Moral Values in Norm Decision Making*

Marc Serramia
Maite Lopez-Sanchez
Universitat de Barcelona, Spain
maite_lopez@ub.edu

Juan A. Rodriguez-Aguilar
Manel Rodriguez
Artificial Intelligence Research Institute, IIIA
Spanish National Research Council, CSIC
{jar,manel.rodriguez}@iiia.csic.es

Michael Wooldridge
Javier Morales
University of Oxford, UK
{mjw,javier.morales}@cs.ox.ac.uk

Carlos Ansotegui
University of Lleida, Spain
carlos@diei.udl.es

ABSTRACT

Most often, both agents and human societies use norms to coordinate their on-going activities. Nevertheless, choosing the 'right' set of norms to regulate these societies constitutes an open problem. Firstly, intrinsic norm relationships may lead to inconsistencies in the chosen set of norms. Secondly, and more importantly, there is an increasing demand of including ethical considerations in the decision making process. This paper focuses on choosing the 'right' norms by considering moral values together with society's partial preferences over these values and the extent to which candidate norms promote them. The resulting decision making problem can then be encoded as a linear program, and hence solved by state-of-the art solvers. Furthermore, we empirically test several optimisation scenarios so to determine the system's performance and the characteristics of the problem that affect its hardness.

KEYWORDS

Normative systems, value-based reasoning, norm decision making.

ACM Reference Format:
Marc Serramia Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar Manel Rodriguez, Michael Wooldridge Javier Morales, and Carlos Ansotegui. 2018. Moral Values in Norm Decision Making. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 9 pages.

1 INTRODUCTION

Norms have been extensively studied as coordination mechanisms within both agent and human societies [8, 39]. Within agent societies, problems such as norm synthesis [1, 40], norm emergence [18, 42], or norm learning [9, 36, 37] have been widely tackled. Regarding human societies, e-participation and e-governance ICT systems are currently attracting a lot of attention [13, 26, 43]. Thus, for example, some regulatory authorities in European cities –such as Reykjavik [34], Madrid [12], or Barcelona [11] municipalities– are opening their policy making to citizens. This is also the case for some countries: New Zealand authorities are opening consultations about legislations related to different topics such as family violence [29] or pensions [22]; whereas governamental site Parlement et Citoyens [16] opens French Parliament to citizens so that they can participate in national law elaboration.

Beyond the intrinsic complexity due to the number of norms to manage –either if they are proposed by humans or automatically generated– choosing the norms to regulate a society constitutes a complex process. The reasons are twofold.

On the one hand, norms can be related. Norm relationships have been previously studied in the literature. Thus, for example, Grossi and Dignum [19] study the relation between abstract and concrete norms, whereas Kollingbaum, Vasconcelos et al. [25, 41] focus on norm conflicts —and solve them based on first-order unification and constraint solving techniques. In this paper, we borrow some of the relationships identified in Morales et al. [33] and follow the work by Lopez-Sanchez et al. [27], which characterises three different binary norm relationships, namely, generalisation, exclusivity, and substitutability. Thus, we can consider a set of norms and the fact that some norms in this set generalise some specific norms; that some other norms are pair-wise incompatible (i.e., mutually exclusive); or interchangeable (that is, substitutable). When this is the case, a regulatory authority should not select these norms to be simultaneously established in the society. We encode these relationships in terms of restrictions in a linear program that computes the norm subsets (subsets of the given set of norms) compliant with the constraints imposed by the associated norm relations.

On the other hand, the work in [27] characterises the problems that regulation authorities confront when considering two different preference criteria over the norms to impose. In this manner, they specify the optimisation problem of finding the subset of norms that, in addition to complying with the relation constraints, maximizes represented norms while minimizing associated norm costs (since norms have deployment costs). This paper presents an empirical analysis of different optimisation scenarios so to characterise the hardness of the problem at hand and to assess how norm relationships affect its performance.

Our main contribution is the consideration of moral values associated to norms as an additional decision criterion. This has to be necessarily the case if we consider the different initiatives that are flourishing revolving the application of Artificial Intelligence

*Research supported by projects TASSAT3: TIN2016-76573-C2-2-P and Collectiveware TIN2015-66863-C2-1-R (MINECO/FEDER). Michael Wooldridge was supported by the European Research Council under Grant 291528 (âĂIJRACEâĂİ). Javier Morales was funded by the H2020-MSCA-IF project number 707688.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

in many aspects of our daily lives [21]. Obviously, this includes automatic decision making, since most resulting decisions have ethical implications. Thus, along with the lines of the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems [3], with a committee devoted to "Embedding Values into Autonomous Intelligent Systems", we include moral values in the selection of the norms that will be enacted in a society.

Values have been studied in argumentation. Some representative examples are Bench-Capon et al.[4, 7] or Modgil [30], which work with different Value-based Argumentation Frameworks. Values have also been considered in artificial moral agents (AMA) [44], introduced by Kohler et al. [24] in multi-agent institutions and, more recently, Luo et al. [28] use them when considering agent opportunistic behaviours. However, to the best of our knowledge, no other authors have considered the values that norms support. Instead, the normative multi-agent systems research area has focused in different normative concepts such as minimality and simplicity [17, 32], liberality [33], compactness [31], or stability [39].

In this paper we assume that a regulatory authority has available a collection of candidate norms (and their relationships) together with a moral system that reflects the moral values shared within a society. Then, we model the problem that pursues to maximize the set of norms to establish under (a combination of) several criteria regarding norm representation, associated costs, and supported moral values. Subsequently, despite the computation complexity of these problems, state-of-the-art linear programming solvers are used to automatically compute their solution.

The paper is structured as follows. Sections 2 and 3 introduce norm related concepts and a norm decision-making problem. Section 4 introduces moral values into this multi-objective optimisation problem. Sections 5 and 6 describe our empirical evaluation, and section 7 discusses its application and future work.

2 BASIC DEFINITIONS

This section introduces norms and other related concepts as the fundamental building blocks of the problems at hand.

Most normative multiagent systems literature [2] considers norms as coordination mechanisms. However, they differ in their formal definition. Our notion of norm is based on a simplification of the one in [45]. Thus, here we formally consider a norm n_i as a pair $\theta(\rho, ac)$, where: θ is a deontic operator (prohibition, permission, or obligation); ρ is a description of the addressee entity, namely, the agent required to comply with the norm; and ac is an action –from a set of actions– that entities can perform in a specific domain. Example 2.1 illustrates this definition.

As aforementioned, norm relations have also been previously studied [19, 25, 41]. Here, we borrow some of the relations from [33] and follow the work in [27], which characterises three norm relationships (namely, *exclusivity*, *substitutability*, and *generalisation*). Informally, it is considered that: two norms are mutually exclusive when they are incompatible; two norms are substitutable if they are interchangeable; and a norm is more general than another one when it subsumes its regulation (has wider regulation scope).

Specifically, considering N as a non-empty set of norms, the *exclusivity* relation is a binary relation $R_x \subseteq N \times N$. If $(n_i, n_j) \in R_x$ we say that n_i, n_j are incompatible or mutually exclusive. R_x is an irreflexive, symmetric, and intransitive relation.

The *substitutability* relation is a binary relation $R_s \subseteq N \times N$. If $(n_i, n_j) \in R_s$, we say that norms n_i, n_j are interchangeable or substitutable. Based on substitutability relationships, we introduce the notion of *substitution chain* as follows. Given two norms, $n_i, n_k \in N$, we say that n_k is *connected by substitutabilities* to n_i if there is a non-empty subset of norms $\{n_1, \ldots, n_p\} \subseteq N$ such that $(n_1, n_2), \ldots, (n_{p-1}, n_p) \in R_s$, $n_1 = n_i$, and $n_p = n_k$. Henceforth, a new relationship $\mathcal{S} \subseteq N \times N$ will contain the pairs of norms that are connected by substitutabilities. In particular, notice that if $(n_i, n_j) \in R_s$, then $(n_i, n_j) \in \mathcal{S}$. R_s is an irreflexive, symmetric, and transitive relation.

Finally, the *direct generalisation* relation is a binary relation $R_g \subseteq N \times N$. If $(n_i, n_j) \in R_g$, we say that n_i is more general than n_j (i.e., it generalises n_j). R_g is irreflexive, anti-symmetric, and intransitive (this is so because if $(n_i, n_j) \in R_g$, $\nexists n_k \in N$ s.t $(n_i, n_k), (n_k, n_j) \in R_g$). Transitivity is captured through the notion of indirect generalisation and the so-called *ancestors*. Given two norms, $n_k, n_i \in N$, we say that n_k is an ancestor of n_i if there is a non-empty subset of norms $\{n_1, \ldots, n_p\} \subseteq N$ such that $(n_k, n_1), \ldots, (n_p, n_i) \in R_g$. Henceforth, given a norm $n_i \in N$, we will note its ancestors as $\mathcal{A}(n_i)$. Notice that if $(n_j, n_i) \in R_g$ then $n_j \notin \mathcal{A}(n_i)$.

Previous concepts allows us to define the so-called norm net.

Def. 1. *A norm net is a pair NN= $\langle N, R \rangle$, where N stands for a set of norms and $R = \{R_g, R_x, R_s\}$ contains generalisation, exclusivity and substitutability relationships over the norms in N. The relationships in R are mutually exclusive, namely $R_g \cap R_x = \emptyset$, $R_g \cap R_s = \emptyset$, and $R_x \cap R_s = \emptyset$.*

Furthermore, given a norm net *NN*= $\langle N, R \rangle$, we will refer to any subset of the norms in N as a *norm system*. The challenge then lies in selecting a norm system out of a norm net. In general, we will be interested in norm systems incorporating as many norms as possible but excluding overlapping nor conflicting norms. Thus, considering that exclusivity relationships capture conflicts between norms whereas substitutability and generalisation relationships capture *redundancy* or overlap (or in the case of generalisation, subsumption), the following characterisation of norm systems naturally follows.

Def. 2. *Given a norm net NN= $\langle N, R \rangle$, we say that a norm system $\Omega \subseteq N$ is conflict-free iff for each $n_i, n_j \in \Omega$, $(n_i, n_j) \notin R_x$.*

Def. 3. *Given a norm net NN= $\langle N, R \rangle$, we say that a norm system $\Omega \subseteq N$ is non-redundant iff for each $n_i, n_j \in \Omega$: (i) $(n_i, n_j) \notin R_g$ and $n_j \notin A(n_i)$; and (ii) $(n_i, n_j) \notin \mathcal{S}$*

Def. 4. *Given a norm net NN= $\langle N, R \rangle$, we say that a norm system $\Omega \subseteq N$ is sound iff it is both conflict-free and non-redundant.*

Thus, we aim at finding sound norm systems that satisfy certain criteria. Next section is devoted to further elaborate on that. Before that, though, we illustrate norm nets with an example.

Example 2.1. Figure 1 illustrates an example of a norm net that includes some norms (rules) of border control at an international airport. Norms are depicted as circles labeled as $n_1, \ldots, n_5$ respectively. In particular, they are defined as follows:
n_1 : *Permission*(*all_passengers*, *cross_border*)
n_2 : *Obligation*(*all_passengers*, *register_passport*)
n_3 : *Obligation*(*all_passengers*, *fulfil_form*)

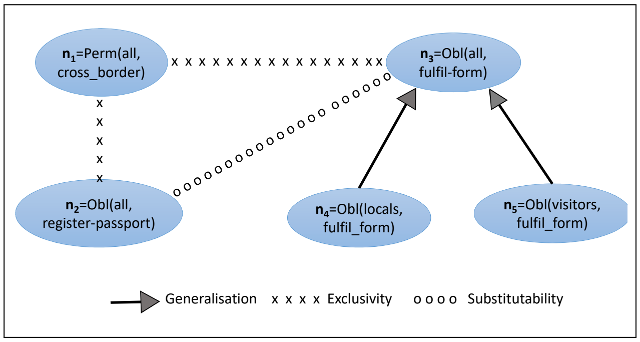

Figure 1: Norm net example: rules of border control at an international airport.

n_4 : *Obligation*(*locals*, *fulfil_form*)
n_5 : *Obligation*(*visitors*, *fulfil_form*)
Norm n_1 rules free movement of passengers, allowing all passengers to cross the border without any additional action. Conversely, norm n_2 requires all passengers to register their passport, and there is still a third rule n_3 that requires them to fulfil a form asking for passport information. As Figure 1 depicts, norm n_1 has an exclusive relation with both norms n_2 and n_3 ($(n_1, n_2), (n_1, n_3) \in R_x$)). Additionally, there is a substitutability relationship between second and third norms ($(n_2, n_3) \in R_x$). Finally, norm n_3, –which requires all passengers to fulfil a form– is a generalisation of norms n_4 and n_5, which respectively require local and foreign passengers (visitors) to fulfil a form ($i.e.$, $(n_3, n_4), (n_3, n_5) \in R_g$). Hence, considering these restrictions $\Omega = \{n_1, n_4\}$ constitutes a possible sound norm system.

3 NORM DECISION MAKING

As mentioned above, the work in [27] characterises the problem of determining sound norm systems that meet some optimization criteria. This section details how soundness constraints and two specific criteria –namely, norm representation and norm deployment cost– can be encoded as a linear program that can in turn be solved with state-of-the-art solvers (e.g. CPLEX [20], Gurobi [35]).

On the one hand, norm representation is a preference criteria that encapsulates the purpose of incorporating as many norms as possible out of those proposed in a norm net, since they are all considered to be acceptable. Thus, informally, we will aim at the norm system that *represents* the largest number of norms in the norm net. On the other hand, we cannot ignore the fact that norm deployment has associated costs and that the expenses derived from imposing norms should be bounded by the available budget. Norm costs may represent monetary expenses derived from regulatory processes –such as norm establishment or norm enforcement– as well as non-monetary aspects –such as social implications or political correctness– as long as they can be somehow quantified. In any case, we intend to minimise incurred costs while maximizing norm representation.

Formally, we consider a representation power function, $r : N \longrightarrow \mathbb{R}$ to be a linear function that computes a norm's representation power: a real value that encapsulates the fact that a norm cannot only represent itself but also all the norms it generalises. Besides linearity, the only condition that we impose on r is that $r(n_i) \leq r(n_j)$ for each $n_j \in \mathcal{A}(n_i)$. Hence, the representation power of a normative system Ω can be readily obtained by adding the representation power of its norms, namely $\rho(\Omega) = \sum_{n \in \Omega} r(n)$.

Similarly, we assume that the cost of a norm system can be obtained by adding the individual costs of its norms, namely $cost(\Omega) = \sum_{n_i \in \Omega} c(n_i)$, where $c(n_i)$ stands for the cost of norm n_i. Furthermore, we make the (reasonable) assumption that $cost(\Omega)$ is bounded by a maximum budget b that is available to cover the expenses of imposing those norms in the resulting norm system.

From that, we can cast the decision problem as the following multi-objective optimisation problem.

Problem 1. *Given a norm net NN=* $\langle N, R\rangle$[1]*, a representation power function r, and a fixed budget b, the maximum norm system problem with limited budget (or the maximum norm system problem, for short) is the problem of finding a sound norm system* $\Omega \subseteq N$ *with maximum representation power and minimum cost limited by some non-negative budget b (i.e., there is no other norm system* $\Omega' \subseteq N$ *such that* $\rho(\Omega') > \rho(\Omega)$, $cost(\Omega') < cost(\Omega)$, *and* $cost(\Omega) \leq b$*).*

Lemma 1. *The complexity of the maximum norm system problem is at least NP-Hard.*

Proof. 1. *The proof goes trivially by reduction of the maximum independent set problem, which is known to be an NP-Hard optimisation problem [23], to the maximum norm system problem. Consider that we want to find the maximum independent set of a graph* $G = (V, E)$*. Now say that each vertex in V stands for a norm and each edge in E stands for an exclusivity relationship in* R_x*. From this follows that finding the maximum independent set of G amounts to solving the maximum norm set problem on the norm net* $\langle V, \{R_x\}\rangle$*, where the representation power and cost functions are defined as* $r(v) = 1$*,* $cost(v) = 1$ *for each* $v \in V$ *and* $b = |V|$*.*

Next, we show how to solve the maximum norm system problem by encoding the optimisation problem as a linear program. Thus, consider a norm net *NN*= $\langle N, R\rangle$, and a set of binary decision variables $\{x_1, \ldots, x_{|N|}\}$, where each x_i encodes the decision of whether norm n_i is selected (taking value 1) for a norm system or not (taking value 0). Thus, solving the maximum norm system problem amounts to balancing the maximisation of $\sum_{i=1}^{|N|} x_i \cdot r(n_i)$ with the minimisation of $\sum_{i=1}^{|N|} x_i \cdot c(n_i)$ subject to a number of constraints (including soundness restrictions).

First, *exclusivity constraints* prevent that two mutually exclusive (incompatible) norms are jointly selected to be part of a norm system. Thus, the following constraints must hold:

$$x_i + x_j \leq 1 \quad \text{for each } (n_i, n_j) \in R_x \tag{1}$$

Second, *substitutability constraints* avoid that interchangeable norms are simultaneously selected. This amounts to enforcing that any pair of norms that are connected by substitutabilities cannot be simultaneously selected, namely:

$$x_i + x_j \leq 1 \quad \text{for each } (n_i, n_j) \in \mathcal{S} \tag{2}$$

[1]Notice that we assume knowledge about candidate norms to enact and the relationships between such norms.

Third, *generalisation constraints* avoid redundancy by imposing that a norm cannot be selected together with any of the norms that it directly generalises. Given a norm n_i, the set of norms generalised by n_i is defined as $Children(n_i) = \{n_j | (n_i, n_j) \in R_g\}$. Then, formally, the following constraints must hold:

$$x_i + x_j \leq 1 \quad n_j \in Children(n_i) \;\; 1 \leq i \leq |N| \tag{3}$$

Moreover, all the children of a norm cannot be simultaneously selected. Formally:

$$\sum_{n_j \in Children(n_i)} x_j < |Children(n_i)| \quad 1 \leq i \leq |N| \tag{4}$$

Additionally, a norm cannot be simultaneously selected together with any of its ancestors, namely:

$$x_i + x_k \leq 1 \quad n_k \in \mathcal{A}(n_i) \quad 1 \leq i \leq |N| \tag{5}$$

Fourth, we must also consider the binary constraints corresponding to the norm decision variables, namely:

$$x_i \in \{0, 1\} \quad 1 \leq i \leq |N| \tag{6}$$

Finally, a further constraint ensures that the cost of the norm system does not exceed the limited budget $b \geq 0$:

$$\sum_{i=1}^{|N|} c(n_i) \cdot x_i \leq b, \tag{7}$$

The linear program encoding the maximum norm system problem requires $|N|$ binary decision variables, $2 \cdot |R_g| + |R_x| + |\mathcal{S}|$ pairwise constraints (equations 1, 2, 3, and 5), and $|P(Rg)|$ inequality constraints (equation 4), where $P(R_g) = \{n_i | (n_i, n_j) \in R_g\}$. Hence, the number of constraints is linear with the number of norm relationships in a norm net.

4 CONSIDERING MORAL VALUES

So far we have considered quantitative criteria that regulation authorities can take into account when choosing the norms to enact in a society. However, as stated in the introduction, there is an increasing demand of including ethical considerations in the norm decision making process. Bringing in moral values into the norm decision making can be regarded as a qualitative criterion.

In fact, both the law (or its philosophical approach, deontology) and moral philosophy (ethics) act as systems of recommendations on which possible actions are to be considered 'right' or 'wrong' [10]. Law's basic component are norms, whereas values constitute the core concept of moral philosophy. Values serve as criteria to guide the evaluation of actions taking into account the relative priority of values [15]. Moreover, the abstract nature of values distinguishes them from norms, which usually refer to specific situations [38]. Thus, in this paper, we will assume that: i) the society has preferences over moral values; and ii) a norm is related to a given moral value when the norm's goal either promotes or demotes that value.

In order to reason about norm systems based on moral value preferences, we must be able to compare them in terms of the values that they support. The principle that we adhere to is: *the more preferred the values supported by a norm system, the more preferred that norm system.* Thus, ideally, the decision maker would like to opt for the norm system that supports the most preferred values out of all the sound norm systems. This section is devoted to extend the multi-objective decision-making problem introduced in section 3 to account for the moral values supported by norms. But first, subsequent subsections introduce values and how they relate to norms.

4.1 Value systems

As aforementioned, moral values (e.g., equality or transparency[2]) serve as criteria to guide the evaluation of actions taking into account their relative priority [15]. Since our main goal is to be able to quantitatively reason about norm systems based on the qualitative preferences over the values that they support, we adapt some value-related definitions by Bench-Capon et al. [7] and consider V to be a non-empty set of shared moral values in a society. Furthermore, we consider a *value system*, similar to the one in [28], consisting of a set of values V and a partial order $\succeq$ encapsulating value preferences.

Def. 5. *A value system is a pair $VS = \langle V, \succeq \rangle$, where V stands for a non-empty set of moral values and $\succeq$ represents a partial order over these values such that $\forall v_i, v_j \in V$ s.t. $v_i \succeq v_j$, it implies that v_i is more or equally preferred to v_j.*

Notice that we can have equally preferred values (when $v_i \succeq v_j$ and $v_j \succeq v_i$) as well as incomparable values, which are those for which neither exists a relation between them nor a path of relations that connects them. It is also worth noticing that this ordering extends the total order in [28].

We consider value preferences are expressed in terms of a *partial order*, therefore, incomparable values or equally preferred values may exist. Hence, we assume a regulation authority will reflect both the moral values shared within a society and their relative preferences into a mixed acyclic graph where: nodes represent values; directed edges mean that the starting node value is more preferred than the end node value; and undirected edges connect equally preferred values. Not connected nodes –neither with an edge nor with a path– represent incomparable values (namely, the regulation authority cannot assess which one is preferred or if they are equally preferred).

For these preferences to be congruent (without inconsistencies), we require for equally preferred values in the graph to have the same preferences over other values. Formally, this preference graph will be transformed to a DAG (Directed Acyclic Graph), as in [14], where edges represent preferences of starting nodes over ending nodes. This transformation ensures that there are no ties of preference over sets of values. In this manner, the undirected edges in the graph are changed to directed edges by randomly (with a 50% / 50% chance) assigning a direction in order to maintain a fair equilibrium between equally preferred values. Therefore, pairs of equally preferred values are formally represented by each value being equiprobably more preferred than the other. Next example illustrates this transformation.

Example 4.1. Figure 2 depicts an example of the transformation of the 'raw' preferences into those formal preferences the system will use. Initially, the regulation authority provides a mixed acyclic graph (see Figure 2 a)). In this case, it represents the set of values

[2] Please refer to [38] for an overview on ten basic values that are recognized in cultures around the world.

$V = \{v_1, v_2, v_3, v_4\}$ with preferences: v_1 preferred over v_3 (i.e., $v_1 \succeq v_3$) and v_1 equally preferred to v_4 (i.e., $v_1 \succeq v_4$) and $v_4 \succeq v_1$). Afterwards (see Figure 2 b)), for this graph to be congruent, v_4 has to be preferred over v_3 (i.e., $v_4 \succeq v_3$) since v_4 is equally preferred to v_1, which is more preferred than v_3. Finally, in order to convert the congruent mixed acyclic graph into a DAG, we have to convert the undirected edge to a directed one, by randomly selecting a direction. Thus, we randomly define $v_1 \succeq v_4$ in Figure 2 c).

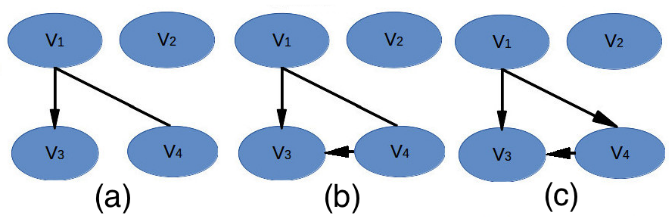

Figure 2: Value preference specification: a) Mixed acyclic graph (MAG); b) Congruent MAG; and c) DAG.

Notice that we only perform this transformation when 'raw' value preferences are ill-defined. Thus, there is no need to perform such processing for congruent DAGs. Either way, equation 8 defines the children relationship to denote immediate preference in a congruent DAG.

$$children(v_i) = \{v_j \mid v_i \succeq v_j \text{ , } \nexists v_k \text{ s.t. } v_i \succeq v_k \succeq v_j\} \tag{8}$$

4.2 Norms' value support

In addition to considering values and their preferences, we need to define how they relate to norms. As previously mentioned, Philosophy literature states that the abstract nature of values distinguishes them from norms, which usually refer to specific situations [38]. Thus, we consider that norms regulate specific situations by promoting (or demoting) some moral values. For instance, consider the specific situation of a funeral, and a norm (or convention) that dictates to wear dark[3]. In this case, this norm promotes the value of respect. From this consideration, we take inspiration in [6] and define the support rate function $val : \mathcal{N} \to [-1, 1]^{|V|}$ which attributes the support of the norm to each value by means of a rate in $[-1, 1]$, where: rate -1 stands for a norm totally demoting a value; 1 means that the norm totally promotes that value; and 0 means that the norm is neutral to the value (that is, it has no relation to it). Thus, $val(n_i) = (x_1, \ldots, x_{|V|})$, encompasses the rate of support of n_i to all values in V.

Using the DAG value structure, we define the following recursive utility function to calculate the preference utility of each value:

$$u(v_i) = \epsilon_i + \sum_{v_j \in children(v_i)} u(v_j) \tag{9}$$

where $1 \leq i \leq |V|$ and $\epsilon_i \in (0, 1]$ is a randomly selected number for each v_i. The randomness of ϵ ensures (with high probability) that there will be no ties between possible sets of values. Moreover, as indicated by equation 8, v_j values correspond to v_i's children (i.e., its immediately less preferred values).

In this case, it is clear that $v_i \succeq v_j \Rightarrow u(v_i) \geq u(v_j)$. Nevertheless, if v_i and v_j are unrelated, although we can ensure (with high probability) that one utility will be greater than the other (no ties), it is a matter of randomness which one will be the largest. Note that this methodology allows us to assure that our solution will be optimal but, since we are choosing ϵ randomly, not all optimal solutions are equiprobable.

Independently of the order used, we can readily calculate the *value support* of a norm n_i by adding the utility of the values supported by the norm as follows:

$$u_n(n_i) = val(n_i)(u(v_1), \ldots, u(v_{|V|}))^{\mathsf{T}} \tag{10}$$

And from the individual utilities of norms we can compute the *value support* for a given norm system $\Omega \subseteq N$ by adding the utility of the values supported by each one of its norms as:

$$u_N(\Omega) = \sum_{n \in \Omega} u_n(n) \tag{11}$$

It is worth noticing that utility function u_N allows us to lift the preferences defined as an order of preferences over single moral values to a preference relation over bundles of norms. Thus, we will say that $\Omega \succeq \Omega' \Leftrightarrow u_N(\Omega) > u_N(\Omega')$. Interestingly, the lifting of preferences provided by the u_N utility function satisfies *responsiveness* [5] which informally states that if in a norm system $\{n_2, n_3\}$, n_3 is replaced by a *better* (supporting more preferred values) norm, e.g. n_1, then $\{n_2, n_1\}$ makes a better norm system.

Lemma 2. *The utility function u_N guarantees responsiveness.*

Proof. 2. *To prove the lemma it suffices to show that given a norm system Ω such that $n_i \in \Omega$, $n_j \notin \Omega$, and $n_j \succeq n_i$, then $\Omega \setminus \{n_i\} \cup \{n_j\} \succeq \Omega$. Let us note $\Omega_{-i} = \Omega \setminus \{n_i\}$. Since $u_N(\Omega) = u_N(\Omega_{-i}) + u_n(n_i) < u_N(\Omega_{-i}) + u_n(n_j) = u_N(\Omega_{-i} \cup \{n_j\})$, then $\Omega \setminus \{n_i\} \cup \{n_j\} \succeq \Omega$ holds.*

4.3 Value optimisation

At this point, we can quantitatively compare norm systems based on the values that they support. Hence, we are ready to define a new multi-objective optimisation problem involving values as an extension of problem 1.

Problem 2. *Given a Norm Net NN= $\langle N, R\rangle$, a representation power function r, a fixed budget b, and value system $(V, \succeq)$, the value-based maximum norm system problem with limited budget (or value-based norm optimisation problem, for short) is the problem of finding a sound norm system $\Omega \subseteq N$ with maximum representation power, minimum cost limited by some non-negative budget b, and maximum value support.*

This problem can be encoded as a linear program by extending the one in section 3. But first, we require: (i) some prioritisation weights w_r, w_c, w_v that measure, respectively, the relative importance of maximizing representation, minimizing cost, and maximising value support; and (ii) normalisation constants. We normalise moral values by considering $\mathcal{V}_{max} = \sum_{i=1}^{|N|} u_n(n_i)$ and normalise representation values by means of $\mathcal{R}_{max} = \sum_{n_j \in G_N} r(n_j)$, where G_N stands for the set of norms that are not directly generalised by

[3]Notice that different cultures differ in their dressing conventions.

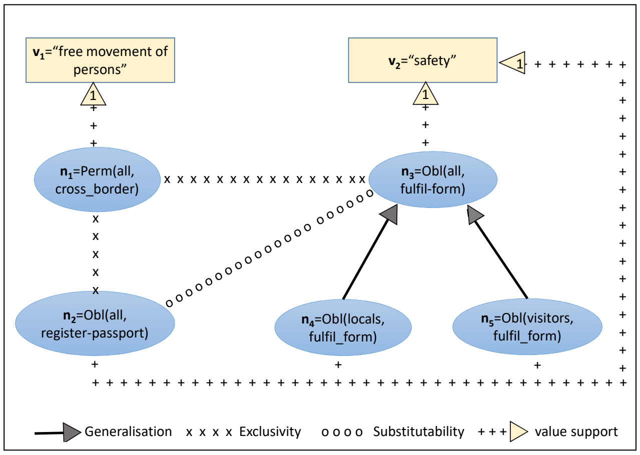

Figure 3: Example of rules of border control ($n_1, \ldots, n_5$) together with the values they promote (v_1, v_2). Total promotion is represented as a 1 within value support relation triangles. For simplicity, value demotion is not depicted.

any other norm. Then, solving the value-based norm optimisation problem amounts to solving the following linear program:

$$max \left[\frac{w_r}{\mathcal{R}_{max}} \cdot \sum_{i=1}^{|N|} x_i \cdot r(n_i) + w_c \cdot (y - \frac{1}{b} \sum_{i=1}^{|N|} x_i \cdot c(n_i)) + \right. \\ \left. + \frac{w_v}{\mathcal{V}_{max}} \cdot \sum_{i=1}^{|N|} x_i \cdot u_n(n_i) \right] \quad (12)$$

subject to constraints from equations 1 to 7 together with the following constraint related with prioritisation weights:

$$w_r + w_c + w_v = 1 \quad w_r, w_c, w_v \in [0, 1] \quad (13)$$

and two additional constraints related with y: a binary indicator variable that allows us to turn the cost minimisation into a maximisation, since finding the norm system with minimum (normalised) cost ($\frac{1}{b} \sum_{i=1}^{|N|} x_i \cdot c(n_i)$) amounts to maximising expression $y - \frac{1}{b} \sum_{i=1}^{|N|} x_i \cdot c(n_i)$. Hence, y must satisfy that:

$$y \in \{0, 1\} \quad (14)$$

$$y \leq \sum_{i=1}^{|N|} x_i \leq M \cdot y \quad (15)$$

where M is a very large number[4]. Furthermore, it is worth mentioning that this indicator variable guarantees that no cost is added to the objective function if no norm is chosen.

Finally, notice that the specification above corresponds to a maximization problem whose constraints are all inequalities. Hence, it is in standard form and it can be solved with state-of-the-art linear program solvers such as CPLEX [20] or Gurobi [35].

Example 4.2. In our example, as Figure 3 shows, we just consider two values: v_1, which corresponds to "free movement of persons"; and v_2, which stands for "safety". Then, we can consider that n_1 totally promotes the "free movement of persons" value but totally demotes "safety" (that is, $val(n_1) = (1, -1)$); whereas $n_2, \ldots, n_5$ totally support the "safety" value whist totally demoting "free movement" (that is, $val(n_2) = \ldots = val(n_5) = (-1, 1)$). Let us consider that: i) the society prefers "free movement" to "safety" (namely $v_1 \succ v_2$), then, $u(v_2) = \epsilon_2$ and $u(v_1) = \epsilon_1 + \epsilon_2$, and ii) moral values are the only criterion to consider (i.e., $w_r = w_c = 0$ and $w_v = 1$). Therefore, our problem amounts to finding the sound norm system that has maximum value support. Then, if we encode the problem, a linear program solver results in two alternative solutions $\Omega = \{n_1, n_4\}$, $\Omega = \{n_1, n_5\}$. In other words, they constitute two different value-optimal sound norm systems.

5 EMPIRICAL HARDNESS ANALYSIS: PROBLEM STRUCTURE

Before analysing the hardness of our problem, we decided to start by empirically studying the effect of the number of norms, their relationships, and available budgets without considering values yet. Thus, next subsection is devoted to describe how experiments on the maximum norm system problem (described in Section 3) are generated. Subsequent subsection presents the obtained results.

5.1 Experimental settings

Initially, we aimed at checking if the norm selection process scales well with the number of norms. With this aim, we conducted experiments for $|N| = 500, 1000, 5000$ norms. Moreover, for each of this predefined number of norms, we generated different experiments varying: the number of exclusivity relations; generalisation relations; and the budget.

Norm relations define the norm net's topology. On the one hand, in order to study generalisation relationships, we automatically generated several generalisation trees for each experiment so that each overall norm net had a forest structure. Different experiments were set up by changing the number of generalisation siblings. Thus, we defined our topologies based on two parameters: height (h), the tree's depth; and width (w), the branching factor. Figure 4 shows different generalisation trees for different heights and widths, including single norm trees —i.e., those trees with $h = 0$ and $w = 0$.

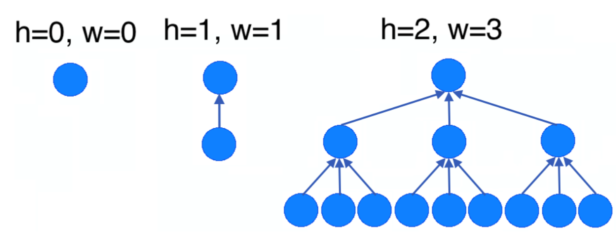

Figure 4: Example of alternative generalisation tree structures together with their heights and widths.

With the aim of avoiding uniform forest structures –i.e., having all trees with the same height h and width w–, we introduce single norm trees in the norm net by considering a g parameter representing the probability of each tree actually having the h and w from its norm net configuration (and thus, $1 - g$ is its probability of being a single norm tree). Considering g requires to compute the number

[4] In our problem M can be defined to be strictly larger than $|N|$.

of trees to generate in the norm net so that its number of norms is close to $|N|$. Our experiments consider $g = 0.5$ and the following formula:

$$trees = round\left(\frac{|N|}{g \cdot ((\sum_{i=0}^{h} w^i) - 1) + 1}\right)$$

Example 5.1. In order to generate a norm net with about 5000 norms, $h = 2$, $w = 3$, and $g = 0.5$, this formula tells us to create 714 trees. From these, we expect 50% of them (that is, 357) to correspond to single norm trees whereas the remaining 357 will have 13 norms each (as in the right-hand side of Figure 4), which amounts for 4641 norms. Hence, the total number of norms expected will be around 4998, which is close to our targeted 5000 norms.

Notice that when generating norm nets with $h = 0$ and $w = 0$ we are just creating $|N|$ single norm trees, without generalisation relations ($R_g = \emptyset$). On the other hand, we have considered 3 different tiers on the generation of exclusivity relations between norms: low exclusivity; medium exclusivity; high exclusivity.

Finally, once we have both generalisation and exclusivity relations, we parameterise the maximum budget to study its possible effects on our problem resolution. First, we assign random individual costs to leaf norms in our trees and then compute the cost of each general norm as the addition of the costs of its children norms:

$$c(n_i) = \sum_{n_j \mid (n_i, n_j) \in R_g} c(n_j)$$

Second, since root norms are the most expensive norms that could be selected, we take as a reference a maximal cost that adds the costs of all root norms (without considering exclusivity relations) in our norm net. Third, we define three alternative budgets: low budget, which is set to be 25% of this maximal cost; medium budget, which corresponds to 50%; and high budget, which is set to 75%.

So to sum up, we have considered the following parameters:

- $|N| = 500, 1000, 5000$
- $h = 0, 1, 2, 3$, $w = 0, 1, 2, 3$
- Low exclusivity, medium exclusivity, high exclusivity
- $budget$ = 25%, 50%, 75% of maximal cost
- $g = 0.5$, $w_r = \frac{1}{2}$, $w_c = \frac{1}{2}$, $w_v = 0$.

We consider representation power as important as cost. Moreover, notice that height and width configurations result in 10 different combinations as configurations of the form $(0, w)$ or $(h, 0)$ are effectively the same as (0,0). Now, for each possible parameter configuration, we performed 100 experiments, which accounts for a total of 27,000 experiments. Next subsection analyses the results using a PC with processor Intel Core i5-6500.

5.2 Results and analysis

As aforementioned, we conducted 27,000 experiments to analyse which problem parameters have a larger impact on resolution performance. Along this empirical evaluation, it became apparent that the larger height (h) and width (w) we considered in an experiment, the smaller computational times we got. This lead us to believe that generalisation relations might make the problem easier. However, the way we generate exclusivities make that given a norm net with a fixed $|N|$ and exclusivity tier, if we reduce the generalisation relations, the exclusivity relations are increased. Therefore, we analyse

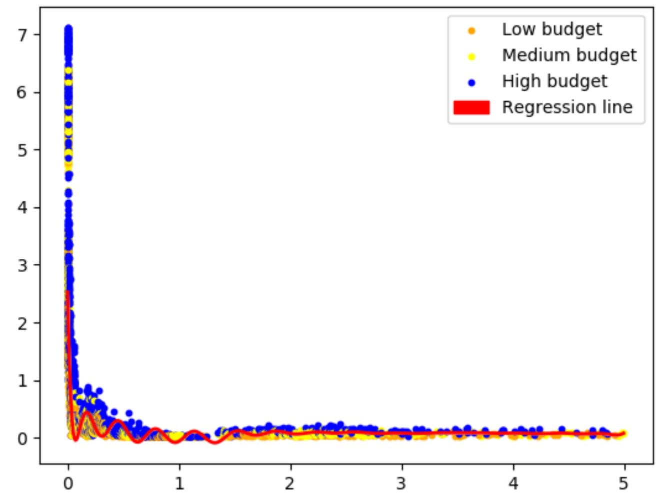

Figure 5: Solving times in seconds (y-axis) for problem instances with 1,000 norms. X-axis corresponds to the *GER*(*NN*) ratio. Each point represents a problem instance. The (red) line depicts a regression on solving times.

the results by considering the ratio of the number of generalisation relations over the number of exclusivity relations. Formally, we define this ratio for a norm net as: $GER(NN) = \frac{|R_g|}{|R_x|}$.

We have seen that this characteristic affects largely the computational time and therefore the problem's hardness. Thus, for example, Figure 5 shows the solving times for 9000 different problem settings with $|N| = 1000$. We observe that solving times rapidly increases as the ratio decreases. In fact, for ratios near zero, tiny changes largely increase hardness. Conversely, solving times get close to zero for larger ratios, without big differences between them. This is caused by a well-defined transition point, separating easy problems corresponding to large ratios from harder problems corresponding to low ratios. The reasons explaining these results are twofold. On the one hand, generalisation relations ease the problem's resolution because, once a root norm is included in the solution, all norms in its generalisation tree cannot belong to the solution, which decreases the number of norms that the solver needs to inspect. On the other hand, exclusivity relations –which just relate root norms– will push the solver to check on siblings (i.e., generalised norms), thus increasing the norms to inspect and slowing the task. Additionally, the less generalisation relations, the more single norm trees (and the more exclusive relations), making finding the best possible solution a matter of checking far more cases. Hence, this makes the case $h = 0$ and $w = 0$ the hardest case to solve.

Regarding maximum budget, Figure 5 also depicts in different colours experiments with low, medium, and high budget. From these results we consider that increasing the budget just affects the time slightly as, the more budget, the more norm combinations are possible solutions, and, therefore, the solver has a harder time selecting them. But it does not represent a substantial difference.

Tests have also shown that, even very large problems bearing thousands of norms have resulted in manageable times. Indeed, the hardest cases considering 500, 1000 and 5000 norms take an average maximum time of 1.6 s, 7 s, and 280 s (i.e., 4'40") respectively.

6 EMPIRICAL HARDNESS ANALYSIS: MORAL VALUES

Next, we study the problem's hardness once we add moral values.

Empirical settings. Here study how values affect performance by focusing on the hardest tree structure from previous section: single norm trees ($h = 0$, $w = 0$) without generalisation relations. For the sake of simplicity (and clarity), all support rates were set to 1. Moreover, we have considered the value cardinality $|V| = \frac{1}{33}|N|$, which we have empirically found to be the hardest.

As for social value preferences, we generated mixed acyclic graphs by assigning, for all possible value pairs, a relation that is randomly chosen from these four possibilities: 1st value is preferred over the 2nd; 2nd value is preferred over the 1st; both values are equally preferred; and values are not related. We then proceeded as described in sections 4.1 and 4.2, producing directed acyclic graphs, values' utilities, and computing the value support of each norm.

Our experiments' objectives are twofold. First, we aim to assess whether or not *norm polarisation* towards certain values affects solvers' performance. Hence, we related norms and values by following four different distributions: i) random distribution; ii) 20%-80%, meaning that 20% of norms support the most desirable values whilst the remaining 80% support the least desirable ones; iii) 50%-50%; and iv) 80%-20%. Second, we aim to study if our problem becomes harder as we increase the variability on the number of relations between norms and values. For this purpose, we generate scenarios with i) low variability, so that all norms relate to few values; ii) medium; or iii) high variability.

In all cases, we considered the weights on equation (13) to be $w_r = w_c = w_v = 0.333$ so as to consider representation power, cost, and value support equally important.

Results and analysis. Table 1 shows the median computation time for each experiment configuration, and compares it to the computation time required with no values (see first row). The second row shows results on problems with randomly generated value relations, increasing time up to 20%. If we bias value support (rows 3-5), the problem tends to become easier, since when there are norms supporting less preferred values, their utilities become lower and they can be discarded. In fact, for large norm nets ($|N| = 5,000$) having more than 50% of norms supporting less preferred values, the resulting times are lower than the problem without values. Regarding rows 6-8, the results improve the times of randomly generated value relations, as although also being generated randomly, in this case variability guides the solver. Thus, for large norm nets, the higher the variability, the higher the improvement.

Overall, we conclude that if the problem's characteristics lead to similar value supports, then the problem will become harder. Otherwise, clearly different value supports ease the problem with respect to those without values, as norms having the highest value supports will tent to be chosen for the resulting norm system.

From these experiments we may also conclude that adding values to the problem does not make the problem much harder. For smaller norm nets, there is barely a noticeable difference on computational times, while for larger cases we have seen that this times can be slightly affected (for good or bad) depending on the values the norms are supporting.

Experiment	500 norms	1,000 norms	5,000 norms
No values	1.67s	6.28s	199.39s
Random values	1.67s	6.82s	227.27s
50%-50%	1.67s	6.36s	192.15s
20%-80%	1.67 s	6.37s	190.91s
80%-20%	1.68s	6.28s	221.99s
Low variability	1.59s	6.17s	226.83s
Medium variability	1.59s	6.2s	216.37s
High variability	1.59s	6.14s	204.97s

Table 1: Median of the computational times for experiment settings with $|N| = 500, 1000, 5000$.

7 CONCLUSIONS AND FUTURE WORK

Parlement et Citoyens [16] is an official web site that enables French citizens to participate in law making. As for November 2017, up to 11 consultations have been completed, leading to a total of 1,411 article propositions that may be submitted to the National Assembly. We argue that our model constitutes a useful decision support system for the complex task of deciding which propositions to approve. We have proven that our model can deal with the expected amount of norms, plus it could solve potential mutually exclusive norm proposals (such as articles regulating open data and those for data privacy) or could eliminate redundancy given similar (generalised) norm proposals. Furthermore, norm implementation costs could also be considered when facing a limited government's budget.

Most importantly, values could also be included explicitly, reflecting shared value preferences or those that can be drawn from the political program of the government, which in turn was selected democratically. Overall, our solver would produce the best possible sound norm system constrained by budget and aligned with the society's value preferences. This would ultimately ease the government's task and encourage citizen participation.

To conclude, we advance the state of the art in norm decision-making by incorporating ethics. Thus, the problem becomes that of choosing the "right norms" by considering, among other criteria, the moral values that candidate norms promote. A problem specification requires the definition of shared moral values, together with the social preferences over these values and the extent to which candidate norms promote them. Then, problem resolution amounts to computing a maximal sound norm system aligned with budget limitations and social moral preferences. We show that CPLEX (a state-of-the-art LP solver) helps solve large problems (in a few seconds when considering up to 1000 norms and at most 4 minutes when considering up to 5000 norms).

As to future work, we plan to perform the automated discovery of norm relationships as well as to further investigating and formalising the concept of value from a philosophical perspective.

REFERENCES

[1] T. Agotnes and M. Wooldridge. 2010. Optimal Social Laws. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 667–674.

[2] Giulia Andrighetto, Guido Governatori, Pablo Noriega, and Leendert WN van der Torre. 2013. *Normative multi-agent systems*. Vol. 4. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[3] IEEE Standards Association. 2016. The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems. http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html. (2016).

[4] Katie Atkinson, Trevor J. M. Bench-Capon, and Peter McBurney. 2006. PARMENIDES: Facilitating Deliberation in Democracies. *Artificial Intelligence and Law* 14 (2006), 261–275.

[5] Salvador Barbera, Peter J. Hammond, and Christian Seidl (Eds.). 2004. *Handbook of utility theory*. Kluwer, Boston, Mass. [u.a.].

[6] Trevor Bench-Capon. 2016. Value-based reasoning and norms. *Artificial Intelligence for Justice* (2016), 9–17.

[7] Trevor J. M. Bench-Capon and Katie Atkinson. 2009. Abstract Argumentation and Values. In *Argumentation in Artificial Intelligence*. 45–64. http://dx.doi.org/10.1007/978-0-387-98197-0_3

[8] Guido Boella, Leendert van der Torre, and Harko Verhagen. 2006. Introduction to normative multiagent systems. *Computational & Mathematical Organization Theory* 12, 2-3 (2006), 71–79.

[9] Jordi Campos, Maite López-Sánchez, Maria Salamó, Pedro Avila, and Juan A. Rodríguez-Aguilar. 2013. Robust regulation adaptation in multi-agent systems. *ACM Transactions on Autonomous and Adaptive Systems* 8 (2013), 1–27.

[10] V. Charisi, L. Dennis, M. Fisher, R. Lieck, A. Matthias, M. Slavkovik, J. Sombetzki, A. F. T. Winfield, and R. Yampolskiy. 2017. Towards Moral Autonomous Systems. *ArXiv e-prints* (March 2017). arXiv:cs.AI/1703.04741 http://adsabs.harvard.edu/abs/2017arXiv170304741C

[11] Ajuntament de Barcelona. 2016. Decidim Barcelona. https://decidim.barcelona. (2016).

[12] Ayuntamiento de Madrid. 2017. Decide Madrid. https://decide.madrid.es/. (2017).

[13] Charles DeTar. 2013. *InterTwinkles: Online Tools for Non-Hierarchical, Consensus-Oriented Decision Making*. Ph.D. Dissertation. Media Arts and Sciences at the Massachusetts Institute of Technology.

[14] Emanuele Di Rosa and Enrico Giunchiglia. 2013. Combining approaches for solving satisfiability problems with qualitative preferences. *AI Communications* 26, 4 (2013), 395–408.

[15] Virginia Dignum. 2017. Responsible Autonomy. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*. 4698–4704.

[16] Parlement et Citoyens. 2017. Parlement et Citoyens: institutional participation portal for the discussion of French laws. https://parlement-et-citoyens.fr/. (2017).

[17] David Fitoussi and Moshe Tennenholtz. 2000. Choosing social laws for multi-agent systems: Minimality and simplicity. *Artificial Intelligence* 119, 1-2 (2000), 61–101.

[18] N. Griffiths and M. Luck. 2010. Norm Emergence in Tag-Based Cooperation. In *Proceedings of COIN*. 79–86.

[19] Davide Grossi and Frank Dignum. 2005. From abstract to concrete norms in agent institutions. In *Proceedings of the Third international conference on Formal Approaches to Agent-Based Systems (FAABS'04)*. Springer-Verlag, Berlin, Heidelberg, 12–29.

[20] IBM. 1988. CPLEX. https://www.ibm.com/analytics/data-science/prescriptive-analytics/cplex-optimizer. (1988). Accessed: 2018-02-28.

[21] The AI Initiative. 2015. The AI Initiative, Civic Debate on the Governance of AI. http://ai-initiative.org/. (2015).

[22] Pensions Policy Institute. 2016. Lessons from New Zealand and State Pension Reform: The Consultation Response. http://www.pensionspolicyinstitute.org.uk/press/press-releases/citizens-pension-lessons-from-new-zealand-and-state-pension-reform-the-consultation-response. (2016).

[23] Richard M Karp. 1972. Reducibility among combinatorial problems. In *Complexity of computer computations*. Springer, 85–103.

[24] T Kohler, J-P Steghoefer, D Busquets, and J Pitt. 2014. The Value of Fairness: Trade-offs in Repeated Dynamic Resource Allocation. IEEE, 1–10.

[25] Martin J. Kollingbaum, Timothy J. Norman, Alun Preece, and Derek Sleeman. 2007. *Norm Conflicts and Inconsistencies in Virtual Organisations*. Springer Berlin Heidelberg, Berlin, Heidelberg, 245–258.

[26] Loomio. 2016. Loomio: software to assist collaborative decision-making processes. https://www.loomio.org/. (2016).

[27] Maite Lopez-Sanchez, Marc Serramia, Juan A. Rodríguez-Aguilar, Javier Morales, and Michael Wooldridge. 2017. Automating decision making to help establish norm-based regulations. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. 1613–1615.

[28] Jieting Luo, John-Jules Meyer, and Max Knobbout. 2017. Reasoning about Opportunistic Propensity in Multi-agent Systems. In *AAMAS 2017 Workshops, Best Papers*. 1–16.

[29] New Zealand's ministry of Justice. 2016. Better family violence law. https://consultations.justice.govt.nz/policy/family-violence-law/. (2016).

[30] S. Modgil. 2006. Value Based Argumentation in Hierarchical Argumentation Frameworks. In *Proceedings of the 2006 Conference on Computational Models of Argument: Proceedings of COMMA 2006*. IOS Press, Amsterdam, The Netherlands, 297–308.

[31] Javier Morales, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, Wamberto Vasconcelos, and Michael Wooldridge. 2015. On-line Automated Synthesis of Compact Normative Systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 10, 1 (March 2015), 2:1–2:33.

[32] Javier Morales, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, Michael Wooldridge, and Wamberto Vasconcelos. 2014. Minimality and Simplicity in the On-line Automated Synthesis of Normative Systems. In *AAMAS 2014*. IFAAMAS, Richland, SC, 109–116.

[33] Javier Morales, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, Michael Wooldridge, and Wamberto Vasconcelos. 2015. Synthesising Liberal Normative Systems. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS '15)*. 433–441.

[34] City of Reykjavík. 2016. City of Reykjavík participation portal. http://reykjavik.is/en/participation. (2016).

[35] Gurobi Optimization. 2010. Gurobi Optimizer. http://www.gurobi.com/. (2010). Accessed: 2018-02-28.

[36] R Riveret, EG Nepomuceno, J Pitt, and A Artikis. 2014. Self-Governance by Transfiguration: From Learning to Prescription Changes. IEEE, 70–79.

[37] BastinTonyRoy Savarimuthu, Stephen Cranefield, MaryamA. Purvis, and MartinK. Purvis. 2013. Identifying prohibition norms in agent societies. *Artificial Intelligence and Law* 21, 1 (2013), 1–46. https://doi.org/10.1007/s10506-012-9126-7

[38] Shalom Schwartz. 2006. An Overview Basic Human Values: Theory, Methods, and Applications Introduction to the Values Theory. *Jerusalem Hebrew University* (2006).

[39] Rajiv Sethi and Eswaran Somanathan. 1996. The evolution of social norms in common property resource use. *The American Economic Review* (1996), 766–788.

[40] Y. Shoham and M. Tennenholtz. 1995. On social laws for artificial agent societies: off-line design. *Artificial Intelligence* 73, 1-2 (February 1995), 231–252.

[41] Wamberto Weber Vasconcelos, Martin J. Kollingbaum, and Timothy J. Norman. 2009. Normative conflict resolution in multi-agent systems. *Autonomous Agents and Multi-Agent Systems* 19, 2 (2009), 124–152.

[42] D. Villatoro, J. Sabater-Mir, and S. Sen. 2011. Social Instruments for Robust Convention Emergence.. In *IJCAI*. 420–425.

[43] Vishanth Weerakkody and Christopher G Reddick. 2012. *Public sector transformation through e-government: experiences from Europe and North America*. Routledge.

[44] Colin Allen Wendell Wallach. 2008. *Moral machines: teaching robots right from wrong*. Oxford University press.

[45] Fabiola López y López, Michael Luck, and Mark d'Inverno. 2002. Constraining autonomy through norms. In *AAMAS*. ACM, 674–681.