))))
    =true).
```

Fig. 2. Event calculus rules for the fishing domain

Once an agent has joined the fishing team (if they ever do—loners prefer to keep fishing alone), a further rule becomes available as a result of taking on the fisherman role:

Fish With Team Rule: If (a) I am expected to show up to fish with the team, and (b) there is at least one other agent who is also expected to show up to fish with the team, then I should (prima facie) show up to fish with the team.

The agent associates a valuation with each action. Loners value fishing alone higher than joining a team (as they have no interest in the latter). Cooperators value joining a team more than fishing alone. However, their valuation of fishing with a team is affected by the presence or absence of the third expectation rule in Figure 2. If the rule exists, their valuation of fishing with the team is increased to become higher than that of fishing alone. This reflects their knowledge of the social consequences of defection (we assume this rule of expectation is well founded, i.e. that sanctioning is generally performed when applicable). If the rule does not exist, then the agents become shirkers. Shirkers are members of the fishing team, but place a higher value on fishing alone than on fishing with the team, despite the social expectations that are on them.

When a simulation is run with agents who are all loners or shirkers, we see them falling into the everyone-defects Nash equilibrium, where every agent barely maintains their subsistence-level diet, but there is no collective action and so they do not thrive. However, running a simulation which includes some cooperators (agents who take seriously the social state—particularly the social expectations on them and on others—and trust that they will be enforced through a social sanction on defectors) results in the collective action of fishing as a team, and thus thriving due to the higher reward for all cooperators.

Our aim, here, is not to propose that our model of cooperators will explain all instances of collective action, but to illustrate the flexibility of our approach in modelling "non-rational" behaviour that is informed by social state, including social expectations. Indeed, the cooperative agents take the social state into account when deciding which actions are relevant and when selecting an action to perform, and this leads them to cooperative behaviour that would otherwise be considered "non-rational."

6 Conclusion

This paper has proposed that the social coordination needed to achieve collective action can arise from agents explicitly reasoning about the social expectations that arise in the problem domain. It presented an approach for investigating this proposal via agent-based simulation (ABS), using a simulation framework that extends an ABS platform with the ability to query rules of social expectations expressed in a variant of the event calculus. The framework also provides a mechanism for choosing relevant actions using decision rules associated with an agent's roles in an institution.

A group fishing simulation scenario that is a variant of an n-person threshold game was presented, and it was shown how agents' social personalities could be modelled by action valuations that can take into account the presence of social expectation rules. For one modelled personality type, the expectation that violations of expectations by members of a group would lead to sanctions was considered as a reason to value cooperation over defection.

Future work includes performing experiments with different combinations of personality types and investigating social learning mechanisms that would allow this type of socially aware personality to spread.

References

[1] Alrawagfeh, W.: Norm representation and reasoning: A formalization in event calculus. In: Boella, G., Elkind, E., Savarimuthu, B.T.R., Dignum, F., Purvis, M.K. (eds.) PRIMA 2013: Principles and Practice of Multi-Agent Systems, LNCS, vol. 8291, pp. 5–20, Springer (2013)

[2] Artikis, A., Sergot, M.: Executable specification of open multi-agent systems. Logic Journal of the IGPL **18**(1), 31–65 (2010)

[3] Artikis, A., Sergot, M.J., Paliouras, G.: An event calculus for event recognition. IEEE Transactions on Knowledge and Data Engineering **27**(4), 895–908 (2015)

[4] Bach, L., Helvik, T., Christiansen, F.: The evolution of n-player cooperation–threshold games and ESS bifurcations. Journal of Theoretical Biology **238**(2), 426–434 (2006)

[5] Castelfranchi, C.: Mind as an anticipatory device: For a theory of expectations. In: De Gregorio, M., Di Maio, V., Frucci, M., Musio, C. (eds.) Brain, Vision, and Artificial Intelligence, pp. 258–276, Springer Berlin Heidelberg (2005)

[6] Castiglione, F.: Introduction to agent based modeling and simulation. In: Encyclopedia of Complexity and Systems Science, pp. 197–200, Springer (2009)
[7] Chesani, F., Mello, P., Montali, M., Torroni, P.: Commitment tracking via the reactive event calculus. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence, pp. 91–96, Morgan Kaufmann (2009)
[8] Cranefield, S.: Agents and expectations. In: Balke, T., Dignum, F., van Riemsdijk, M.B., Chopra, A.K. (eds.) Coordination, Organizations, Institutions, and Norms in Agent Systems IX, Lecture Notes in Computer Science, vol. 8386, pp. 234–258, Springer (2014)
[9] Erickson, T.: Chapter 4: Social computing. In: Soegaard, M., Dam, R.F. (eds.) The Encyclopedia of Human-Computer Interaction, https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/social-computing, 2nd edn. (nd)
[10] Fehr, E., Fischbacher, U.: Why social preferences matter: The impact of non-selfish motives on competition. The Economic Journal **112**, C1–C33 (2002)
[11] Hardin, G.: The tragedy of the commons. Science **162**(3859), 1243–1248 (1968)
[12] Hardin, R.: Collective action as an agreeable n-person prisoner's dilemma. Behavioral Science **15**, 472–481 (1971)
[13] Hardin, R.: Collective Action. Johns Hopkins University Press (1982)
[14] Hardin, R.: The free rider problem. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/archives/spr2013/entries/free-rider/, Metaphysics Research Lab, Stanford University, Spring 2013 edn. (2013)
[15] Höllander, H.: A social exchange approach to voluntary cooperation. American Economic Review **80**, 1157–1167 (1990)
[16] Holzinger, K.: The problems of collective action: A new approach. MPI Collective Goods Preprint No. 2003/2, SSRN (2003), doi:10.2139/ssrn.399140
[17] Kowalski, R., Sergot, M.: A logic-based calculus of events. New Generation Computing **4**, 67–95 (1986)
[18] Leyton-Brown, K., Shoham, Y.: Essentials of Game Theory: A Concise, Multidisciplinary Introduction. Morgan and Claypool Publishers (2008)
[19] Malinowski, B.: Crime and custom in savage society. K. Paul, Trench, Trubner & Co., Ltd. (1926)
[20] Mueller, E.T.: Commonsense Reasoning. Morgan Kaufmann (2006)
[21] Nikolai, C., Madey, G.: Tools of the trade: A survey of various agent based modeling platforms. Journal of Artificial Societies and Social Simulation **12**(2), article 2 (2009), ISSN 1460-7425, URL http://jasss.soc.surrey.ac.uk/12/2/2.html
[22] North, M.J., Collier, N.T., Ozik, J., Tatara, E.R., Macal, C.M., Bragen, M.J., Sydelko, P.: Complex adaptive systems modeling with Repast Simphony. Complex Adaptive Systems Modeling **1**, 3 (2013)

[23] Olson, M.: The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard University Press (1965)

[24] Ostrom, E.: Governing the Commons. Cambridge University Press (1990)

[25] Ostrom, E., Ahn, T.K.: Foundations of Social Capital. Edward Elgar Publishing (2003)

[26] Pacheco, J.M., Santos, F.C., Souza, M.O., Skyrms, B.: Evolutionary dynamics of collective action in N-person stag hunt dilemmas. Proceedings of the Royal Society of London B: Biological Sciences **276**(1655), 315–321 (2009)

[27] Papadimitriou, C., Roughgarden, T.: Computing equilibria in multi-player games. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 82–91, Society for Industrial and Applied Mathematics (2005)

[28] Petruzzi, P.E., Busquets, D., Pitt, J.: Social capital as a complexity reduction mechanism for decision making in large scale open systems. In: 2014 IEEE 8th International Conference on Self-Adaptive and Self-Organizing Systems, pp. 145–150, IEEE (2014)

[29] Petruzzi, P.E., Busquets, D., Pitt, J.: A generic social capital framework for optimising self-organised collective action. In: 2015 IEEE 9th International Conference on Self-Adaptive and Self-Organizing Systems, pp. 21–30, IEEE (2015), ISSN 1949-3673, https://doi.org/10.1109/SASO.2015.10

[30] Petruzzi, P.E., Pitt, J., Busquets, D.: Inter-institutional social capital for self-organising 'nested enterprises'. In: 2016 IEEE 10th International Conference on Self-Adaptive and Self-Organizing Systems, pp. 90–99, IEEE (2016)

[31] Pitt, J., Nowak, A.: The reinvention of social capital for socio-technical systems. IEEE Technology & Society Magazine **33**(1), 27–33 (2014)

[32] Pitt, J., Schaumeier, J., Artikis, A.: Axiomatization of socio-economic principles for self-organizing institutions: Concepts, experiments and challenges. ACM Transactions on Autonomous Adaptive Systems **7**(4), 39:1–39:39 (2012)

[33] Reuben, E.: The Evolution of Theories of Collective Action. MPhil thesis, Tinbergen Institute (2003)

[34] Schuler, D.: Social computing. Communications of the ACM **37**(1), 28–29 (1994)

[35] Shanahan, M.: The event calculus explained. In: Artificial Intelligence Today, Lecture Notes in Computer Science, vol. 1600, pp. 409–430, Springer (1999)

[36] Wielemaker, J., Schrijvers, T., Triska, M., Lager, T.: SWI-Prolog. Theory and Practice of Logic Programming **12**(1-2), 67–96 (2012), ISSN 1471-0684

[37] Yolum, P., Singh, M.: Reasoning about commitments in the event calculus: An approach for specifying and executing protocols. Annals of Mathematics and Artificial Intelligence **42**, 227–253 (2004)

[38] Young, H.P.: Social norms. In: Durlauf, S., Blume, L. (eds.) The New Palgrave Dictionary of Economics, Palgrave Macmillan, 2nd edn. (2008)

Constructing an Agent Taxonomy from a Simulation through Topological Data Analysis

Samarth Swarup and Reza Rezazadegan

University of Virginia,
Charlottesville, VA, USA
{swarup, rr7nz}@virginia.edu

Abstract. We investigate the use of topological data analysis (TDA) for automatically generating an agent taxonomy from the results of a multiagent simulation. This helps to simplify the results of a complex multiagent simulation and make it comprehensible in terms of the large-scale structure and emergent behavior induced by the dynamics of interaction in the simulation. We first do a toy evacuation simulation and show how TDA can be extended to apply to trajectory data. The results show that the extracted types of agents conform to the designed agent behavior and also to emergent structure due to agent interactions. We then apply the method to a sample of data from a large-scale disaster simulation and demonstrate the existence of multiple emergent types of agents.

Keywords: topological data analysis · simulation analytics · agent taxonomy.

1 Introduction

A common question that is raised, when a multiagent simulation is presented, is, "Do you have different types of agents?" Generally the intent of the question is with respect to the design of the simulation, i.e., whether the simulation has different types of agents by design. An example might be a disaster simulation that has civilians and emergency responders, or adults and children, etc. A typology by design helps to understand the structure of the simulation, since different types of agents might have different behaviors, which result in different types of trajectories through the state space of the simulation, and ultimately manifest in different outcomes.

However, the same question can be asked with respect to an analysis of the outputs of the simulation. In this case, the intent of the question is with respect to emergent behavior in the simulation. While there is still considerable debate about the definition of emergence, here we simply mean differences in agent behaviors that are not explicitly designed into the simulation, but are induced by the dynamics of interaction within the simulation.

Constructing a typology of agents from the outputs of a simulation may, in a sense, be more instructive because *(1)* agents that are different by design may not exhibit significant differences in behavior during the actual running of the simulation, and *(2)* agents that are not different by design might still exhibit significant differences induced by the dynamics of interaction within the simulation. Thus the emergent typology offers insight into how the interaction dynamics drive the simulation to exhibit particular emergent outcomes, which can be more complex and subtle than the design of the simulation might suggest.

Our goal here is to devise a method for generating a *taxonomy* of agents from the results of a simulation. A taxonomy goes beyond a typology in that it not only identifies meaningful types from a data set, but also establishes relationships among those types. For example, a taxonomy of biological organisms generally groups them into "taxa" by shared morphological characteristics. It can also create a ranking by grouping the taxa, like a hierarchical clustering method.

Generating a taxonomy of agents in a multiagent simulation is useful not just for understanding the emergent structure of the simulation. It is also a very useful way to present the simulation to end-users. For example, operational end-users who actually have to implement response plans during a disaster recognize the existence of emergent roles and behavior [10, 9], and would benefit greatly from this type of information. This would, in turn, allow progress towards using simulations in a prescriptive way [5], i.e., to use simulations to suggest operationalizable courses of action and response plans.

The rest of this paper is organized as follows. We begin by describing a toy simulation of an evacuation scenario. We analyze output agent spatial trajectories from this simulation and show that clustering alone is not sufficient to extract the different types of trajectories from the data. After that we describe the topological data analysis method with a simple example and show how we can extend it to trajectory analysis. We then apply TDA to the evacuation simulation data and show the resulting taxonomy of agents. To assess the method on a more complex data set, we use sample trajectory data that we obtained from a recent disaster simulation [8]. The result of applying TDA to this data set is considerably more complicated, but we show that a set of emergent categories of agents can still be extracted from the results. We end with a discussion of the method and possible extensions.

2 Evacuation Simulation

We created a simple simulation of an evacuation scenario, where we have a population moving over a road network and trying to reach some marked "exit" nodes. This is not meant to be a realistic evacuation simulation. It is a toy test-bed where we can design simple interactions between agents and observe their effects on the resulting agent spatio-temporal trajectories. This will help us evaluate the effectiveness of the TDA method for generating a taxonomy.

The main components of the simulation are 1. A population of agents, 2. A road network, and 3. A behavior model. We describe each of these next, as well as the format of the resulting outputs.

Population: The agent population is organized into groups of different sizes, from 1 to 4. Groups of size 1 are referred to as individuals, and the rest are referred to as group agents. Each agent is assigned an *age* and *gender*, though only the age is relevant to behavior, as we describe further below. Groups correspond to families and are assigned appropriate ages and genders. In particular, children (i.e., agents with age less than 18) are always group agents. A distribution over group sizes governs the relative numbers of individuals and group agents that are generated. In the experiments in this paper, we generated a population of 100 agents, of whom 50 were individuals and 50 were group agents. The latter were divided into 10 groups of size 2, 6 groups of size 3, and 3 groups of size 4.

Road Network: The evacuation is assumed to be taking place over a road network. We model this as a graph embedded in two dimensions. We construct the graph by generating a collection of random points in a square area and connecting each point to its k nearest neighbors. The points correspond to the nodes in the network and, thus, each node has a corresponding (x, y) location. The road network used in the simulations presented in the following sections is shown in Figure 1. Two nodes were randomly selected as exit nodes. These are marked in green in Figure 1. In our simple model, agents are assumed to be at the nodes (corresponding to intersections), and to move exactly one hop in a time step (if they choose to move at all). In the simulations used in this paper, we generated a road network with 100 nodes, where each node is connected to its four nearest neighbors.

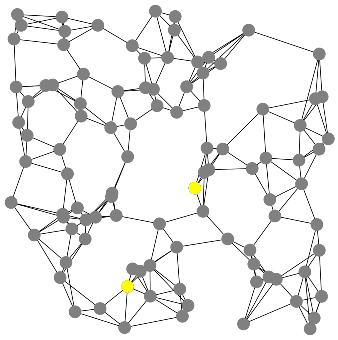


Fig. 1: Road network for the evacuation simulation. The exit nodes are marked in yellow.

Behavior: We implement five different behaviors:

- **Evacuation**: This is the behavior where the agent is heading towards the closest exit node. This is implemented efficiently by using precomputed shortest paths from all nodes to the closest exit node. Individual agents always execute this behavior until they reach the exit node.
- **Rendezvous**: In this behavior, agents move towards their nearest group member. Once all the group members are at the same node, this behavior ends, and the entire group switches to the evacuation behavior.
- **Stay**: In this behavior, agents stay at their current node and do not move. This behavior is executed by all child agents until at least one of the adults from their group arrives at the same node. Thereafter child agents switch to rendezvous or evacuation, matching the collocated group members.
- **Exited**: Once an individual agent reaches an exit node, it has exited and so continues to be at that node for the rest of the simulation. Group agents switch to the exited behavior only if the entire group is together when they

reach the exit node. Otherwise, in the rendezvous behavior, they can pass through an exit node without switching to the exited behavior.
- **Do-nothing**: With a small probability, an agent in the evacuation or rendezvous behavior stays at its current node for one time step.

Simulation Outputs: Initially the agents are randomly distributed over the road network nodes. The simulation is then run for 20 time steps, which is sufficient for ∼85 of the agents to reach the exit nodes. We output the spatio-temporal trajectories of all the agents as series of (x, y, t) tuples, where the (x, y) coordinates are the coordinates of the road network node the agent is at, at time step t, for $t \in [0, 20]$.

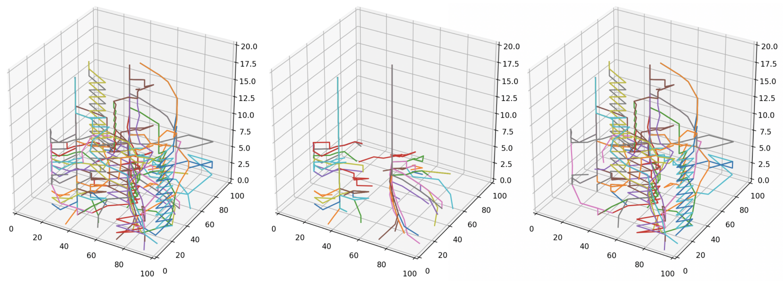


Fig. 2: Agent trajectories. The left panel shows the trajectories of all agents. When we split them into individuals and groups, we see that the trajectories of individuals (middle panel) look quite different from the trajectories of groups (right panel). The x and y axes show spatial locations. The z-axis is time. The colors are arbitrary.

The resulting trajectories are shown in Figure 2. Though all the trajectories taken together are hard to parse visually, there is a hidden structure or typology induced by the dynamics, which is made clear when we separate out the individual and group agents.

The middle panel shows just the individual agents. Since these agents only do the evacuation and exited behaviors, these trajectories are qualitatively simple. They correspond to the shortest paths from the initial nodes of the agents to the closest exit nodes. Essentially, there are two types of agents here: those that go to one exit node, and those that go to the other exit node. This difference is entirely due to their initial location.

The right panel shows just the group agents, who have significantly more complicated trajectories. The rendezvous behavior results in trajectories that go away from the closest exit node, trajectories that have loops, and trajectories that show oscillations between two adjacent nodes. These last are due to group members ending up at adjacent nodes and then each trying to move to the other group members' location at each time step. The do-nothing behavior helps to escape this trap over time, but we see that in a couple of cases, the simulation hasn't run for long enough for the agents to stop oscillating.

3 Analysis by clustering

We first analyze the set of trajectories by clustering, as follows. Paths in a graph G can be regarded as points in the high dimensional space of all the paths in G. As such a path is equivalent to a vector of adjacent vertices in the graph. To have a measure of similarity between paths, one can extend the graph distance into a distance between paths by using any of the L^p norms. This means if c, c' are two paths then we set $D_p(c, c') = \left(\sum d(c(i), c'(i))^p\right)^{1/p}$, where $c(i)$ refers to the i^{th} element of c, and correspondingly for c'. The metric d can be the graph distance or, if the graph is embedded in Euclidean space, the Euclidean distance. Of particular interest are the cases $p = 2, \infty$. The latter gives us the maximum distance between the corresponding vertices of the graph.

To cluster the trajectories shown in Figure 2, we used complete linkage clustering with the number of clusters set to either 2 or 4. In this method a dendrogram is obtained by recursively merging clusters of closest distance. The distance between two clusters is given by the maximum of distances between their elements. At the beginning each data point is a cluster of its own and at the end of the process all points are merged into one cluster. Given the desired number n of clusters, one uses the last level where there were n clusters in the dendrogram [6].

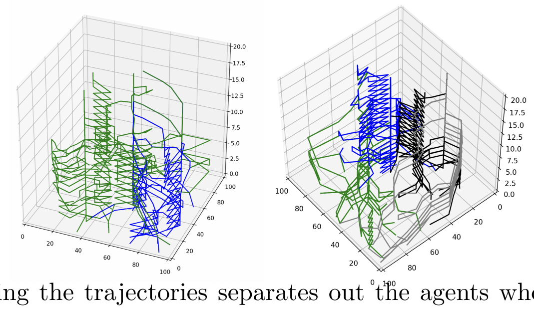


Fig. 3: Clustering the trajectories separates out the agents who go to one exit node vs. the other (green vs. blue), but does not separate the individual and group trajectories.

The results of clustering are shown in Figure 3. The left panel shows the results with number of clusters set to 2, and the right panel shows the results with the number of clusters set to 4. We see, in each case, that the nearby trajectories are grouped together, without distinction as to the structure or "complexity" of the trajectories. This is as expected, of course, but it demonstrates the inability of simple clustering to extract the real structure in the data, which is the distinction between the simple trajectories of individual agents and the complex trajectories of group agents. Thus, a taxonomy based on simple clustering would not give a meaningful set of categories of agents. Intuitively, the property we are trying to extract is captured by the shape of the trajectories, which suggests that a topological method might be better suited. So, we now turn to topological data analysis as a possible route to constructing a taxonomy. We first introduce TDA with a simple example, and then apply it to our simulation.

4 Topological Data Analysis (TDA)

Topology is the study of spaces equipped with a notion of neighborhood between their elements. Metric spaces (in particular graphs) are a particular example of topological spaces, though in general we do not need a metric to know which elements are neighbors. Other examples of topological spaces include simplicial complexes which are hypergraphs in which any subset of a hyper-edge is itself a hyper-edge. The dimension of a simplicial complex is the size of its largest hyper-edge minus one. In particular a graph is a simplicial complex of dimension one.

TDA aims to find a hypothetical topological space to which a given data set belongs. For example, we can take the proximity graph of a data set i.e., the graph obtained from the data set by connecting pairs of points whose distance is less than a given threshold. Once a topological space is associated to the data set, one can apply various topological methods and invariants to study the data and extract its inherent characteristics. As topology is the study of properties invariant under continuous transformations, TDA can be thought of as studying the properties of data which are robust w.r.t. continuous deformations of data. One prominent example of such an invariant is persistent homology [1] which has been applied to studying data in various different fields. However a precise description of persistent homology is beyond the scope of this paper.

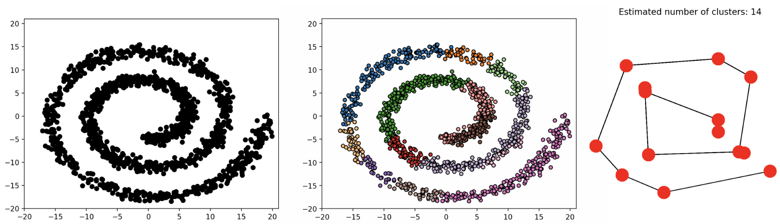


Fig. 4: Example of topological data analysis.

In a recent example, Lum et al. [4] use an enhanced method of clustering to assign a graph to the data. The data set is equipped with a filtration function and one uses this function to divide the dataset into a set of overlapping bins. One then clusters the data in each bin. Each such cluster gives us a vertex of the output graph and two such vertices are connected by an edge if their corresponding clusters have elements in common. This method is then applied to several data sets such as gene expression from breast tumors, voting data from the United States House of Representatives and player performance data from the NBA to obtain new insight on associations in among data points. In each case the authors find stratifications of the data which are more refined than those produced by traditional methods.

Figure 4 shows an example of TDA applied to a data set (left panel) which contains points along a spiral manifold. We used principal components analysis (PCA) to choose a direction for projection to a single dimension. The points were binned into five overlapping bins along this axis and the points in each bin were clustered in the original 2D space using DBSCAN [2]. The results of the clustering are shown in the middle panel. Each cluster was then replaced by a graph node placed at the cluster centroid (using Manhattan distance) and connected to neighboring nodes if the corresponding clusters shared any common points. This resulted in the graph structure shown in the panel on the right, which captures the essential spiral structure of the original data set.

Though it uses clustering, TDA is solving a fundamentally different problem. The spiral manifold data set doesn't have meaningful clusters. TDA is capturing the essential shape of the data set as a graph. The structure of the graph shows the linear structure of the data manifold. In addition, since the graph is embedded in two dimensions (each node has an $(x, y$ location derived from the data), it also captures the spiral structure of the manifold.

5 Analysis by TDA

To apply TDA to the study of agent trajectories, instead of using a filtration function as in [4], we can restrict the paths to possibly overlapping temporal regions. In this case we divide the runtime interval $[0, N]$ into intervals $[0, k], [k + 1, 2k], [2k + 1, 3k], \ldots$ and then cluster the restrictions of the paths to each subinterval as before. This way we obtain sets of clusters $C_0, C_1, \ldots C_n$ where $n = N/k$. We then connect with an edge the clusters in C_i and C_{i+1} that contain the restrictions of the same path.

An important special case is when $k = 0$. In this case the above procedure is equivalent to clustering the positions of the agents at each time step i and then connecting any cluster in C_i to those in C_{i+1} which contain the position of the same agent. This can be though of as a coarsening of the trajectories.

In the other extreme, i.e., when $k = N$ we obtain the clustering of section 3 back. Therefore we can regard our adaptation of TDA as a parametrized clustering method for agents.

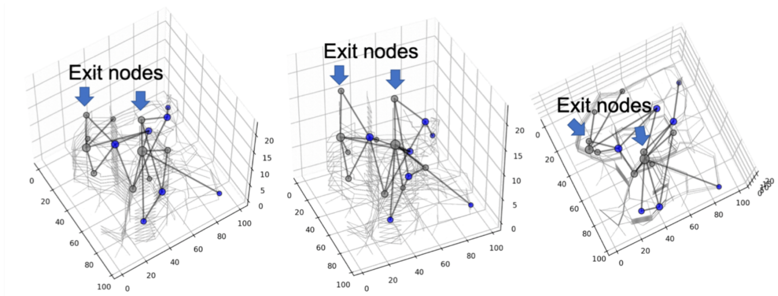


Fig. 5: Topological data analysis of agent trajectories. The three panels show three different views of the same graph, to help with understanding its 3D shape. The actual trajectories are shown in thin gray lines. We see that TDA is able to separate out several of the group trajectories, especially ones where the groups don't reach the exit.

Figure 5 shows the result of applying TDA to the agent trajectories. The graph has a somewhat complicated structure, so three different views are shown. The graph nodes corresponding to the exit nodes of the road network are marked. If all the trajectories assigned to a node are group agent trajectories, then the node is colored blue, otherwise it is colored gray.

Taxonomy We see that TDA is able to separate out several group trajectories, especially ones that don't reach the exit nodes. There are just two clusters (graph nodes) at the top level, five at the middle level, and ten at the bottom level. This gives us a nice taxonomy of the agents. Broadly, there are the two groups of agents that reach the two exit nodes. Even the agents that don't end up reaching one of the exit nodes are mapped to the closest exit node. Following the edges from the two top-level nodes gives us the five mid-level "taxa". Here the agents that don't reach the exit nodes are split off into their own categories. The lowest level taxa correspond to the early part of the simulation, and are therefore reflective of the starting locations of the agents.

We will now turn to a much more complex disaster simulation. We describe the simulation and the data set briefly first, and then present results from applying TDA.

6 Analysis of a disaster simulation

We obtained a sample of agent trajectories from a recent disaster simulation [8]. In this section, we briefly describe the simulation before going on to show the results of our method applied to the data set.

The scenario of the simulation is that an improvised nuclear device is detonated in Washington DC, USA. This hypothetical disaster is known as National Planning Scenario 1 (NPS-1) and has been studied extensively for many years We will refer to the simulation [8] as the NPS-1 simulation.

In the NPS-1 simulation, they modeled a detailed "synthetic" population of the region, including agent demographics, household structure, daily activity patterns, road networks, and various kinds of locations, such as workplaces, schools, government buildings, etc. This was a highly data-driven simulation, using data from multiple sources, such as the American Community Survey, the National Household Travel Survey, Navteq (road network data), Dun & Bradstreet (business location data), and more.

The simulation also contained models of multiple infrastructures, including power, communication, transportation, and health. Damage to these infrastructures affects the behavior and mobility of agents in the simulation in multiple ways. For instance, cell towers are inoperative close to ground zero, which means that people can't get in touch with family members, can't make 911 calls, and can't receive emergency broadcasts advising them to shelter in place. This lack of information affects agents' behavioral choices.

Similarly, damage to roads, as well as injuries and radiation sickness, prevent or limit agent mobility. Slow movement through areas close to ground zero also increases radiation exposure and exacerbates loss of health.

The simulation modeled six behaviors [8], as mentioned in Section 1. The behaviors were *household reconstitution*, *shelter-seeking*, *worry*, *evacuation*, *healthcare-seeking*, and *aiding & assisting*. These behaviors were implemented as specific policies, specified as short programs, over an action space that contained just two actions: moving (towards a destination) and calling (a family member, 911, etc).

The simulation was run for 100 time steps. The first six time steps corresponded to 10 minute intervals of real-time each, and the next 94 to 30 minutes of real-time each, giving a total of 48 hours. We obtained a sample of 10,000 agents, out of a total of 730,833 agents modeled in the simulation. The variables included in the data set are *distance from ground zero* in meters, *level of radiation exposure* in centiGrays, *health state*, which is an integer in the range $[0, 7]$, and *behavior*, which is nominal, indicating which of the above behaviors an agent is executing at each time step.

6.1 Results

To enable viewing the results in a 3D plot as before, we restrict our analysis to pairs of variables (plus time). We also limit our TDA graph construction to time step 20 because we found that agent states don't change very much after that. To run TDA for the full sample of 10,000 agents takes a few hours (on a MacBook Pro with 2.6 GHz Intel Core i7 and 16GB RAM), and results in a plot that is too cluttered to understand easily. Therefore, we demonstrate results with a random sample of 100 agents. We tried the analysis with multiple random samples of 100 agents, and the results are qualitatively similar each time.

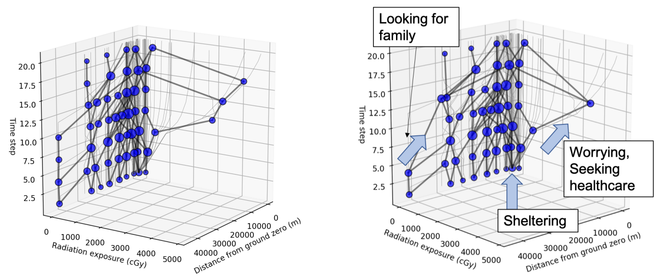


Fig. 6: Topological data analysis of 100 randomly chosen agent trajectories in the disaster simulation, where the variables are distance from ground zero and level of radiation exposure. The left panel shows the result of TDA, while the right panel shows the graph after homeomorphic smoothing (edge contraction).

Figures 6, 7, and 8 show the results. In each case the actual trajectories are shown with thin gray lines, while the TDA graph is shown in blue. The right panel of each of the figures shows a simplified version of the graph in the left panel, generated by homeomorphic smoothing (edge contraction) [3].

The idea of homeomorphic smoothing is to simplify a graph by removing nodes of degree 2 and connecting their neighbors to each other. The graphs generated by TDA often exhibit long paths where it follows a trajectory of an agent that doesn't interact with other agents. Examples can be seen on the left and right side of the left panel in Figure 6.

For the purpose of generating a taxonomy, these intermediate nodes in paths in the TDA graph don't add any information and can be removed. Depending on the structure of the graph, the impact of homeomorphic smoothing can be small (as in Figures 6 and 7), or large (as in Figure 8). Yet, though the right panel in Figure 8 is greatly simplified compared to the left panel, it preserves the essential distinction between agents who have low radiation exposure and remain healthy vs agents whose radiation exposure increases over the course of the simulation, leading to a deterioration of their health condition.

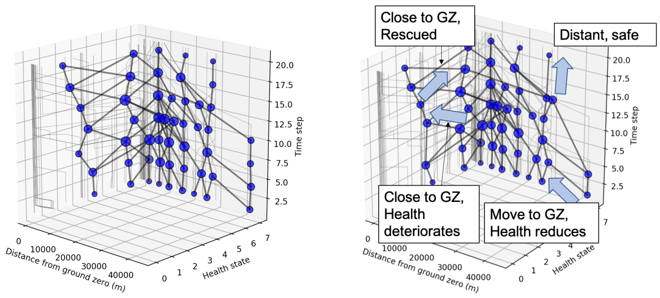


Fig. 7: Topological data analysis of 100 randomly chosen agent trajectories in the disaster simulation, where the variables are distance from ground zero and health state. The left panel shows the result of TDA, while the right panel shows the graph after homeomorphic smoothing (edge contraction). "GZ" is Ground Zero, i.e., the location where the bomb is detonated.

Taxonomy The annotations in Figures 6, 7, and 8 show some of the taxa that emerge. In this case also, we can treat the graph nodes at the end of the simulation as the top level taxa in our emergent taxonomy. Thus, for example, in Figure 7, the top level taxa correspond to agents who are

1. close to ground zero and in poor health,
2. close to ground zero and in good health,
3. at an intermediate distance from ground zero and in good health, and
4. far from ground zero and in good health.

As we follow the graph edges and move down from the top level, we can describe how the agents got to the states in the top level. These categories are annotated in the right panel of Figure 7. Similarly, we can come up with descriptive categories of agents from the graph-based taxonomy in Figures 6 and 8, as shown in the right panels of those figures. Thus the method gives us a ranked classification, i.e., a taxonomy, not just a typology of agents in the simulation. Importantly, these categories are not designed into the simulation, but emerge from the interactions induced by agent movement, communication, and behavior.

7 Discussion

The method presented here is a beginning to the solution of the problem posed in this paper. Topological data analysis, though an elegant idea, relies on clustering, which is still more of an art than a science. There is no doubt that the results presented here could be further improved through more experimentation. One possible direction for future research along these lines is to test the predictive power of the discovered taxonomy, e.g., can we use the taxa to predict the behavior the agents are engaged in? If that were to be the case, it would suggest applicability of this method beyond simulations, to predictive analysis of real-world disaster data.

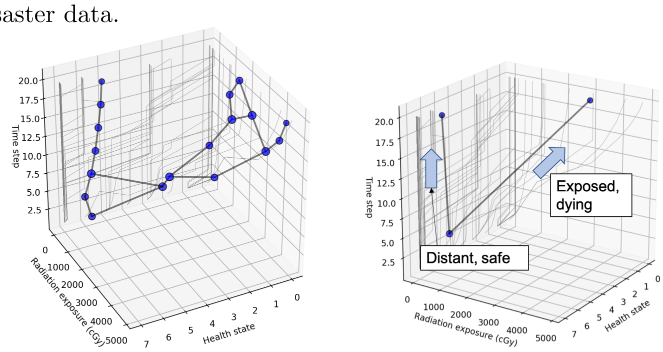


Fig. 8: Topological data analysis of 100 randomly chosen agent trajectories in the disaster simulation, where the variables are level of radiation exposure and health state. The left panel shows the result of TDA, while the right panel shows the graph after homeomorphic smoothing (edge contraction).

The method presented also generalizes to higher dimensions, though we chose to stick to two dimensions (plus time) for our experiments for ease of presentation and understanding of the results. There is a need for a more rigorous method of evaluating the results from TDA in order to be able to use it well in higher dimensions. Presently, there isn't a good method for deciding how well the TDA graph captures the topology of the underlying data set. An important direction for future research is to connect TDA to more rigorously theoretical methods in topology like persistent homology.

Our experience here also suggests that, for complex simulations, the graph resulting from TDA might itself be too complex to understand. We further simplified it using homeomorphic smoothing, but other methods could possibly be developed for that. Other generalizations that are possible are to use tensor factorization to discover good filtrations when time is not one of the variables, and to develop a method for doing TDA on graphs when the graph is not embedded in Euclidean space. While distances are still well-defined in that case (shortest path distance), filtration and binning don't have obvious analogs.

More generally, we believe this is a helpful method for the broader goal of making simulations more usable and useful. Many simulation analytics methods are being developed which address different facets of simulation use, and these methods need to be brought together into a common framework. For example, the problem of simulation summarization [7] is clearly related to the problem of generating a taxonomy of agents. A user study could also be done to assess if operational users, such as emergency responders and planners, find this taxonomy useful or interesting. This would help improve the simulation as well as build trust in the methods on the part of the end-users, which is ultimately the biggest barrier to mainstream adoption of MAS methods.

Acknowledgments S.S. was supported in part by DTRA CNIMS Contract HDTRA1-17-0118.

References

1. Carlsson, G.: Topology and data. Bull. Amer. Math. Soc. **46**, 255–308 (2009)
2. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In: Proc. KDD. pp. 226–231. AAAI Press (1996)
3. Gross, J.L., Yellen, J.: Graph Theory and Its Applications, Second Edition. CRC Press (1998), pp. 263
4. Lum, P.Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., Carlsson, J., Carlsson, G.: Extracting insights from the shape of complex data using topology. Scientific Reports **3**, Article 1236 (2013)
5. Marathe, M., Mortveit, H., Parikh, N., Swarup, S.: Prescriptive analytics using synthetic information. In: Hsu, W.H. (ed.) Emerging Trends in Predictive Analytics: Risk Management and Decision Making, pp. 1–19. IGI Global, Hershey, PA (2014)
6. Müllner, D.: fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. Journal of Statistical Software **53**(9), 1–18 (2013)
7. Parikh, N., Marathe, M.V., Swarup, S.: Summarizing simulation results using causally-relevant states. In: Osman, N., Sierra, C. (eds.) Autonomous Agents and Multiagent Systems: AAMAS 2016 Workshops, Visionary Papers, Singapore, Singapore, May 9-10, 2016, Revised Selected Papers, pp. 88–103. Springer International Publishing (2016)
8. Parikh, N., Swarup, S., Stretz, P.E., Rivers, C.M., Lewis, B.L., Marathe, M.V., Eubank, S.G., Barrett, C.L., Lum, K., Chungbaek, Y.: Modeling human behavior in the aftermath of a hypothetical improvised nuclear detonation. In: Proc. AAMAS. Saint Paul, MN, USA (May 2013)
9. Provitolo, D., Dubos-Paillard, E., Müller, J.P.: Emergent human behaviour during a disaster: Thematic *vs* complex systems approaches. In: Proceedings of EPNACS 2011 within ECCS'11. Vienna, Austria (September 15 2011)
10. Stallings, R.A., Quarantelli, E.L.: Emergent citizen groups and emergency management. Public Administration Review **45**, 93–100 (1985), special Issue: Emergency Management: A Challenge for Public Administration

On Developing A More Comprehensive Decision-Making Architecture for Empirical Social Research: Lesson from Agent-Based Simulation of Mobility Demands in Switzerland

Khoa Nguyen[1] and René Schumann[1]

HES-SO Valais Wallis, SiLab, Rue de Technpole 3, 3960 Sierre, Switzerland
{khoa.nguyen,rene.schumann}@hevs.ch
http://silab.hevs.ch

Abstract. Agent-based simulation is an alternative approach to traditional analytical methods for understanding and capturing different types of complex, dynamic interactive processes. However, the application of these models is currently not common in the field of socio-economical science and many researchers still consider them as intransparent, unreliable and unsuitable for prediction. One of the main reasons is that these models are often built on architectures derived from computational concepts, and hence do not speak to the selected domain's ontologies. Using Triandis' Theory of Interpersonal Behaviour, we are developing a new agent architecture for choice model simulation that capable of combining a diverse number of determinants in human decision-making and being enhanced by empirical data. It also aims to promote communication between technical scientists and other disciplines in a collaborative environment. This paper illustrates an overview of this architecture and its implementation in creating an agent population for the simulation of mobility demand in Switzerland.

Keywords: Agent architecture · Multi-agent system · Agent-based modelling · Discrete choice analysis.

1 Introduction

The use of a specific architecture can facilitate the application of agent-based methodology in a particular domain. Traditionally, economists tend to give importance to the selfish and rational part (*homo economicus*), while sociologists focus on the social capabilities (*Aristotle's zoon politikon*) and psychologists tend to see humans as mainly irrational and emotional. Thus, explicitly or not, agent-based models often follow one or another of these perspectives (e.g [4, 9, 11]).

In recent years, we observe a trend of applying agent-based techniques to combine the views from different domains to provide more reliable descriptions for real-world phenomena [23] (e.g. self-organisation, the emergence of counter-intuitive behaviours [13]). This leads to the search for a generic computational platform that has a higher degree of abstract, while can also be adapted as an illustration of a specific theory or hypothesis [7]. There is still, however, a lack of decision-making architecture that is expressive and flexible enough to build arguments both micro-macro levels in the socio-economical context [3, 30].

This paper introduces an agent architecture for choice modelling simulation, which is inspired by Triandis' Theory of Interpersonal Behaviour (TIB) [34]. TIB states that behaviour is primarily a function of the intention to engage in the act, habit and facilitating conditions. It provides a meaningful set of determinants that contribute to decision-making in socio-psychology and can be used to produce statements about behaviours at society level as well as its individual members. In addition, the function given in TIB allows us to calculate the probability that a particular action will take place. By enhancing it with statistical data, this architecture can enable an agent-based model to have not only theoretical support from an established concept but also the capability to include empirical findings in scenario design. We demonstrate the implementation of this architecture in BedDeM (i.e. Behaviour-Driven Demand Model) - a simulation tool that aims to address both micro and macro perspectives of modal choice for mobility domain in Switzerland.

After considering some of the popular strategies for decision-making simulation in Section 2, a specification of the new architecture is presented (Section 3). Next, its contextualisation in the studied problem, Behavioural-driven Demand Model (BedDeM), is carried out in Section 4, especially focusing on the attribute definition, micro-behaviour and calibration. We then conclude our experience with the whole process and suggest further development in Section 5.

2 Related works

For models that aim to understand the aggregate consequences of real-world phenomena, it is important to specify an agent's behaviours in a way that is both theoretically and empirically defensible [12]. There are different approaches for this issue in choice modelling, ranging from as basic as a reactive mechanism to the level of a complex entity using a cognitive model.

A simple design involves agents follow some sets of behaviour rules (i.e. decision-tree or production-rule systems), which apply both in information-gathering stage and when making a final choice. It is typically used in conjunction with a set of assumed preferences for the agent to rank outcomes by desirability order. Examples include heuristics that update agent's behaviours according to the accumulated experience (e.g. [33]) or pick the next option that satisfies the qualities identified from empirical data analysis (e.g. [16]). In this setup, modellers have a straightforward job to trackback any changes in agents' behaviour but have to face a significant increase in computational complexity when a new rule is introduced [22].

Alternatively, researchers can choose to assign agents with beliefs, values or world views that correspond to observation from ethnographic data or stakeholder's assessment. A range of cognitive inspired agent architectures has been developed in recent years for this purpose. Mostly supported by process-based theories [30] and a bounded rationality approach [27], they aim for providing a framework for a psychological mechanism through specifying essential structures, divisions of modules and their relations while always embodying fundamental theoretical assumptions [29]. One of the most well-known architecture is Belief-Desires-Intentions (BDI) [20]. It provides a robust standard framework for any agent-based simulation that wants to take into account human's decision-making process. However, these methods are often criticised for the lack of experimental grounding [6] and the agent choice of being homogeneous, completely rational and selfish [20].

Taking into account the dual nature of social processes, working on individual and societal levels requires the consideration of both and the interaction dynamics among them [8]. Thus, other cognitive models that add complexity to the classical rational agent, have emerged. Representatives for this category are CLARION [28], ACT-R [32], SOAR [17] etc. They usually take into account social theories and focus on different issues that were ignored in the rational agent. For example, Conte et al. [5] empower the social learning capabilities or Sun et al. [31] focus on organisational theories and the agent roles while others stress on the importance of beliefs in cognition [25]. There have been attempts in finding a global unifying principles for cognitive architecture (e.g. [6]), but it still remains an open debate [29,30]. Balke et al. [3] make a comparison between their features, which reveals none of the mentioned models is currently cover all socio-psychological aspects of decision-making (i.e. cognitive, affective, social, norm and learning).

Another popular approach is to enhance the agent's preferences, strategies and likelihood of making a particular decision with discrete choice models (e.g. [14]). Giving some defined set of possible options, it specifies a ranking order of these choice outcomes, which can then be converted into predicted probabilities. To produce an actual choice, a random component (representing human-error) can be introduced by sampling from a multinomial distribution with these probabilities. Alternatively, one can assume the computed value reflect the underlying desire of the agent and specify it to always pick the option with the highest utility value. By incorporating empirical data (such as observed choices, survey responses to hypothetical scenarios or administrative records), the discrete choice model provides one flexible framework for estimating the parameter of choice behaviour, especially when there is a lack of information on which determinants affecting individual choice decisions. Despite that, without comprehensive support from a socio-psychological theory, current discrete choice models are often difficult for non-experts to understand the underlying implications of different modelling scenarios and associated behavioural assumptions [15].

3 New architecture design

As an effort to produce a more comprehensive agent architecture for empirical researches, we decide to implement Triandis' Theory of Interpersonal Behaviour (TIB) [34] (Fig 1). The first level is concerned with the way personal characteristics and prior experiences shape personal attitudes, beliefs and social determinants related to the behaviour. The second level explains how cognition, affect and social determinants and personal normative beliefs influence the formation of intentions with regards to a specific behaviour. Finally, the third level states that intentions regarding the behaviour, prior experience and situational conditions predict whether or not the person will perform the behaviour in question.

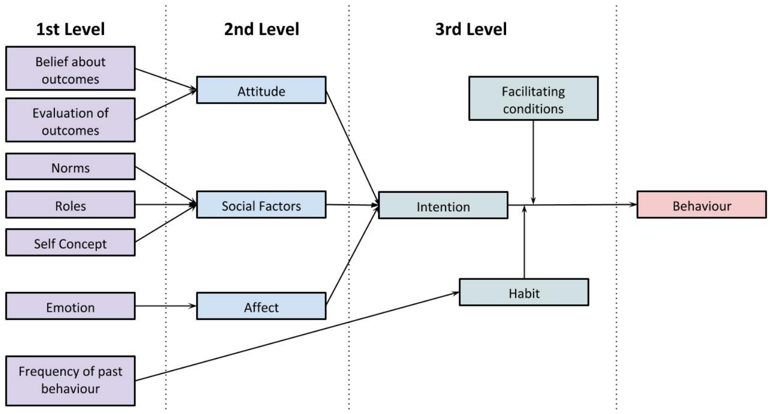


Fig. 1: Triandis' tri-level model [34]

A full decision-making cycle with an example of a mobility application is illustrated in Fig. 2. An agent first selects an isolated decision-making task from the list that is sequentially executed. Its personal desire/goal is then combined with means provided by the external environment to generate a set of possible options. For all determinants (d), each option (opt) is given a referenced value which comes from comparing its property with other's ($R_d(opt)$). In the first level, this can be done using either a real numerical system (for determinants such as price or time) or ranking function (for determinants such as emotion). Both can be derived from empirical data (e.g. census/survey) or calibrated with expert's knowledge/stakeholder's assessment.

The results for these determinants are then normalised and multiplied with an associated weight (called w_d); the sum of which becomes the referenced value for the option in the next level (see Eq.1). The weight, in this case, represents the importance of a decision-making determinant compare to others at the same level and emphasises on the heterogeneity of individuals. It also allows the modeller to express a certain theory by cutting of determinants (by setting their values to

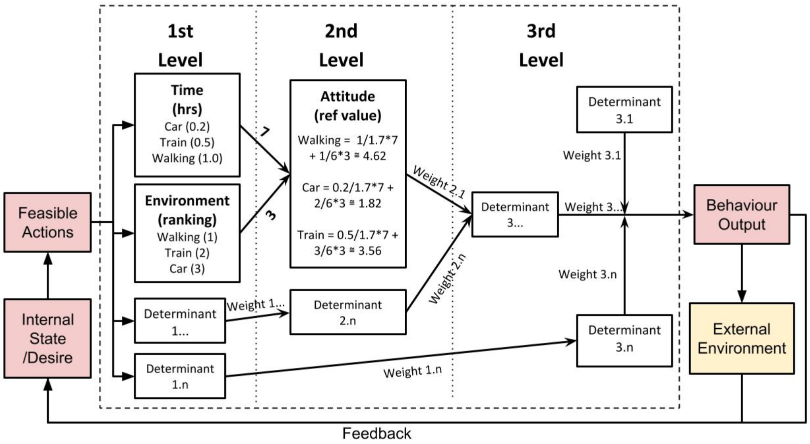


Fig. 2: Agent's decision-making procedure

0) that are not relevant to a case study. The combination process then continues until it reaches the behaviour output list; the referenced value of which can be interpreted as the probabilities that an agent will perform that option. If the agent is assumed to be deterministic, it can pick the option that is correlated to best-evaluated value.

$$R_d(opt) = \sum_{c=1}^{C}(R_c(opt)/(\sum_{o=1}^{O} R_c(o)) * w_c)$$

where
- $R_d(opt)$ is the reference value of an option (opt) at determinant d.
- C is the set of the children of d (i.e. determinants connects with d in the previous level).
- O is the set of all available options.
- w_c is the weight of child determinant c.

(1)

In our mobility example (see Fig. 2), the agent has access to 3 options: *walking*, using *car* or taking *train*. For a working trip of around 10 kilometres distance, according to *time*, their referenced values are: R_{time} =car(0.2), train(0.5), walking(1.0) (measured in hours); which combine to 1.7. According to *environmental friendly* determinant, they can be ranked as $R_{environment}$ = walking(1), train(2), car(3) (from best to worst); the sum of which is 6. If w_{time} and $w_{environment}$ are 7 and 3 respectively, the new referenced value in next level list($R_{attitude}$) of *walking* would be 1/1.7*7 + 1/6*3 $\approx$ 4.62, *car* would be 0.2/1.7*7+2/6*3 $\approx$ 1.82 and *train* would be 0.5/1.7*7+3/6*3 $\approx$ 3.56. Hence, according to *attitude*, *car* would have the highest chance to be picked for this individual agent, followed by *train* and *walking*.

4 A case study - BedDeM [1]

BedDeM is being developed in Java using Repast library for agent-based modelling [21], aiming to generate yearly mobility data at the individual household level that can be interpreted at the granularity of a historical *evolution of transportation* for Switzerland. In this section, we describe technical details of the agent population design starting with mapping data sources with their attributes, followed by an overview of the simulation process and the calibration procedure.

4.1 Agent specification

As mentioned in Section 3, the decision-making architecture requires 2 elements to calculate the probabilities for a set of options: (1) how to specify a ranking order of the option according to a determinant ($R_d(opt)$) and (2) the weight of the determinant (w_d). For this purpose, we utilise the Swiss Household Energy Demand Survey (SHEDS) [26]. There are several questions that compared the criteria for mobility mode choices, which answer can be interpreted as the weights(w_i) for different psychological determinants in TIB. A typical example is "Please rate how important the following aspects are for choosing this mode of transportation (from 1 to 5) - •Choosing the cheapest option; •Travelling as fast as possible, etc.". A large number of similar questions can be categories into TIB determinants. However, as the first step into this experimental design, we decided on a mapping of a smaller set (see Table 1), which is based on some of the past researches [2] and what properties can be measured or ranked objectively (using common sense). Note that in this case, the determinant *belief* is omitted since the system assumes that the knowledge/perception of agents is always correct.

Having the decision-making components figured, the next step is parametrising the profiles to build a synthetic population. This is accomplished by utilising another data source - the Mobility and Transport Microcensus [18], which includes the attributes listed in Table 2. Its entries (N = 57,091) are placed in a latent space (socio-matrix) that is represented by a symmetric Gower distance matrix [10]. All pairwise distances/dissimilarities are created based on the common features of the two data sources (e.g. age group, gender, region, household size, income level, number of personal vehicles). This matrix also provides a way to calculate the recommendation for agents from the same network (i.e. R_{role} - see Table 1). We then find the most similar peers that have the lowest distance towards each other and join them with entries from SHEDS (N=5,515). A random number of representatives for each geographical region in Switzerland are selected to become our agent population (N=3,080).

Along with the attributes in Table 2, a weekly schedule is also derived for each agent from microcensus to provide a way to calculate all relative costs for a trip (including purpose, distance, execution time). The agent's main purpose

[1] Behaviour-Driven Demand Model

Determinant	**Layer**	**Measuring/Ranking property ($R(opt)$)**	**Matching question(s) in SHEDS** (w, with scale 1-5)
Evaluation - *Price*	1st	R_{price} = Cost of travelling	w_{price} = •Choosing the cheapest option
Evaluation - *Time*	1st	R_{time} = Duration of the trip (including the journey to station)	w_{time} = •Travelling as fast as possible
Norm - *Environment Friendly*	1st	R_{norm} = Motor type of the vehicle (Gas/Electric/No motor)	w_{norm} = •In the Swiss society, it is usually expected that one behaves in an environmentally friendly manner
Role - *Environment Friendly*	1st	R_{role} = Recommend from other agents in its network	w_{role} = •Most of my acquaintances expect that I behave in an environmentally friendly manner
Self-concept - *Environment Friendly*	1st	$R_{self-concept}$ = No data available - **<u>to be calibrated</u>** (see Section 4.3)	$w_{self-concept}$ = •I feel personally obliged to behave in an environmentally friendly manner as much as possible
Emotion - *Enjoyment*	1st	$R_{emotion}$ = Vehicle's comfortableness/luxury	$w_{emotion}$ = •I enjoy this way of travelling
Frequency of past behaviours	1st	R_{freq} = The number of usage over a certain period	w_{freq} = •I am used to taking this means of transport
Attitude	2nd	$R_{attitude} = R_{price}/\sum_{price} * w_{price} + R_{time}/\sum_{time} * w_{time}$	$w_{attitude}$ = •Wealth(material possessions,money)
Social factors	2nd	$R_{soc} = R_{norm}/\sum_{norm} * w_{norm} + R_{role}/\sum_{role} * w_{role} + R_{self}/\sum_{self} * w_{self}$	w_{soc} = Avg(•Equality •Social power •Authority •Protect the environment •Influential •Helpful •Prevent pollution)
Affect	2nd	$R_{affect} = R_{emotion} * w_{emotion}$	w_{soc} = Avg(•Pleasure •Enjoying life •Self-indulgent)
Facilitating conditions	3rd	R_{cond} = Does the trip pass all constrains? (e.g. time, budget, vehicle's availability) (0/1)	Agent filters the options that are possible to be performed that the time of decision-making
Habit	3rd	$R_{habit} = R_{freq} * w_{freq}$	w_{habit} = •Habit and Routine: I do without thinking
Intention	3rd	$R_{intent} = R_{attitude}/\sum_{attitude} * w_{attitude} + R_{soc}/\sum_{soc} * w_{soc} + R_{affect}/\sum_{affect} * w_{affect}$	w_{intent} = MAX_SCALE - •I do without thinking
Decision	Output	$R_{decision} = (R_{intent}/\sum_{intent} * w_{intent} + R_{habit}/\sum_{habit} * w_{habit}) * R_{cond}$	

Table 1: Mapping of TIB's determinants and SHEDS to initiate decision-making weights

Attribute	Brief description
Location	Region (or *Cantons* in Switzerland) in which the agent is living
Budget	Weekly travelling budget
Accessibility set	List of available transportation services for the agent, which can be used to calculate all relative costs from a trip
Owned vehicles and Discounts	List of vehicles that the agent own
Weight to universe	The proportion of population that the agent represents

Table 2: An agent's state attributes

is to select a mode of transportation (including rail, car, bus, tram, biking, walking, others) to perform a task on its schedule. There is also an option of not performing the scheduled activity due to the constraints from the agent's states or environment (e.g. exhaustion of budget or exceeded travelling time on all available modes). Agents perform this filtering procedure before any decision-making activities (see determinant *Facilitating conditions* in Table.1).

4.2 Simulation procedure

The simulation process starts with a central controller creating all the agents with all their attributes and assigned them to their respective regions. Initial values for these attributes are coming from the mapping process above. The agent then looks at its individual schedule and creates decision-making events to be activated. At the time of simulation, the controller triggers these events simultaneously, waits for them to finish, then skips to the next scheduled point (i.e. event-driven). At this developing stage, no learning technique is applied for feedback loop inside the agent's decision-making process. Agents simply keep track of the number of times its used a vehicle for trips of the same purpose, which is used for determinant *habit* (see Table 1). After all the task finished, a reporter component in the region collects the final results.

4.3 Calibration

The purpose of calibration is to improve the compatibility of the current population with the target system. We are focusing on figuring out the most fitted ranking patterns of $R_{self-concept}$. Since the mapping question in SHEDS for this determinant is related to environmental friendly aspect of the option, we divided the agent population into 4 main profiles, depending on their daily main transportations: (1) soft-mobility modes (walking/biking), (2) public vehicles (tram/bus/train) (3) private vehicles (car/motorbike) and (4) others. $R_{self-concept}$ for each of them can then be calibrated by permuting the ranking order of all the modal choices.

Objective function: Our main objective is to minimise the error calculated the Eq.2. It is measured from the total differences between the final sum of kilometres in each mobility mode at the end of a period (i.e. a year in this case) and historical data. From microcensus [18], the total kilometres result for one year of all mobility profiles mentioned above can be obtained (i.e. walking/biking, bus/tram/train, car/motorbike, others). Assuming that no two modes can be ranked in the same position, calibration involves using the permutation of these four sets of modes as configurations for the $R_{self-concept}$. We repeat this procedure for all agent's profiles set at either deterministic (choose the best option) or stochastic (choose from a random function with probabilities provided by sampling distribution of final referenced values) to find the smallest error.

$$\underset{conf}{\text{minimise}} \quad err(conf) = \sum_{i=1}^{M} \mid census_i - sim_i(conf) \mid$$

where
- M = {walking/biking, bus/tram/train, car/motorbike, other}.
- $conf = S(M) \oplus S(M) \oplus S(M) \oplus S(M)$, an instance of the concatenation of two permutation sequences of M.
- $census_i$ is census data for mode i (in kilometres).
- $sim_i(conf)$ is the simutation result for mode i (in kilometres).

(2)

Type	**conf**	**CM**	**BTT**	**WB**	**O**	**err(conf)**
Census		72.7	27.5	8.6	3.7	n/a
Deterministic	R_{CM} = (1)CM, (2)BTT, (3)WB, (4)O R_{BTT} = (3)CM, (1)BTT, (4)WB, (2)O R_{WB} = (4)CM, (2)BTT, (1)WB, (3)O R_O = (2)CM, (4)BTT, (3)WB, (1)O	73.1	26.7	3.3	4.4	7.3
Stochastic	R_{CM} = (1)CM, (2)BTT, (4)WB, (3)O R_{BTT} = (3)CM, (1)BTT, (4)WB, (2)O R_{WB} = (4)CM, (3)BTT, (1)WB, (2)O R_O = (4)CM, (2)BTT, (3)WB, (1)O	46.7	6.0	5.0	4.6	51.9

Table 3: Calibration results[2,3]

Result: We list the kilometres in census data and the top results of two types of agents in Table.3. The best configuration is in the deterministic model with an error around 7.3x10^9 kilometres, which accounts for 6.5% of the total scheduled kilometres. The main differences are in the *public* (i.e. walking/biking) numbers. We also observe that the stochastic error are much larger - above 51.8x10^9

[2] All units are in 10^9 kilometres
[3] Abbreviation - CM: Car/ Motobike, BTT: Bus/Tram/Train, WB: Walking/Biking, O:Others

kilometres, which is only 46% accuracy. This is expected since agents in stochastic mode choose options based on a random function of probabilities derived from the referenced values. Currently, there is no pattern shown in the ranking function $R_{self-concept}$ of the results of *stochastic* mode, and hence additional runs with different distribution functions are needed in order to have a broader picture for this setting.

5 Conclusion and future direction

The tree-like and layered structure of TIB has inspired us to develop a new agent architecture that can combine many different determinants in human decision-making; each of which can also be enhanced by empirical data. This is potentially a useful tool to facilitate the engagement of socio-psychologists, economists and the general public with research projects. We aim to demonstrate its practicality by creating a fully-working model to predict trends in the mobility domain for Switzerland - BedDeM. An agent population has been created and calibrated with the data of Mobility and Transport Microcensus and SHEDS.

There is some small margin error from the calibration process (around 6.5% of the total scheduled kilometres). To address this, we are planning to focus on learning in the upcoming developing stage. As mentioned in Section 4.2, agents are currently keeping track of the number of times they used a mode on trips with the same purpose, which accounts for *habit* in decision-making. We also aim to capture the influence of past experience to the ranking function of elements such as *enjoyment*, and/or enable self-reflection by changing the weights of determinants. Reinforcement Learning techniques (e.g. [19]) can be utilised for these updates.

The next important step is assessing the model's uncertainty, variability and sensitivity. This can be done by selecting different representatives for the population when joining the two data sources. Although we have acquired the help of an economist specialised in environmental substantiality, it is also necessary to receive inputs from sociologist to derive alternative mappings of empirical data to TIB determinants (see Table 1) for more agent profiles. Another potential research direction is comparing the efficiency of Triandis' Theory with other similar behavioural theories (e.g. Theory of Planned Behaviour [1]) by also changing the mapping of determinants. The next wave of microcensus (available in 2020) is a potential source for this test.

In term of validation, one of the good direction for our model is determining whether the key relationship or mechanisms highlighted in the agent-based model seem to be plausible explanations of real-world phenomena, which often involves analysis of empirical data that is separate from the agent-based model. A good data source is SCCER-CEST [24], which can be used to indicate the pattern in demand for the transportation sector. Another way to do this is to design an experimental scenario aimed at capturing mechanisms of interest. It can be done with the support of an expert in sociology.

We close with a few words about software and documentation. As mentioned above, the core agent framework and BedDeM are developed in Java using an agent-based platform called RePast [21]. Although facing some problem with documentation, it is easy to understand and has reduced the learning curve for the development process. RePast is also actively updated for newer Java version and functionalities. We are using the R language to take care of handling and analysis to empirical input data. We also plan to publish the core architecture along with BedDeM's agent implementation to gather peer review. This will allows us to have feedback from multiple perspectives to improve the platform so that it can be employed for researches across different domains.

Acknowledgement

This project is part of the activities of SCCER CREST, which is financially supported by the Swiss Commission for Technology and Innovation (Innosuisse). The current version also utilises data from the *Mobility and Transport Microcensus* - 2015 edition, which provided by the Federal Office for Spatial Development (ARE) in October 2017.

References

1. Ajzen, I.: The theory of planned behavior. Organizational behavior and human decision processes **50**(2), 179–211 (1991)
2. Ampt, E., et al.: Understanding voluntary travel behaviour change. Transport Engineering in Australia **9**(2), 53 (2004)
3. Balke, T., Gilbert, N.: How do agents make decisions? a survey. Journal of Artificial Societies and Social Simulation **17**(4), 13 (2014)
4. Bowles, S., Gintis, H.: Social preferences, homo economicus, and zoon politikon. In: The Oxford handbook of contextual political analysis, pp. 172–186. Oxford University Press Oxford (2006)
5. Conte, R., Paolucci, M., et al.: Intelligent social learning. Journal of Artificial Societies and Social Simulation **4**(1), U61–U82 (2001)
6. Deffuant, G., Moss, S., Jager, W.: Dialogues concerning a (possibly) new science. Journal of Artificial Societies and Social Simulation **9**(1) (2006)
7. Edmonds, B.: Simulation and complexityhow they can relate. Virtual worlds of precision: computer-based simulations in the sciences and social sciences. Lit-Verlag, Münster pp. 5–32 (2005)
8. Gilbert, N.: When does social simulation need cognitive models. Cognition and multi-agent interaction: From cognitive modeling to social simulation pp. 428–432 (2006)
9. Gintis, H., Van Schaik, C., Boehm, C., Chapais, B., Flack, J.C., Pagel, M., Pruetz, J.D., Spikins, P., Whiten, A., Erdal, D., et al.: Zoon politikon: The evolutionary origins of human political systems. Current Anthropology **56**(3), 340–341 (2015)
10. Gower, J.C.: A general coefficient of similarity and some of its properties. Biometrics pp. 857–871 (1971)
11. Heckbert, S., Baynes, T., Reeson, A.: Agent-based modeling in ecological economics. Annals of the New York Academy of Sciences **1185**(1), 39–53 (2010)

12. Hedström, P.: Dissecting the social: On the principles of analytical sociology (2005)
13. Helbing, D.: Systemic risks in society and economics. In: Social Self-Organization, pp. 261–284. Springer (2012)
14. Hörl, S., Balac, M., Axhausen, K.W.: A first look at bridging discrete choice modeling and agent-based microsimulation in matsim. Procedia computer science **130**, 900–907 (2018)
15. Keane, M.: Current issues in discrete choice modeling. Marketing Letters **8**(3), 307–322 (1997)
16. Lady, B.B.F.: From pride and prejudice to persuasion satisficing in mate search. Simple heuristics that make us smart p. 287 (1999)
17. Laird, J.E.: The Soar cognitive architecture. MIT press (2012)
18. Mobility and Transport Microcensus. https://www.are.admin.ch/are/en/home/transport-and-infrastructure/data/mtmc.html, [Last accessed 16-Jan-2019]
19. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529 (2015)
20. Rao, A.S., George, M.P.: Modeling rational agents within a bdi-architecture, proceedings of the second international conference on principles of knowledge representation and reasoning, kr91, edited by allen, j., fikes, r., and sandewall, e. KR91 Allen, j., Fikes, R., Sandewall, E.(eds) Morgan Kauffman, San Matteo, CA (1991)
21. The Repast Suite. https://repast.github.io, [Last accessed 24-Jan-2019]
22. Safavian, S.R., Landgrebe, D.: A survey of decision tree classifier methodology. IEEE transactions on systems, man, and cybernetics **21**(3), 660–674 (1991)
23. Sawyer, R.K.: Artificial societies: Multiagent systems and the micro-macro link in sociological theory. Sociological methods & research **31**(3), 325–363 (2003)
24. The SCCER-CEST website. https://www.sccer-crest.ch/research-database/, [Last accessed 24-Jan-2019]
25. Schill, K., Zetzsche, C., Hois, J.: A belief-based architecture for scene analysis: From sensorimotor features to knowledge and ontology. Fuzzy Sets and Systems **160**(10), 1507–1516 (2009)
26. Swiss Household Energy Demand Survey (SHEDS). https://www.sccer-crest.ch/research/swiss-household-energy-demand-survey-sheds/, [Last accessed 16-Jan-2019]
27. Simon, H.A.: Bounded rationality. In: Utility and probability, pp. 15–18. Springer (1990)
28. Sun, R.: The clarion cognitive architecture: Extending cognitive modeling to social simulation. Cognition and multi-agent interaction pp. 79–99 (2006)
29. Sun, R.: Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge University Press (2006)
30. Sun, R.: The Cambridge handbook of computational psychology. Cambridge University Press (2008)
31. Sun, R., Naveh, I.: Simulating organizational decision-making using a cognitively realistic agent model. Journal of Artificial Societies and Social Simulation **7**(3) (2004)
32. Taatgen, N.A., Lebiere, C., Anderson, J.R.: Modeling paradigms in act-r. Cognition and multi-agent interaction: From cognitive modeling to social simulation pp. 29–52 (2006)
33. Todd, P.M., Billari, F.C., Simao, J.: Aggregate age-at-marriage patterns from individual mate-search heuristics. Demography **42**(3), 559–574 (2005)
34. Triandis, H.C.: Interpersonal behavior. Brooks/Cole Pub. Co. (1977)

Complexity metrics for Agent Based Models of Social Systems ⋆

Kiran Lakkaraju, Asmeret Naugle, Stephen J. Verzi, Laura P. Swiler, Michael Livesay, Christina Warrender, Michael Bernard, and Vicente Romero

Sandia National Laboratories

Abstract. We develop a multi-tiered approach to measure the complexity of agent based models of social systems, incorporating four interacting but complementary aspects of complexity: system intricacy, information theoretic complexity, behavioral capacity and social organization. We apply these metrics on the classic Schelling model of segregation as an example.

1 Introduction

Agent based modeling (ABM) has a long and rich history in studying social phenomena. The benefit of ABM is in developing simulations in which complex patterns emerge. The extent to which underlying micro-processes and the resulting patterns of behavior are similar to the real world is the question of validity and is an extensively studied question [7,2].

We focus on an aligned question, what does it mean for an agent based model to be "complex"? While there has been study of the computational complexity of multi-agent systems (MAS) [19] and the complexity of MAS software [12] there is very little work that studies the complexity of multi-agent systems in a way similar to that of the real world. Thousands of models with widely differing micro processes have explored historical, fictional and futuristic domains. How do we distinguish between different agent based models? How do we quantitatively compare across different models over different domains?

To address these questions, we propose a quantitative, multi-tier definition of complexity that can be used in studying agent based models of social systems. Our metric is founded on insights from complexity theory, the social sciences, and software engineering. By integrating multiple domains in the development of our metric we are able to better capture the multi-faceted nature of complexity.

Using our metric, agent based models of social systems can be quantitatively compared with each other, allowing us to better understand their utility.

⋆ Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525.

2 Ground truth

All agents have an underlying decision process that integrates perceptions of the environment, signals from other agents, and their own goals to determine behavior autonomously. We call this underlying decision model the "ground truth" of the model, since it dictates the behavior of agents in the simulation (the "micro level") and, thusly, the behavior of the simulation as a whole (the "macro level" effects [23]).

As a simple example, consider a 2-D cellular automata. Each cell perceives its neighbor's states and autonomously changes its state in response based on the rules of the model. The decision rule for each cell is the ground truth of the model. If the decision rule changes, the macro level simulation behavior can dramatically change [3].

A variety of models of agent decision making exist such as the Belief-Desire-Intention (BDI) cognitive model [20], or the Partially Observable Markov Decision Process (POMDP) model [18].

Our goal is to identify a simple representation of agent decision making that will capture the following:

- How agents in the simulation make decisions.
- How agents in the simulation interact with each other.
- How agents in the simulation interact with their environments.
- Any environmental factors that influence each other within the simulation.

We will constrain the ground truth to focus only on causal connections between variables and parameters within the model. Causality is defined by interrogation of the decision rules and equational relationships for the model. For example, if variable A is used in the equation/algorithm for calculating variable B, then we say that A causally impacts B.

We represent the ground truth as a graph, Nodes represent variables and functions/aggregations of variables, and edges between nodes represent causal relationships between nodes. For example, if variable A is used in the equation/algorithm for calculating variable B, then we include a link from Node A to Node B.

We allow nodes to represent functions/aggregations over agents or the environment (for instance, a node could be the average value of a state across all agents in the simulation).

In identifying the ground truth, we follow these principles:

- Nodes should be combined where possible. If there are multiple simulation variables that represent similar concepts and have the same causal structure (ie: the same causal influences), then those variables can be represented as a single node.
- Relationships between entities should be represented as simply as possible. For example, the ground truth does not need to represent the entire influence network between agents in the simulation; instead, links can represent types of causal relationships between generic agents.

– The exact functional form of equations and parameterization are not represented in the ground truth. The ground truth diagram is only meant to specify causal relationships.

Section 4 outlines the ground truth for the classic Schelling Segregation model.

3 Proposed multi-tier complexity metrics.

Table 1. Complexity metrics organization

	Not tied to social/behavioral science	Inspired by social/behavioral science
Requires knowledge of system structure	System Intricacy	Behavioral Capacity
Does not require knowledge of system structure	Information-Theoretic	Social Organization

Delineating between simple, intricate (sometimes also referred to as complicated), complex, and chaotic systems is a difficult task. Many definitions of complexity have been proposed in the literature [9,14] but no definition is widely accepted. We are focused on assessing the complexity of models of social systems. To measure the complexity of social simulations, we have identified a multi-tiered suite of metrics that captures different elements of complexity. Using a carefully chosen combination of methods, we can gain a deeper and more nuanced understanding of simulation complexity than could be achieved with a single metric.

The complexity metrics are organized along the two dimensions in Table 1. The first dimension (rows) differentiates between metrics that require knowledge of the system structure (i.e., ground truth) of a simulation and those that do not. Metrics that require knowledge of the system structure may be useful for causal simulations, but we generally do not have knowledge of the causal structure of real-world systems. However, if we can develop methods to infer this causal structure, these complexity metrics may apply on real-world systems.

The second dimension (columns) relates to the original intended application space of the metric. The right-hand column includes metrics that are inspired by the social and behavioral sciences, while metrics in the left-hand column measure more abstract properties of the simulation, and might be inspired by other application spaces or might be purely mathematical. We focus on the social and behavioral sciences since our focus is on modeling of social systems.

The four metrics are described in more detail below.

3.1 System Intricacy

Measures of system intricacy capture the complexity of a simulation's causal structure, or ground truth. These metrics are inspired by the notion that the more components and causal relationships a system has, the more complicated it is.

System intricacy is intimately tied to the causal structures, processes and interactions that determine the dynamics of the system. One approach for measuring system intricacy in simulations is to evaluate the complexity of the structure of the underlying software implementation, however these are not pure metrics of the system. We evaluate the system intricacy of a simulation by evaluating the simulation's ground truth.

Cyclomatic complexity (initially proposed in [13]) was initially developed for studying the complexity of software, however it has since been used in other domains [16]. We use it as a concise summary of the complicatedness of a simulations ground truth.

Cyclomatic complexity (M), captures the interconnectedness of a graph by counting the nodes (N), edges (E) and the number of connected components (P) in a graph:

$$M = E - N + 2P$$

3.2 Behavioral Capacity

Behavioral capacity measures capture the potential for rich and diverse interaction potential among agents in a system. The underlying hypothesis is that complex simulation of social processes will include significant and varied interaction between agents. Humans participate in a wide variety of groups, at multiple scales (from country membership to family groups). A complexity measure that captures this will capture an important part of human behavior.

A variety of metrics can be used to represent behavioral capacity of a social simulation, such as the number of interactions between agents, or the number of groups an agent participates in.

We focus on a measure that explicitly counts the number of relationships an agent has, the number of differentiated relationships, because of its intuitive appeal, ability to quantify, and prior work in the literature [1].

Intuitively, an agent that has multiple different types of relationships must juggle different goals and needs. An individual must do the same when they interact with a shopkeeper vs. family member. The difference in relationship can naturally track that of group membership.

Quantification of this measure can be done by viewing the ground truth of the model. Since the ground truth specifies all interactions between agents we should see evidence of differentiated relationships as types of influences and interactions agents can have with each other.

3.3 Information-Theoretic

Information-theoretic complexity measures capture information content related to the dynamics of a system. These metrics are inspired by the notion that a more complex system will generate more information over time. These metrics account for uncertainty, and are calculated using a systems (or simulations) input and/or output data (see [22] for a review). Information-theoretic complexity metrics have been developed and used in several fields. These metrics may not always capture our intuition of complexity; for example, these measures might consider randomness to be a form of complexity, since uncertainty and information content are entangled. We address this by considering information-theoretic complexity metrics in conjunction with the other three metric categories. These metrics are calculated using data directly from the social system or simulation results.

Many information theoretic metrics have been proposed in the literature, such as entropy [4], mutual information [4], autocorrelation [11], and compression ratios [10]. We focus on forecasting complexity (C) [22], which captures the minimum amount of information (H) needed for optimal prediction within a time-series, where part of the time-series, X^-, is used to predict the rest, X^+(such that $X = (X^-, X^+)$), using a model f in M (where M is a specific space of models):

$$C = \min_{f \in M} H(f(X^-))$$

Forecasting complexity captures an intuitive notion of complexity based on prediction, but it is hard to compute and requires a space of models (M) to search. Here, we use an approximation to forecast complexity involving compression ratio which itself is an approximation to normalized information distance [10] (which we call approximate NID).

For the information theoretic complexity on a time series of data, normalized information distance between the past and future information is defined as follows.

$$NID(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}},$$

We approximate this using split points computed over the entire time series giving a series of approximate NID complexities. This series of complexities is then averaged. The equation is given by

$$\text{approximate}NID(x, y) = \frac{\max\{Z(x|y), Z(y|x)\}}{\max\{Z(x), Z(y)\}},$$

where Z takes time-series information and gives the size of information after conditional compression. The LempelZiv-Markov chain algorithm (LZMA) compression is used because it takes and creates dictionaries for compression as it compresses and uses these dictionaries to compress new strings processed in the future, thus implementing a notion of conditional compression used here in approximate NID.

Let $S = \{S_1, ..., S_n\}$ be the data series. Then we are interested in the series approximate$NID(S_t, S_{t+1})$ for $t = 1, 2, ..., n-1$.

In theory approximate$NID(x, y) \in [0, 1]$, because in theory $Z(x|y) < Z(x)$ and Z is positive valued. The normalized p-norm is $(\sum_{i=1}^{n} f(i)^p \frac{1}{n})^{1/p}$. If $p = 1$ then indeed the mean is given. If $p = \infty$ (that is the limit as p gets arbitraily large), then the maximum is given. Note that $\lim_{p \to \infty} \frac{1}{n}^{1/p} = 1$. Since the information theoretic complexity is the normalized 2-norm it is in $[0, 1]$ and it is moderately influenced by the maximum and the average. But this complexity is resilient to maximums which are outliers, as well as large sequences of constant values which may be due to poor choice of cut-off times for the simulation.

3.4 Social Organization

Measures of social organization capture information about how individuals form groups, how groups combine to form larger groups, and how individuals and groups interact. These metrics are inspired by the idea that complex social systems demonstrate emergent hierarchical organization and complicated interactions between individuals and groups [17]. This category of metrics addresses the interaction between different levels of analysis, i.e., micro, meso and macro scale patterns within the system [6]. These metrics will be calculated using simulation results – the characteristics, states, and actions of agents during the simulation.

To preserve generality, we focus on measures that apply to a social network generated by a simulation. The social network represents interactions between agents in a simulation. A node in the social network represents an agent, and an edge represents interaction. We can extract interactions between agents from simulation output to create a social network.

Quantitative characteristics of a social network have been used to characterize real world social systems ([5]). Existing literature suggests a variety of means to capture different aspects of social networks, including clustering coefficients, community detection algorithms, and centrality measures.

We focus on measures that can capture the hierarchy within a social system. Hierarchies are an important concept in the social and physical sciences, and have been considered a fundamental characteristic of complex systems [17]. To quantify the hierarchy within a simulation we focus on the Global Reaching Centrality measure (GRC) as defined in [15], which uses the local reach centrality.

Let $C_R(i)$ be the local reach centrality of node I, the proportion of nodes that can be reached from node I via outgoing edges. Then the Global Reach Centrality (GRC) is defined as:

$$GRC = \sum_{i \in V} \frac{[C_R^{\max} - C_R(i)]}{(N-1)}$$

This definition can be easily extended to undirected networks by considering the weights on the edges (with a default weight of 1) for every edge. The GRC can range from 0 to 1, with a higher value indicating a higher level of hierarchy in the social network [15].

4 Example application: Schelling segregation model

As an illustrative example, and to highlight potential difficulties, we consider applying our complexity metrics to the classic Schelling segregation model [21]. Our goal is not to extensively evaluate the Schelling model, but rather to highlight the promise, and understand the pitfalls, of our multi-tier complexity metric.

We use the NetLogo implementation of the Schelling model [24]. Agents are characterized by a color that is fixed throughout the simulation and are located on a 2-d lattice. Only one agent can be at any single point on the lattice.

Agents have a preference to be with like-minded (i.e., same color) agents. The premise is that if an individual has a preference to be in a neighborhood where a larger percent of neighbors have similar traits to themselves than their current neighborhood, then and only then are they motivated to move to a new residence (empty lattice site).

By leaving, the agent has positively reinforced the current dominating trait; while on the other hand, this agent's presence at their new location reinforces their own trait at the new location. Such an act reinforces the average mindset of the neighborhood, causing any unlike-minded neighbors to be even more outnumbered, hence they have reinforced segregation on multiple fronts. Due to this positive feedback, even a slight intolerance (such as the need for 26% like-minded neighbors) could potentially lead to highly segregated regions.

We consider a simple version of the Schelling model here, but there has been extensive study of this model, see [8]

Ground truth for the Schelling model Agents in the Schelling model have the following characteristics (fixed features of the agent) and states (dynamic features of an agent):

Characteristics :
- **Color** An actor has a color that is fixed throughout the simulation.
- **Preference Ratio** The percentage of neighbors of an actor that should share the same color, denoted as p_{pref}.

Behaviors :
- **Change Location** An actor can take the action to change it's location to another empty location in the grid.

States :
- **Location** Location of an actor on the grid.

Let p_{matching} be the fraction of neighbors of an actor that have the same color as the actor. On a time step, the decision rule an actor executes is the following:

$$\textbf{if } p_{\text{matching}} < p_{\text{pref}} \textbf{ then } \text{Actor Changes Location}$$

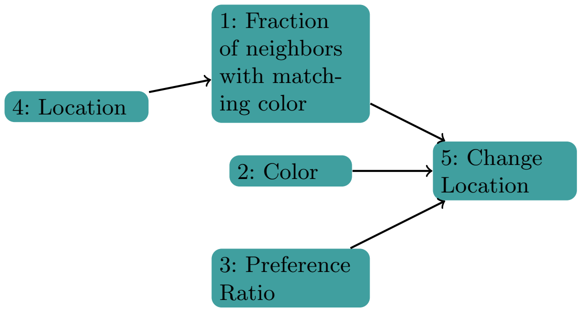


Fig. 1. Ground Truth for the Schelling model

Fig. 1 is the ground truth for the Schelling segregation model.

Node 1 is an aggregation of the characteristics of neighboring agents. Node 2 is the color of an agent. Node 3 is the preference ratio for an agent. Node 4 is the location of an agent, and Node 5 is the behavior an agent undertakes.

In this specific implementation, all agents have the same preference ratio. Even if agents had different preference ratios, the ground truth would not change as it only captures the fact that there is a relationship between preference ratio and the behavior to change a location.

5 Parameters and Types of Behaviors

The Information-Theoretic and Social Organization measures are computed on simulation output and are thus dependent upon the parameterization of the simulation. There are several parameters in the Schelling model that can influence the system behavior.

Density The fraction of locations on the 2-D lattice on which an agent resides. If the density of the population is too low then all the agents may have no desire to move, as they have no neighbors. If the density were 100% then the agents could not move. Moreover, given a good value of density there exist high values of preference so that the agents never settle; likewise there are such low values of preference that they settle immediately.

Preference Ratio Setting If the preference is above 80% (and even above 75% for most densities) then enough agents choose to move that the simulation becomes chaotic and uninformative. If the preference ratio is too low, agents do not desire to move at any time and the simulation exhibits no dynamics.

The range of density in NetLogo's segregation model is $50\% - 99\%$, all of which are determined to be reasonable. For densities less than 50%, the simulation converges too quickly or gets stuck in a random cycle. When the preference for the agents is too high, then on the boundary of the segregated regions the agents will choose to move and move randomly, making the boundary grow. Eventually, the entire population is moving randomly, rarely staying in any location.

6 Application and discussion of Complexity Metrics

6.1 System Intricacy & Behavioral Capacity

For the Schelling model, the system intricacy is $M = 4 - 5 + 2 * 1 = 1$. The behavioral capacity is simply the number of differentiated relationships, which is determined to be 1.

This aligns with intuition and general perception. In fact, its importance derives from the fact that so few elements are needed to produce what is thought to be a complex pattern of behavior.

All agents in the Schelling model interact with each other in the same way, by evaluating their color. A counter argument would be that since color determines action, and there is a different action for agents that are of a different color, that would indicate a different relationship. However, note that the action of an agent does not have a subject – no agent does anything to another agent. This is the underlying characteristic of a differentiating relationship, one in which there are different actions towards different agents, of which there are none here.

6.2 Information Theoretic Complexity

To apply our information theoretic measure, we need to identify the appropriate information to collect at each time step. In this simple model we can use the states of every location on the grid as a representation of the simulation at each time step.

Table 2 shows the information theoretic complexity values for a variety of parameter settings, chosen to highlight different behaviors.

This metric aligns with our intuition. When there is near instant convergence, the information theoretic complexity is low (0.38 for density = 50% and preference ratio at 50%) vs. situations in which there is lots of movement (density = 99% and preference ratio of 60%).

We acknowledge a weak correlation with the number of timesteps, but note that our 2-norm method alleviates some of that. Compare the values for density=90% vs. density=50% for the preference ratio value of 50%, we can see that even with fewer time steps to converge the information theoretic measure was higher.

Density	Preference 50%	Preference 60%
99%	0.64 (35)	0.96 (1001)
90%	0.59 (19)	0.72 (99)
80%	0.46 (25)	0.63 (33)
70%	0.49 (21)	0.58 (30)
60%	0.47 (20)	0.59 (23)
50%	0.38 (28)	0.62 (19)

Table 2. Information Theoretic complexity examples at convergence (time steps to convergence indicated in parentheses, simulation stopped at 1000 timesteps if not converged), given density and preference.

6.3 Social Organization

We define the social network for the Schelling model in the following way. An unweighted edge is established between two agents if they were neighbors at any point during the simulation. The global reaching centrality (GRC) was calculated on this network, see table 3.

Density	Preference 50%	Preference 60%	Preference 70%
99%	0.1648	0.1628	0.1583
90%	0.0931	0.1012	0.0898
80%	0.0880	0.1129	0.0537
70%	0.0801	0.0979	0.0494
60%	0.0825	0.0954	0.0654
50%	0.0870	0.0910	0.0717

Table 3. Global Reaching Centrality (GRC) examples after 500 time steps, given density and preference. Given a particular density, it would appear that the GRC is at its highest for the preferences which converge the most quickly. This table does not include parameters for when the agents move randomly forever.

The GRC measure of the simulation runs can be compared to real world examples. [15] calculates the GRC for a variety of real world graphs. Food webs have a high GRC. Surprisingly, trust in an organization has low GRC scores. Our results show quite a low value of GRC for the parameter setting, indicating that the social network we defined based on neighbors is not very hierarchical. This makes sense, as agents move around in the grid. We can also notice a pattern of decreased hierarchy as the density increases. There should be a correlation between increase in density and agent moving (if they desire to move, but can't find a place to move, they will continue to desire to move).

System Intricacy	Behavioral Capacity
1	1
Info.-Theoretic	**Social Organization**
0.59	0.10

Table 4. Summary of complexity metric for the Schelling model. Info. Theoreticand Social Organization values are means over parameters settings defined in Table 2 and Table 3, respectively.

7 Discussion & Conclusion

Table 4 summarizes our assessment of the Schelling model using the multi-tier complexity metrics. In any simulation there will be critical parameters that can impact simulation behavior. Two of our metrics are sensitive to those parameters since they are dependent upon the simulation output. As we have done, to appropriately use our complexity metrics it is necessary to characterize behavior in the parameter spaces. This may require significant resources (computational time).

Each individual metric, taken by itself, will have some drawbacks. However, when considered together we believe they capture important classes of models. As a guide, consider classifying systems based on the system structure metrics (System Intricacy and Behavioral Capacity) vs. the metrics that are based on simulation output (Information-Theoretic and Social Organization). We can determine four different cases based on whether the values for system structure and simulation output, respectively, are "low" or "high":

Low-Low These are systems which are not intricate and do not produce unpredictable behavior. A simple linear system, such as $x_t = 1.1 * x_{t-1}$ may be an example of this.

High-Low These are intricate systems that produce simple behavior. This is interesting as the intuition is that an increase in the complexity of the causal structure of a model should result in an increase in complexity. However, these systems do not exhibit this characteristic.

Low-High These are systems, where a simple, non-intricate set of rules can determine complex behavior. Examples abound of these types of models, for instance Rule 110 in the Cellular Automata literature is a simple decision rule but is shown to be Turing complete [3].

High-High These are systems that capture our intuitive notion that more structurally complex systems will have higher complexity. Most real world system, which will contain a multitude of entities interacting over continuous time and space would fall into this category.

We have developed these metrics to be used in conjunction with each other. Together, they can help characterize and organize different model.

References

1. Thore J Bergman and Jacinta C Beehner. Measuring social complexity. *Animal Behaviour*, 103:203–209, May 2015.
2. Carlo Bianchi, Pasquale Cirillo, Mauro Gallegati, and Pietro Vagliasindi. Validating and Calibrating Agent-Based Models: A Case Study. *Computational Economics*, 30(3):245–264, 2007.
3. Keith C. Clarke. Cellular Automata and Agent-Based Models. In Manfred M. Fischer and Peter Nijkamp, editors, *Handbook of Regional Science*, pages 1217–1233. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
4. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, New York, NY, USA, 2006.
5. R. I. M. Dunbar, Valerio Arnaboldi, Marco Conti, and Andrea Passarella. The structure of online social networks mirrors those in the offline world. *Social Networks*, 43:39–47, October 2015.
6. Jessica C. Flack and David C. Krakauer. Challenges for complexity measures: A perspective from social dynamics and collective social computation. *Chaos (Woodbury, N.Y.)*, 21(3):037108, September 2011.
7. Mattia Guerini and Alessio Moneta. A method for agent-based models validation. *Journal of Economic Dynamics and Control*, 82:125–141, September 2017.
8. Erez Hatna and Itzhak Benenson. The Schelling Model of Ethnic Residential Dynamics: Beyond the Integrated - Segregated Dichotomy of Patterns. *Journal of Artificial Societies and Social Simulation*, 15(1):6, 2010.
9. James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67, Jan 2013.
10. Ming Li, Xin Chen, Xin Li, Bin Ma, and P. M. B. Vitanyi. The similarity metric. *IEEE Transactions on Information Theory*, 50(12):3250–3264, Dec 2004.
11. Benoit B. Mandelbrot. *Fractals and Scaling In Finance: Discontinuity, Concentration, Risk*. Springer Publishing Company, Incorporated, 1st edition, 2010.
12. Toufik Marir, Farid Mokhati, Hassina Bouchelaghem-Seridi, and Zouheyr Tamrabet. Complexity Measurement of Multi-Agent Systems. In *Multiagent System Technologies*, pages 188–201. Springer, Cham, Cham, September 2014.
13. T.J. McCabe. A Complexity Measure. *IEEE Transactions on Software Engineering*, SE-2(4):308–320, December 1976.
14. Melanie Mitchell and Mark Newman. Complex Systems Theory and Evolution. In *Encyclopedia of Evolution*. Oxford University Press, 2002.
15. Enys Mones, Lilla Vicsek, and Tamás Vicsek. Hierarchy Measure for Complex Networks. *PLOS ONE*, 7(3):e33799, March 2012.
16. Fivos Papadimitriou. Artificial intelligence in modelling the complexity of mediterranean landscape transformations. *Comput. Electron. Agric.*, 81:87–96, February 2012.
17. Denise Pumain. *Hierarchy in Natural and Social Sciences*. Springer Science & Business Media, November 2005. Google-Books-ID: lq3jgwe96H8C.
18. David V Pynadath and Stacy Marsella. PsychSim - Modeling Theory of Mind with Decision-Theoretic Agents. In *International Joint Conference on Artificial Intelligence*, 2005.
19. Zinovi Rabinovich, Claudia V Goldman, and Jeffrey S Rosenschein. *The complexity of multiagent systems: the price of silence*. the price of silence. ACM, New York, New York, USA, July 2003.

20. Anand S. Rao and Michael P. Georgeff. BDI Agents: From Theory to Practice. In *In Proceedings of the First International Conference on Multi-Agent Systems (icmas-95*, pages 312–319, 1995.
21. Thomas C. Schelling. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186, July 1971.
22. Cosma Rohilla Shalizi. Methods and Techniques of Complex Systems Science: An Overview. In Thomas S Deisboeck and J Yasha Kresh, editors, *Complex Systems Science in Biomedicine*, pages 33–114. Springer US, Boston, MA, 2006.
23. Edward Bishop Smith and William Rand. Simulating Macro-Level Effects from Micro-Level Observations. *Management Science*, 64(11):5405–5421, December 2017.
24. Uri Wilensky. Netlogo 4.0. 4. 2008.