Value-Guided Synthesis of Parametric Normative Systems

Nieves Montes
Artificial Intelligence Research Institute (IIIA-CSIC)
Barcelona
nmontes@iiia.csic.es

Carles Sierra
Artificial Intelligence Research Institute (IIIA-CSIC)
Barcelona
sierra@iiia.csic.es

ABSTRACT

In recent years there has been an increasing interest in ensuring that autonomous systems behave consistently with human values. A popular approach to this challenge is through the incorporation of norms that regulate behaviour in an ethical way. However, such norms must be effective at promoting the values we consider most important. In this work, we introduce a systematic methodology for the automated synthesis of parametric normative systems based on value promotion. We introduce the new concepts of *Shapley values of norms* and *value compatibility*. To quantify the effectiveness of norms at upholding the values we consider relevant, we adopt the value alignment indicator from a previously established framework. We apply our model to a toy system which we use to illustrate our approach from end to end.

KEYWORDS

Normative Multiagent Systems; Norm Synthesis; Value Alignment; Shapley Value

ACM Reference Format:
Nieves Montes and Carles Sierra. 2021. Value-Guided Synthesis of Parametric Normative Systems. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 9 pages.

1 INTRODUCTION

In recent times, research in Artificial Intelligence has led to giant leaps forward in the field of autonomous intelligent systems. With such advances, however, new concerns arise related to our ability to control such systems and to ensure that they behave in ways that are ethically consistent with our shared human values. That consistency (or lack thereof) is usually labelled as the *value alignment problem* [17, 23].

Some previous work in the field focuses on equipping participating agents with internal reasoning schemes guided by individual values [3, 5, 6]. In practice, however, this type of approach requires access to the internal workings of the participants. In general, this is not possible in shared social spaces where the intentions of participating agents are not public knowledge. To overcome this shortcoming, another line of research targets the introduction of norms and regulations that enforce or ban specific agent actions, and impose sanctions upon detecting violations [1, 8].

Nonetheless, a major issue comes up when considering this approach: how to carefully design norms that promote our esteemed values without incurring in undesired consequences. In this work, we take the view that norms have to be *provably aligned* with our values [7, 17] and rely on an established value alignment model [22] to quantify how effective is a given set of norms, aka a *normative system*, at promoting some values of choice. Such metric allows us to proceed with an optimisation search over the norm instance space with an assessment of the value alignment degree as the target function to maximise.

1.1 Related work and contributions

There is abundant literature on the automated synthesis of norms in multiagent systems. There are two main approaches to tackle this task: off-line and on-line design. Early research on the design of norms (often referred to as *social laws*) for technologically enabled agents focused on the off-line situation [15, 21]. Norms are conceived as restrictions (essentially prohibitions) on the actions that can be taken under some specified circumstances, while still allowing agents with sufficient margin to achieve their individual goals. While the general problem is proven to be intractable, certain reasonable restrictions (such as state factorisation), allow significant complexity reduction.

Off-line design is suitable for applications where knowledge on the composition of the system is readily available prior to deployment. However, off-line design is unable to successfully regulate open multiagent systems, where participants might enter and leave at any point during run-time, and the norms might need to adapt to the current make-up of the society.

On-line norm synthesis tries to overcome these limitations by adopting regulations tailored to the current circumstances during run-time [9]. One approach to this problem is bottom-up norm emergence, where agents collaborate to agree on the regulations to adopt [13]. Other solutions that do not assume agent collaboration include running a “norm engine” concurrently with the system to be regulated [12]. This engine continuously monitors the effectiveness of the current normative system and updates it as conflicts are detected.

Despite the extensive research on on-line and off-line norm design, the incorporation of moral values into the process is a much more recent concern. Previous work in [19] does take into account preferences over the moral values supported by a society and encodes them as a utility function to be maximised. Their approach to the problem is somewhat opposite of ours. Given a set of specified norms, they seek to find the best norm subset based on various criteria. In contrast, we proceed by pre-defining a family of parametric norms to work with and then looking for the parameter values for each norm that work best.

In this work, we intend to enrich the current research landscape on the synthesis of norms that take into account the moral values entailed by them. The primary purpose is to provide a systematic methodology to craft norms that maximally align with a value or

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.
© 2021 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

set of values. Mathematically, we seek to compute the following:

$$N^* = \underset{N' \subseteq N}{\arg\max}\, \mathsf{Algn}^{G}_{N',V} \tag{1}$$

where N is the set of normative systems, V is the set of values of interest, G is the set of agents in the society and $\mathsf{Algn}^{G}_{N,V}$ is the degree of alignment of norms N with respect to values V for agent set G. Our approach is fundamentally off-line (like that of [19]) as the compliance of norms with values is evaluated through their effect on a simulated society. Each of the norms is related to a single or a set of parameters, and value compliance is the guiding principle for the search of optimal parameter values. We adopt a formal model of value alignment in normative multiagent systems crafted to computationally evaluate the alignment of a set of given norms for a set of values [22]. We leverage this model to work in a different direction, *i.e.* given a set of values, obtain the norms that promote them the most.

Our approach is comprised of the following general steps:

(1) Provide the domain of the multiagent system and define the system's state space $\mathcal{S}$.
(2) Define the normative system N that governs the systems in terms of individual component norms n_i. In turn, define the parameters P_i upon which such individual norms n_i are dependent.
(3) Provide a mathematical expression for the alignment of the norms in N for every value that the system should promote. The previous parametrisation of norms together with the definition of a target alignment function provide an avenue to optimise P_i with respect to the values of interest.
(4) Choose a suitable optimisation algorithm and perform a maximising search through the normative parameters space with the previously defined alignment as the target function.

Beyond obtaining ethically optimal normative systems, we are also interested in tools that allow examining them at a more detailed level. The other important contributions of this paper in that regard are:

- The concept of *Shapley value for a single norm*, which is based on understanding normative systems as coalitions of individual norms and is derived from the Shapley value in cooperative game theory. With this concept, we can evaluate norms on an individual basis within the normative system.
- The concept of *compatibility* of values under a given normative system, which captures the dynamic relationships between different values and quantifies to what extent can a normative system be in simultaneous compliance with them.

Running example

We illustrate all of our contributions with a running example of a toy social model. In this society, agents pay an amount that is dependent upon their current wealth. To emulate the daily struggle of fiscal authorities, we introduce a subset of evader agents who systematically seek to avoid tax payment. However, not all hope is lost, as there is a non-zero probability of the authorities detecting any evader and imposing a fine that is larger than the original tax payment.

The collected taxes plus any claimed fines are merged into a common fund and invested. Consequently, they grow by a fixed amount. Then, the common fund (the original taxes and fines plus the profits) is redistributed back to the agents, also according to their economic status. This process repeats at every time-step, and the evolution of the system is monitored. Within this system, we model the values *equality* and *fairness* by defining alignment functions and obtaining optimal normative systems for the two cases.

This paper is structured as follows. First, we lay out the framework for the problem in Section 2, covering steps (1) and (2). Then, in Section 3, we develop the task as an optimisation problem, which correspond to steps (3) and (4). Sections 4 and 5 introduce the concepts of Shapley value for a norm and value compatibility, respectively. We conclude with Section 6. We illustrate all sections with our running example.

2 PROBLEM FRAMEWORK

The focus of this work is on normative multiagent systems [2], where a set of participating agents G interact with one another and with the environment. At each time step, state s captures the situation that any agent finds itself in. We denote the set of all possible states with $\mathcal{S}$. The global state $\mathbf{s} = (s_i)_{\forall i \in G}$ is composed of the individual states of all agents in the system. Typically, agents will only have access to their representations s_i, which might even be further limited by incomplete or partial observability. The global state might be partially or totally accessible to an outer entity, typically a regulatory agency keeping track of meaningful indicators.

As a consequence of the interactions between agents, at each time-step agents transition from state s to state s', hence, the global state also changes from $\mathbf{s}$ to $\mathbf{s'}$. The interactions giving rise to such transitions happen as specified by a set of regulatory norms. These provide guidelines and behaviour directives for agents to adhere to or dismiss. In this work, we do not consider any internal reasoning schemes that would allow agents to decide whether to follow or disobey a norm. Note that this is not a defining feature of our methodology, since it can certainly be extended to introduce individual reasoning. However, agent reasoning is not the central point of this work and so we leave it out in order to focus on the main topic of this paper, the design of moral norms. For the time being, let us assume a top-down approach where agents have no role at crafting these norms, and that they are externally created. In principle, these norms intend to steer the system towards more desirable states. Whatever "more desirable" means depends on the standard that we wish to hold the system to.

N denotes the normative system in place. It is a set of typically more than one norm, $\{n_i\}$. We conceive norms as *parametric*, where each n_i is related to a set of parameters P_i. The elements in P_i might have continuous or discrete domains, and be unbounded or constrained. For example, a norm regulating parking is associated with a numerical fine if an agent parks in an illegal spot. In contrast, a norm regulating access to television content is related to a Boolean parameter indicating whether the spectator is underage or not.

Associating every norm n_i with a parameter set P_i is flexible enough to include norms that are not parametric at all ($P_i = \emptyset$), as well as norms associated to a single parameter ($|P_i| = 1$) and norms associated to multiple parameters ($|P_i| > 1$). The set of norms, aka the normative system, is then dependent upon the set of parameters in all P_i sets. We refer to this as the *set of normative parameters*, and denote it by P_N. It formally corresponds to the union $P_N = \cup_i P_i$.

The specific normative system N in charge of regulating the society has all the parameters in P_N associated to a numerical value within the allowed bounds. There is a large number of possible normative systems, potentially infinite if any of the normative parameters has a continuous domain. To denote the set of all possible normative systems made up of the set of individual norms $\{n_i\}$ and tied to the non-instantiated variables in P_i we use the term *normative systems family*.

To tackle the purpose of this work as stated in eq. (1), in practise, we need to define the normative systems family from which we will draw the most suitable norms. To do so, we have to describe qualitatively the role of each norm n_i, their associated parameters in P_i as unassigned variables, their domains, and any constraints they might be subject to. We restrict the search space in eq. (1) to the defined family, *i.e.* the subset of all possible normative systems composed of individual norms $\{n_i\}$ related to parameters P_i. An element in the search space has all the parameters in P_N instantiated and corresponds to a fully defined normative system, which is the construct that will actually regulate the society. It is up to the value-guided search to find the optimal numerical or categorical values of the parameters (see section 3).

Running example

In our simple social model at each time-step we have a set of agents ($|G| = 200$) that pay taxes. Ten of those agents try to evade taxes. They might get caught and made to pay a fine, equal to the evaded taxes plus some additional percentage. Collected taxes and fines are merged. They are then invested with a fixed 5% return rate and are redistributed back to the participants. Agents belong to one of five equally populated groups according to their wealth, the wealthiest agents being part of group #5 and the poorest agents part of group #1. The collection of taxes and redistribution of money for every agent is dependent upon their wealth group. Agents' wealth is initialised randomly according to a uniform distribution $U(0, 100)$.

At every state, s, the absolute wealth x_i and wealth group gr_i specify an agent's situation. Thus, the state space is given by $\mathcal{S} = \mathbb{R}^+ \times \{1, 2, 3, 4, 5\}$. For convenience, the model does not allow any member to get into debt, so wealth always stays positive. Payments that would result in negative wealth values are cut down to the maximum amount that the agent can afford.

Within this model, transitions happen under the regulations of a normative system $N = \{n_1, n_2, n_3, n_4\}$. Next is the explanation of every individual norm n_i and their associated parameter sets P_i:

- $\mathbf{n_1}$ is a norm specifying how much each agent must contribute in taxes. It is parametric on the set $P_1 = \{collect_j\}_{j=1,...,5}$, where $collect_j$ is the tax rate that must be paid by agents in the j-th wealth group at each time-step, and whose possible values are bounded between 0 and 1.
- $\mathbf{n_2}$ is a norm specifying how should the invested taxes be redistributed back to the agents. It is parametric on the set $P_2 = \{redistribute_j\}_{j=1,...,5}$, where $redistribute_j$ corresponds to the fraction of the invested taxes to be equally shared by all the members in the j-th wealth group. Its values are bounded between 0 and 1, plus the additional linear constraint:

$$\sum_j redistribute_j = 1 \tag{2}$$

  meaning that the totality of invested taxes is given back to the agents.
- $\mathbf{n_3}$ is a norm specifying how effective the detection of evaders is. It is parametric on a single parameter, $P_3 = \{catch\}$, which denotes the probability of catching an evader during a transition. Assuming the difficulty of law-enforcement tasks, we constrain it between 0 and $\frac{1}{2}$.
- $\mathbf{n_4}$ is a norm specifying the punishment imposed on detected evaders. It is also parametric on a single parameter, $P_4 = \{fine\}$. Whenever an evader is caught, it is obliged to pay a fine equal to the original taxes plus the additional fraction given by *fine*. If the payment of such fine would result in the agent having negative wealth, then the payment corresponds just to the totality of the evader's wealth.

The set of normative parameters defining the normative search space is then $P_N = \{collect_j, redistribute_j, catch, fine\}_{j=1,...,5}$. It should be noted that, despite bearing some resemblance with real-life tax codes, our model is not intended as a reliable substitute for any of them. It is just a proposal to demonstrate our methodology in action.

3 VALUE-GUIDED OPTIMISATION OF NORMATIVE PARAMETERS

So far, we have specified and illustrated how norms regulate a multiagent system. Our goal is then to find instances of P_N parameters such that the evolution of the model under N is effective at promoting some of our most esteemed human values. We quantify such effectiveness through the *alignment* of a set of norms with respect to some value, or set of values.

We take the view adopted in [22], which provides a formal computational model of value alignment. We denote the set of values of interest by V. For each value $v_i \in V$, a preference function is defined over consecutive states. The design of such functions should mathematically capture our understanding of value v_i, *e.g.* by evaluating the fulfilment of the properties relevant to value v_i in the pre- and post-transition states. Preference functions are bounded between +1 and -1 to denote promotion and demotion of such value over a transition.

In [22], the alignment of a normative system N for value v is obtained by averaging preference changes over a sequence of transitions. Due to the possibly stochastic nature of the environment, it is reasonable not to average preferences over a single sequence of transitions, but to perform Monte Carlo sampling over numerous sequences to reduce the variance of the estimated average. For convenience, it is also reasonable to keep the length of the samples sequences fixed. Since the defined range for preferences is set to $[-1, +1]$, the resulting alignment computed by averaging them also falls within this interval. Alignment of +(-)1 denotes that the norms in place are highly (not) compliant towards the value of choice. Beyond being a requirement of the original model, keeping the numerical alignment within a fixed scale is necessary in order to compare alignment of different normative systems with respect to (possibly different) values.

As the normative system N with its associated parameters P_N dictates the transitions $s \rightarrow s'$, changes on the values of the parameters will affect the eventual alignment. The purpose of the search process is then to find which parameter instances lead to the most preferred states concerning value v, according to how we mathematically quantify the value's associated preference function. Consequently, our problem turns into an optimisation task in which a search through the P_N space has to be performed with the alignment Algn of norms N (associated to P_N) with respect to value v as the target function to maximise.

In its most granular form, alignment is computed for a single agent $\alpha \in G$ with respect to a single value v. However, agents have some internal value hierarchy, which allows them to weight the relative importance of the values they hold dear and aggregate over them. Analogously, agents that share their preferences with one another or with a central entity can have them aggregated over the whole group, and hence obtain a quantity representing how well is the society on the whole doing with respect to the value(s).

Our methodology is compatible with any kind of alignment (aggregated or not). Using alignment for a single agent as the target function makes sense from an individual point of view, when a member of the society wishes to know the regulations of its own behaviour that would most promote some esteemed value, or promote them all at once while respecting their hierarchy. From a social perspective, alignments aggregated over the whole agent set are more interesting. However, their computation is possible only if there is some mechanism in place by which agents can share their opinions.

Once we have chosen the values of interest, alongside preference and alignment functions to compute their alignment, the optimisation search is ready. We do not advocate for any particular algorithm to tackle this step. This choice should be based on the specific requirements and characteristics of the problem, namely whether the normative parameters have continuous or discrete domains, and whether the search strategy has to deal with constraints of any kind (linear or non-linear).

A default possibility to perform the search is to implement a Genetic Algorithm (GA) [11]. GAs are very versatile search methods over multidimensional spaces that can be tailored to both discrete and continuous domain variables. In the problem we are dealing with, computing the gradient of the alignment with respect to the normative variables is not, in general, an easy task, and hence gradient descent algorithms [16] are not recommended. In contrast, GAs do not make any requirements concerning the continuity or differentiability of the target function, and so we believe they are the most suitable method to be applied for this undertaking.

In a genetic search applied to our context, a population of candidate normative system instances is maintained, first initialised with random P_N parameters. The "fitness" of candidate solutions equals to the alignment of norms N implemented according to the numerical values of P_N for the candidate solution. Highly promising candidates are selected for recombination in the hopes of obtaining even better solutions. Some form of explorability is introduced through a mutation operation, that is applied to the newly generated candidates. The process is repeated until some stopping criterion is met. Since preference functions, and consequently also alignments, are bounded between -1 and +1, we recommend that the search can be shortened by stopping when a highly aligned normative system has been encountered with Algn ~ 1.

Running example

As advanced in section 1, we are interested in modelling the set of values $V = \{equality, fairness\}$ into our running example. This choice of values exemplifies two goals that are not correlated (as will be shown next) and that are therefore achieved through different taxing strategies.

To compute the alignments, we take a slightly different approach to that of the original proposal [22] in two aspects. First, to compute the alignment, we define the Algn functions with respect to the values directly, and not through preferences over consecutive transitions. In particular, we consider properties of the multiagent system at the final global state $\mathbf{s_f}$ of a state sequence. This approach takes into account how good is the state of affairs in the model eventually and does not worry about how it got there. Nevertheless, we believe that the fundamental premises of [22] still hold: the values and the norms that seek to uphold them are as worthy as the tangible outcomes they generate in the real or virtual world.

Second, we also define the Algn functions directly for the whole set of agents G. That is, we take the view of an external examiner, *e.g.* the tax authority, with complete access to the global state $\mathbf{s}$ who wishes to assess to what degree their regulations are successful at promoting values of general interest. Originally, preference functions were defined at the individual level and then aggregated over the set of agents in G. This deviation from the original proposal is more significant than the first one, as we no longer consider values as inherently individual cognitive constructs, but as tools to help assess the overall performance of specific legislation in action according to the standards of an external evaluator. This is not a feature of the general methodology but of this specific example. As discussed in section 2, the global state might not be accessible. However, in many countries tax agencies routinely monitor income, so in the context of our example we believe that access to the global state $\mathbf{s}$ is justified.

So, overall, by assessing values through the properties of the final global state, we are effectively taking the view of a policy-maker who has defined some targets towards which he would like the society to evolve in the long term. Now, policy and legislation have to be designed, with these standards in mind, to steer the system towards it.

To define our target alignment functions, we introduce the notation $\mathbb{E}_n[\cdot]$, which denotes the expectation of the bracketed statistic under normative systems N. Given a sample of paths that have evolved under the governance of the norms in N, the indicator of choice is computed for each of them and averaged to obtain its expected value.

Alignment with respect to value *equality* is quantified through the Gini Index (GI) [10], a very well-known indicator of widespread use in economics to quantify wealth and income inequality [4]. A set of norms is considered to be in high compliance with equality if it manages to minimise the GI:

$$\mathsf{Algn}^{G}_{N,equality} = \mathbb{E}_N\left[f(\mathrm{S}_f)\right] \tag{3}$$

Table 1: Optimisation results for the running example with respect to the considered values: optimal normative parameters defining the normative system, and their associated optimal alignment.

Value and target function	Optimal normative parameters P_N^*	Optimal alignment $\mathrm{Algn}_{\mathrm{N,v}}^{\mathrm{G}\,*}$
Equality, eq. (3)	*collect* = [20%, 29%, 26%, 35%, 27%]	0.95
	redistribute = [20%, 22%, 19%, 26%, 13%]	
	catch = 44%	
	fine = 61%	
Fairness, eq. (5)	*collect* = [1%, 30%, 37%, 72%, 66%]	0.93
	redistribute = [2%, 23%, 42%, 24%, 9%]	
	catch = 45%	
	fine = 56%	

where the random variable S_f correspond to the global state at the end of a sampled path, and the function f is given by:

$$f(\mathbf{s}) = 1 - 2 \cdot GI(\mathbf{s}), \text{ with } GI(\mathbf{s}) = \frac{\sum_{i,j \in G} |x_i - x_j|}{2 \cdot |G|^2 \cdot \bar{x}} \tag{4}$$

where $GI(\mathbf{s})$ is the Gini Index at the global state, with x_k denoting the wealth of the agents in global state $\mathbf{s}$ and $\bar{x}$ indicating the average. Then, high (low) wealth parity is reflected by a GI close to 0 (+1), and by eq. (4) the resulting alignment with respect to equality is +1 (-1).

High promotion of the value *fairness* is considered to be achieved if, by the end of a random path, as many evaders as possible are, on average, among the poorest individuals in the population:

$$\mathsf{Algn}^G_{N,fairness} = \mathbb{E}_N \left[g(\mathrm{S}_f) \right] \tag{5}$$

where:

$$g(\mathbf{s}) = 2 \cdot \hat{\mathbb{P}}[gr_i(\mathbf{s}) = 1 | evader_i] - 1 \tag{6}$$

and $\hat{\mathbb{P}}[gr_i(\mathbf{s}) = 1 | evader_i]$ denotes the estimated probability that an evader agent belong to the lowest wealth segment at the global state $\mathbf{s}$. It is computed as the proportion of evaders who are in segment #1 at the given global state.

Given that, in our simulated society, there are more agents per wealth group (40) than evaders (10), in the best-case scenario all evaders might end up in wealth group #1. The upper bound for functions $g(\mathbf{s})$, then, is +1. If there were more evaders than agents in each wealth group, this function would need to be modified so that the potential maximum alignment that can be achieved does not fall below +1.

Functions (3)-(6) serve as proxies of the values they represent. Our goal is then to find instances of normative parameters $P_N = \{collect_j, redistribute_j, catch, fine\}_{j=1,\dots,5}$ such that the evolution of the model under the normative system that they implement maximises the alignment functions we have just defined.

We use a Genetic Algorithm for the optimisation search with equations (3)-(6) as target functions. We perform Monte-Carlo sampling of 500 paths of 10 transitions each. In order to adapt the search to the continuous P_N space, the crossover operation is performed by intermediate recombination, which features exploration and hence eliminates the need for a mutation operation [14]. Exploitability is enhanced through the use of elitism, meaning that the best candidates from the previous generation replace the worst candidate from the current one. Additional tweaks to ensure that constraint (2) is fulfilled are introduced.

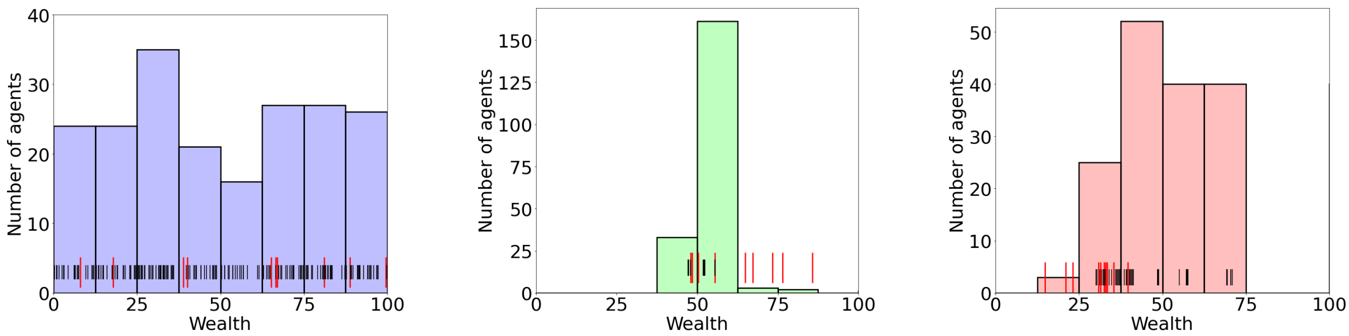
Optimisation results

Table 1 presents the optimisation results for the two values of focus. The optimal alignments obtained are very satisfactory, with large positive values > 0.9.

We provide an intuitive interpretation of the optimal normative parameters. For equality, the differences between *collect* are small across wealth groups. In practise, this means that wealthier agents contribute to the common fund with more resources in absolute terms (as their base wealth is larger). The even redistribution rates across groups then ensure that all agents receive a similar piece of the invested funds. The moderate values for the collecting and redistribution rates in the optimal model with respect to equality correspond to a compromise between funnelling enough resources from rich to poor agents in order to shrink the wealth distribution, but not channelling too many as to swap them, which would then be detrimental towards lowering the Gini Index.

For fairness, the parameters indicate than another mechanism is in place in order to push evaders towards the poorest group. It is worth noting that neither the probability of catching evaders nor the fine they are imposed are particularly large, they are similar or even smaller than those found for the optimal normative system towards equality. Rather, it appears that evaders are pushed towards group #1 by retrieving a lot of resources from the upper groups where undetected evaders manage to sneak, and then redirecting them towards the middle class, which is vastly composed of law-abiding citizens, since detected evaders belong to the lower groups and undetected ones belong to the upper ones. Hence, the norms act by identifying the wealth groups most likely to include cheaters and directing their wealth elsewhere. It does not target group #1, but it keeps the cash flow in and out of that group very limited, so that already poor evaders do not have any avenue to enrich themselves.

To provide a visual representation of the evolution of the society under either optimal normative system, Figure 1 displays the initial wealth distribution (sampled uniformly) alongside the final ones when the optimal normative systems are in place. It is very clear that the optimal norms for both values accomplish what their encoded objective was. The optimal normative system for equality is extremely successful at shrinking the distribution, hence greatly diminishing the Gini Index. Also, evaders, who are initially uniformly distributed within the wealth range, are pushed towards the lower positions after the optimal norms with respect to fairness are enforced.



(a) (b) (c)

Figure 1: Wealth distributions and rug plot indicating the location of law-abiding agents (regular black marks) and evaders (longer red marks): (a) at the initial state; (b) after a random path under the optimal normative system for equality; (c) after a random path under the optimal normative system for fairness.

4 SHAPLEY VALUES OF INDIVIDUAL NORMS

Up to this point, we have been able to synthesise entire normative systems that are highly compliant towards some value. However, we would like to quantitatively assess how big a role each of the individual norms n_i in N plays to achieve this high alignment. Intuitively, not all norms are equally relevant when it comes to promoting a particular value. For instance, in the context of our example, one might wonder whether imposing large fines to detected evaders might actually be detrimental towards ensuring high equality.

To tackle this issue, we take the view of every optimal normative system as a coalition of individual norms working together to achieve high compliance towards their target value. We wish to quantify how relevant each individual norm n_i is at achieving this task. To do so, we adapt the concept of Shapley value from cooperative game theory [20], and define the Shapley value $\phi_i(v)$ of an individual norm n_i within system N with respect to value v as:

$$\phi_i(v) = \sum_{N' \subseteq N \setminus \{n_i\}} \frac{|N'|!\,(|N| - |N'| - 1)!}{|N|!} \cdot \cdot \left(\mathsf{Algn}_{N' \cup \{n_i\}, v} - \mathsf{Algn}_{N', v}\right) \tag{7}$$

where the sum is taken over all normative systems N' not including individual norm n_i.

Eq. (7) displays that the Shapley value of norm n_i is computed through the alignment improvement from its absence, ($\mathsf{Algn}_{N',v}$), to its introduction, ($\mathsf{Algn}_{N' \cup \{n_i\},v}$). However, two questions immediately come up with this approach: (i) which normative systems N' should the sum include? and, (ii) what does it mean for norm n_i to be absent from normative system instance N'? To answer the former, we first need to provide a concise explanation of the latter.

Consider an arbitrary normative system instance N, from which we wish to remove the set of individual norms $\{n_k\}$. We denote the numerical values upon which norm n_i is parametric in the normative system instance N as $P_i^{(N)}$. To proceed with the removal, we first introduce a baseline normative system N_{bsl}. N_{bsl} is an instance of the same norm family as N, *i.e.* it has the same set of norms $\{n_i\}$ related to the same set of normative parameters P_N with the same domains and constraints. However, the numerical values for N_{bsl} are chosen *a priori* in a way as to reflect lack of regulation. Then, to remove norms $\{n_k\}$ from N, we substitute the values of n_k's parameters in the original normative system, $P_k^{(N)}$, by their baseline counterparts, $P_k^{(bsl)}$, for all k. Consequently, the normative parameters of system $N \setminus \{n_k\}$ is composed of $P_i^{(N)}$ for n_i not in the set to be removed, union $P_k^{(bsl)}$ for n_k in the set to be removed.

Now that we know how to remove individual norms from a normative system, we can provide a straightforward answer to the first question. The sum in eq. (7) is taken over $N' \subseteq N \setminus \{n_i\}$, which denotes normative systems from which *at least* the set composed of the single norm n_i has been removed. Hence, to obtain all the terms in the summation, one must first remove n_i from the normative system N, and then substitute other norms' parameters by their baselines according to all possible combinations of the remaining norms $\{n_j\}_{j \neq i}$. Finally, $N' \cup \{n_i\}$ is obtained by setting $P_i^{(bsl)}$ back to the original parameter values $P_i^{(N)}$.

The Shapley value includes factorial terms that refer to the size of the original normative system $|N|$ (which is a fixed quantity across all normative systems in the same family) and to the trimmed one $|N'|$. Note that $|N'|$ only counts the individual norms that have *not* been substituted by baselines, as those that have been are considered as absent.

In this paper, we do not provide a systematic method to find an adequate baseline given a normative system space. For the time being, the simplicity of the example we provide allows us to specify the baseline parameters from mere intuition. However, we can assert that a good choice for a baseline is one such that, for any sampled sequence of transitions, the initial global state is kept unchanged, $\mathbf{s_0} = \mathbf{s}_f$. This requirement is quite strict since we are not referring to expectations over multiple random paths, but deterministic equality over every single sample sequence of transitions.

Adopting the notion of normative systems as a coalition of players lends support for the adoption of concepts from cooperative game theory, namely:

Definition 4.1. Given normative system N, two individual norms n_1 and n_2 are *interchangeable in N with respect to value v* if, for any normative system $N' \subseteq N \setminus \{n_1, n_2\}$ where at least both n_1 and n_2 have been substituted by the baseline, the alignments with respect to value v after introducing either norm are identical, $\mathsf{Algn}^G_{N' \cup \{n_1\}, v} = \mathsf{Algn}^G_{N' \cup \{n_2\}, v}$.

Definition 4.2. An individual norm n is said to be a *null norm* with respect to value v in normative system N if, for any normative system $N' \subseteq N \setminus \{n\}$ excluding n (substituted by the baseline), the introduction of norm n does not change the alignment with respect to value v, $\mathsf{Algn}^G_{N' \cup \{n\}, v} = \mathsf{Algn}^G_{N', v}$.

It follows from eq. (7) and some basic manipulations that two interchangeable norms have identical Shapley values, and that any null norm has Shapley value equal to zero.

Running example

In our example model, setting the baseline parameters can be manually made thanks to its simplicity:

$$N_{baseline} = \left\{ \begin{array}{l} n_1 \sim collect = [0, ..., 0] \\ n_2 \sim redistribute = [\frac{1}{5}, ..., \frac{1}{5}] \\ n_3 \sim catch = 0 \\ n_4 \sim fine = 0 \end{array} \right\} \quad (8)$$

Note that the choice of the *redistribute* list respects the constraint (2). We have experimentally checked that this choice of parameters does indeed leave the initial global state of the system unchanged.

Table 2: Shapley values for all the individual norms conforming the optimal normative systems with respect to the value for which they are optimised.

Value	Norm	Shapley value
Equality	n_1	0.50
	n_2	0.03
	n_3	0.08
	n_4	0.01
Fairness	n_1	0.19
	n_2	0.45
	n_3	0.46
	n_4	0.41

Table 2 presents the results of computing the Shapley values of every individual norm in the optimal normative systems (shown in Table 1) with respect to the value for which they have been optimised. For value *equality*, the norm with the highest Shapley value is by far n_1, which is related to the collection of taxes. All other norms appear to be close to null, with Shapley values ~ 0, including the other norm of economic nature, n_2, linked to the redistribution of the common fund. This results would indicate that the distribution of wealth is already shrunk after taxes are collected. Then, the invested fund is redistributed roughly equally across wealth groups (see Table 1), which would only result in a horizontal shift of the wealth distribution and not on its width. Hence, the results indicate that the Gini Index is mostly reduced after taxes have been collected but not yet paid back.

In contrast, for value *fairness*, the situation is the opposite, with norms n_2, n_3 and n_4 having all similar and large Shapley values significantly above that of norm n_1. This would indicate that to punish evaders, we first need to detect them (n_3), as undetected evaders would automatically rise as the wealthiest members in the society. Then, the common fund needs to be very unevenly redistributed (see the optimal parameters in Table 1) towards the middle class, which is mostly composed by law-abiding citizens, as we have argued in Section 3. Note that imposing a fine on evaders is important, yet it is only the norms with the third largest Shapley value. Once evaders have been detected and their resources directed elsewhere (yet at the expense of harming law-abiding citizens), payment of an additional fine does not appear to be as relevant.

5 VALUE COMPATIBILITY

As seen above, by analysing the Shapley values of individual norms, we can study the interactions and dependencies between them. Now, in this section, we focus on a complementary examination regarding the interaction *among values* given a fixed normative system. Established philosophical theory on the topic of human values remarks the dynamic relationships between them [18]. In other words, aggressively pursuing one value has consequences that may be in direct conflict or congruent with others. We wish to examine these dynamic relationships closely. In particular, we have synthesised normative systems that are optimal for some value v_i, and that therefore promote it very strongly. How well are they aligned with respect to another value v_j?

Such an evaluation provides numerical estimates of whether two values are congruent under the same normative system, indicated by a large positive alignment close to +1, or whether they are in strong opposition, indicated by a large negative alignment close to -1. To mathematically capture this concept, we provide a formal definition of *value compatibility*:

Definition 5.1. Given a fixed normative system N, a set of values $V = \{v_1, v_2, ..., v_k\}$ is said to be *compatible to degree d*, or *d-compatible*, under N, if, for all values $v \in V$, it holds that $\mathsf{Algn}^G_{N,v} \geq d$.

It immediately follows that, if some set of values are compatible to degree d, they are also compatible to any degree $d' \leq d$ under the same norms. Normative systems that are highly specialised towards a particular value at the expense of others will result in low, possibly negative compatibility degrees. Normative systems that compromise on upholding different values should result in much larger compatibility degrees, though it most surely will still be far from the ideal situation, where the norms are all maximally aligned with respect to all the values and hence $d \simeq 1$.

The concept of compatibility works with alignment functions with respect to single values, $\mathsf{Algn}^G_{N,v}$. Nonetheless, we can expand it to aggregations over values, $\mathsf{Algn}^G_{N,V}$:

Definition 5.2. Given a set V of d-compatible values under a normative system N, an aggregated alignment function over the set of values, $\mathsf{Algn}^G_{N,V}$ is said to *preserve the compatibility* if $\mathsf{Algn}^G_{N,V} \geq d$.

Trivially, aggregation functions based on linear combinations of the alignments for the individual values do preserve d-compatibility, by choosing $d = \min_{v \in V} \mathsf{Algn}_{N,v}$.

The challenge regarding value compatibility is to find the combination of parametric norms that maximises the degree d_{max} for a set of relevant values. This task, however, is outside the scope of the present work and is left for future research.

Running example

Table 3: Cross-alignment for the optimal normative systems that have been optimised for value v_i with respect to value v_j.

		v_j	
		Equality	**Fairness**
v_i	**Equality**	-	-0.28
	Fairness	0.60	-

The cross-alignment values of the optimal normative systems produced for the running example of this paper are displayed in Table 3. It shows that the optimal normative system for value equality has very poor negative alignment for fairness, meaning that it is very specialised towards equality, yet at the expense of not being harsh on evaders. Under this normative system, despite the high alignment with respect to equality, the two values are very incompatible. As displayed in Figure 1b, few of the evaders end up with wealth similar to the bulk of the agents, therefore diminishing the chances of them belonging to #1. In fact, some of the evaders manage to gather resources well above the average law-abiding agent. Although this feature is not explicitly captured by eq. (5), it clearly goes against our sense of what a fair tax code should produce. In summary, the optimal normative system with respect to equality is highly specialised towards this value, at the expense of being lax with evader agents.

In contrast, the normative system optimised for value fairness respects equality to a much larger degree. With the help of Figure 1c, we provide an intuitive explanation for this result. In its quest to treat evaders harshly, this normative system also impoverishes law-abiding citizens, most of which move to the lower half of the initial wealth range, since the retrieval of evader agent resources is not done mostly through fining that would target them exclusively. Hence, the wealth distribution is effectively narrower than initially, and the Gini Index is reduced. The squeeze in the wealth distribution is not as acute as that produced by the optimal norms with respect to equality, and hence the compatibility between "fairness" and "equality" is moderate, ~ 0.6.

Interestingly, pursuing equality is detrimental towards fairness, but pursuing fairness is somewhat congruent with equality. The relationship is not symmetrical.

6 CONCLUSIONS

In this work, we have provided a methodology to synthesise value-promoting parametric normative systems. We have illustrated this methodology with a straightforward social model emulating a tax system. However, our proposed approach is general and can be applied to much more complex systems with sophisticated regulations, as long as the intervening norms are tied to optimisable quantities. Also, we have taken the approach of defining alignment functions directly over the final global state of the system and already aggregated over the set of agents. However, one might wish to return to the original proposal in [22] and compute alignments from average individual preferences, as distinct participants might conceive the manifestation of the same value differently. Our methodology works with either choice.

Despite its flexibility, our methodology does possess some limitations. The most obvious one is the discrete and finite nature of the models that it can be applied to. Also, even a relatively small amount of regulating norms can result in a very large search space over many dimensions. This fact might lead to computationally prohibitive optimisation searches, hindering its application to very complex models.

By viewing any normative system as a coalition of individual norms, one can adapt the concept of *Shapley value* from cooperative game theory to the context of norms. While the alignment evaluates how appropriate is a normative system overall, the Shapley values grant a deeper understanding of the role of individual norms and the mechanism by which they achieve high compliance with the encoded values.

Finally, we have provided a formal basis to tackle the concept of *compatible values*. Compatibility is an interesting point since it touches on the field of unintended consequences, *i.e.* how the aggressive promotion of some values can be detrimental or congruent with others. In our example model, we have observed one example for each situation.

Further work building up on our results should investigate which are the most efficient algorithms and techniques to find the optimal normative parameters, hence partially overcoming one of the limitation we have pointed at. Additionally, the incorporation of *aggregated* alignment functions that simultaneously consider various values should shed more light on the issue of value compatibility.

In summary, in this work, we provided methodology and tools for the synthesis and evaluation of norms guided by the need to respect, promote and uphold our human values. We believe that our approach can be helpful in the field of automated policy analysis and design.

7 CODE AVAILABILITY

All the necessary code to go along with this paper has been integrally developed in Python3. It is available under an MIT license at https://github.com/nmontesg/aamas21.

ACKNOWLEDGMENTS

This work has been supported by the AppPhil project (RecerCaixa 2017), the CIMBVAL project (funded by the Spanish government, project # TIN2017-89758-R), the EU WeNet project (H2020 FET Proactive project # 823783) and the EU TAILOR project (H2020 # 952215).

REFERENCES

[1] Thomas Ågotnes, Wiebe van der Hoek, Juan A Rodríguez-Aguilar, Carles Sierra, and Michael Wooldridge. 2009. *A Temporal Logic of Normative Systems.* Springer

Netherlands, Dordrecht, 69–106. https://doi.org/10.1007/978-1-4020-9084-4_5
[2] Giulia Andrighetto, Guido Governatori, Pablo Noriega, and Leon van der Torre. 2013. *Normative Multi-Agent Systems*. Vol. 4. Dagstuhl Publishing, Saarbrücken/Wadern. https://doi.org/10.4230/DFU.Vol4.12111.i
[3] Katie Atkinson and Trevor Bench-Capon. 2016. States, goals and values: Revisiting practical reasoning. *Argument and Computation* 7, 2-3 (2016), 135–154. https://doi.org/10.3233/AAC-160011
[4] The World Bank and Development Research Group. [n.d.]. Gini index (World Bank estimate). https://data.worldbank.org/indicator/SI.POV.GINI
[5] Trevor Bench-Capon and Sanjay Modgil. 2017. Norms and value based reasoning: justifying compliance and violation. *Artificial Intelligence and Law* 25, 1 (2017), 29–64. https://doi.org/10.1007/s10506-017-9194-9
[6] Stephen Cranefield, Michael Winikoff, Virginia Dignum, and Frank Dignum. 2017. No pizza for you: Value-based plan selection in BDI agents. *IJCAI International Joint Conference on Artificial Intelligence* 0 (2017), 178–184. https://doi.org/10.24963/ijcai.2017/26
[7] Virginia Dignum. 2019. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way* (1 ed.). Springer International Publishing. 127 pages. https://doi.org/10.1007/978-3-030-30371-6
[8] Moser Silva Fagundes, Sascha Ossowski, Jesús Cerquides, and Pablo Noriega. 2016. Design and evaluation of norm-aware agents based on Normative Markov Decision Processes. *International Journal of Approximate Reasoning* 78 (2016), 33 – 61. https://doi.org/10.1016/j.ijar.2016.06.005
[9] Emilia Garcia, Adriana Giret, and Vicente Botti. 2015. *Regulated Open Multi-Agent Systems (ROMAS)* (1 ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-11572-6_8
[10] Corrado Gini. 1912. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche. Fascicolo 1° Introduzione - Indici di variabilità - Indici di mutabilità*. Facoltà di Giurisprudenza della R. Università Cagliari, Bologna. 156 pages.
[11] Sean Luke. 2016. Population Methods. In *Essentials of Metaheuristics* (second ed.). Lulu, Chapter 3. http://cs.gmu.edu/~sean/book/metaheuristics/
[12] Javier Morales, Maite López-Sánchez, Juan A. Rodríguez-Aguilar, Michael Wooldridge, and Wamberto Vasconcelos. 2013. Automated synthesis of normative systems. *12th International Conference on Autonomous Agents and Multiagent Systems 2013, AAMAS 2013* 1 (2013), 483–490.
[13] Andreasa Morris-Martin, Marina De Vos, and Julian Padget. 2019. *Norm emergence in multiagent systems: A viewpoint paper*. Vol. 33. Springer US. 706–749 pages. https://doi.org/10.1007/s10458-019-09422-0
[14] Heinz Mühlenbein and Dirk Schlierkamp-Voosen. 1993. Predictive Models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization. *Evolutionary Computation* 1, 1 (1993), 25–49. https://doi.org/10.1162/evco.1993.1.1.25
[15] Shmuel Onn and Moshe Tennenholtz. 1997. Determination of social laws for multi-agent mobilization. *Artificial Intelligence* 95, 1 (1997), 155–167. https://doi.org/10.1016/S0004-3702(97)00045-3
[16] Sebastian Ruder. 2016. An overview of gradient descent optimization algorithms. (2016), 1–14. arXiv:1609.04747 http://arxiv.org/abs/1609.04747
[17] Stuart Russell. 2017. Provably Beneficial Artificial Intelligence. In *The Next Step: Exponential Life*. Turner Libros, Madrid.
[18] Shalom H. Schwartz. 1992. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in Experimental Social Psychology*. Vol. 25. 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6
[19] Marc Serramia, Maite López-Sánchez, Juan A. Rodríguez-Aguilar, Javier Morales, Michael Wooldridge, and Carlos Ansotegui. 2018. Exploiting Moral Values to Choose the Right Norms. *AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (2018), 264–270. https://doi.org/10.1145/3278721.3278735
[20] Lloyd S. Shapley. 1953. A Value for N-Person Games. In *Contributions to the Theory of Games (AM-28)*. Annals of Mathematics Studies, Vol. 2. Princeton University Press.
[21] Yoav Shoham and Moshe Tennenholtz. 1995. On social laws for artificial agent societies: off-line design. *Artificial Intelligence* 73, 1-2 (1995), 231–252. https://doi.org/10.1016/0004-3702(94)00007-N
[22] Carles Sierra, Nardine Osman, Pablo Noriega, Jordi Sabater-Mir, and Antoni Perelló-Moragues. 2019. Value alignment: a formal approach. *Responsible Artificial Intelligence Agents (RAIA) in AAMAS 2019*, 15.
[23] Norbert Wiener. 1960. Some Moral and Technical Consequences of Automation. *Science* 131, 3410 (1960), 1355–1358. https://doi.org/10.1126/science.131.3410.1355 arXiv:17841602