# A Qualitative Approach to Composing Value-Aligned Norm Systems*

Marc Serramia
Artificial Intelligence Research Institute (IIIA-CSIC)
marcserr@iiia.csic.es

Maite Lopez-Sanchez
University of Barcelona
maite_lopez@ub.edu

Juan A. Rodriguez-Aguilar
Artificial Intelligence Research Institute (IIIA-CSIC)
jar@iiia.csic.es

## ABSTRACT

Research in normative multi-agent systems has explored several approaches to compute the norm system (set of norms) required to make coordination possible. More recently, norm selection supposes an already available collection of norms from which to select a norm system to enact. A key aspect in this selection process is the consideration of moral values together with preferences among them, thus the selection follows the principle: the more preferred the values promoted by a norm system, the more preferred the norm system. Unfortunately, norm selection follows a quantitative approach despite the qualitative nature of the information available to the decision maker. In this paper we provide a novel qualitative approach to norm selection by formalising the process to infer a norm system ranking from the value preferences. We provide an encoding of this qualitative problem into a linear program and show that their solutions are equivalent.

## CCS CONCEPTS

• **Computing methodologies → Multi-agent systems**; **Cooperation and coordination**;

**ACM Reference Format:**
Marc Serramia, Maite Lopez-Sanchez, and Juan A. Rodriguez-Aguilar. 2020. A Qualitative Approach to Composing Value-Aligned Norm Systems. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 9 pages.

## 1 INTRODUCTION

Norms can be considered as coordination mechanisms that regulate the behaviour of individuals within a society. Research in normative multi-agent systems (MAS) has explored several approaches to compute the norm system (set of norms) required to make coordination possible. On the one hand, bottom-up norm emergence (or convention emergence) approaches (e.g. [1, 2, 16, 19, 20]) have norms of a MAS emerge from within the agent society at run-time. On the other hand, top-down norm synthesis approaches (e.g. [12, 13]) obtain norms through the external observation of agents' behaviours and their interactions' outcomes.

Alternatively, norms have also been related to moral values [7, 17, 18]. Instead of addressing norm synthesis, in [17, 18] we relied on an available collection of norms from which to *select* a norm system (that is, a set of norms to enact) based on the moral values they support. Specifically, this selection considers value preferences and the degree of value promotion of norms and is performed according to the following principle: the more preferred the values promoted by a norm system, the more preferred the norm system, or, in other words, the more *value-aligned.* Thus, in this context, the decision maker aims at the selection of the *most value-aligned norm system.* Besides that, the norm selection is conceived to consider the relationships between norms (e.g. a norm prohibiting action $a$ is not compatible with another norm obliging to do $a$ and thus, they cannot belong to the same norm system).

Although novel, the norm selection in [18] follows a quantitative approach despite the qualitative nature of the information available to the decision maker. First, since this approach follows [4], the decision maker is forced to quantify the relations between norms and values by specifying the degrees of value promotion of norms. This is hard to ascertain, and as noted in [15], transforming qualitative information into numerical data is prone to errors and biases. Second, that approach relies on building an arbitrary utility function for values from the preferences in the value system. In fact, as we will discuss in this paper, the utility function proposed in [18] can bias the selection towards values being supported by a larger number of norms to the detriment of most preferred values.

Against this background, in this paper we take a different stance. Thus, we show that it is possible to solve the norm selection problem faced by a decision maker while solely relying on the qualitative information available to her, namely the relations between norms, the preferences over values in the value system, and the relations between norms and values.

In order to find the most value-aligned norm system, we must be able to compare norm systems qualitatively, based on their alignment with the value preferences in the value system. For that, we exploit recent, seminal work in the realm of social choice theory [6, 8, 9] that makes possible to rank objects from a preference relation over their subsets. In short, by combining the *lex-cel* ranking described in [6] with the novel *anti-lex-cel* ranking introduced here, we manage to produce a ranking of norm systems. Such ranking exploits the preferences in the value system so that the more preferred the values promoted by a norm system, the more preferred the norm system is. Finally, based on our ranking of norm systems, we show how to encode the norm selection problem as a linear program (LP) so that we can solve it with the aid of state-of-the-art LP solvers without further assumptions.

To summarise, in this paper we provide a novel qualitative approach to solve the norm selection problem originally posed in [18]. With this aim, we make the following two major contributions:

*Funded by AI4EU(H2020-825619), H2020-769142, PGC2018-096212-B-C33, 19S01329-001 by Ajuntament de Barcelona (Fundació Solidaritat UB).

*Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.* 

- A novel ranking that allows to qualitatively rank norm systems based on the values that they promote.
- Based on our ranking, we show how to translate our qualitative optimisation problem, the norm selection problem, into a linear program. Importantly, we prove that the optimal solution to the resulting LP is also the optimal solution to the problem of selecting the most value-aligned norm system.

The paper is organised as follows. Section 2 provides background on norms, values, and order theory, the core components of our decision-making problem and its resolution process. Section 3 poses an informal outline of the problem we want to solve and provides a formalisation of each of the problem's steps of resolution (Sections 3.1, 3.2 and 3.3), once all the problem objects are formalised we provide a formal definition of the problem (Section 3.4). Section 4 provides an encoding to solve the problem. Section 5 provides resolutions for the running example and compares the approach we describe to that of [18]. Finally, in Section 6 we conclude.

## 2 BACKGROUND

In this section we introduce the core components of our decision-making problem. First, we define norms, their relations and the norm net, which we borrow from [18].

### 2.1 Norm net

Hereafter we consider a multi-agent system composed of a set of agents $Ag$, with a finite set of actions $\mathcal{A}$ available to them, which can enact different roles from a set $R$. These roles are hierarchically organised in $H = \langle R, \trianglelefteq \rangle$, where $\trianglelefteq$ is a hierarchical relation of subsumption (for two roles $\rho, \rho' \in R$, $\rho' \trianglelefteq \rho$ means $\rho$ can subsume the functionalities of $\rho'$). Furthermore, we consider a simple, first-order language $\mathcal{L}$ to describe the state of the multi-agent system. With these definitions in place, we formalise the object of our problem, the norm:

Def. 1 (Norm). *A norm is a pair $\langle \varphi, \theta(\rho, a) \rangle$, where $\varphi$ is a precondition in the language $\mathcal{L}$, $\rho \in R$ is the role of the addressee agent, namely the agent role required to comply with the norm; $a \in \mathcal{A}$ is the regulated action, and $\theta \in \{Obl, Per, Prh\}$ is a deontic operator.*

Example 1. *Say that a country has to decide the norms to apply to its airport borders. A possible norm would be $\langle \emptyset, Obl(all, show\_passport) \rangle$ meaning that all agents must show their passports. Another possibility would be a norm such as $\langle \emptyset, Per(all, cross) \rangle$, permitting all agents to cross the border.*

Given a set of norms $N$, as noted in [18], relationships between norms may hold. Thus, we identify norm exclusivity, substitutability and generalisation as norm relations. Such relationships are relations over norms, henceforth noted as $R_x$, $R_s$, and $R_g$ respectively. Two norms $n, n'$ are *mutually exclusive*, noted as $(n, n') \in R_x$, when they cannot be enacted at once; they are *substitutable*, noted as $(n, n') \in R_s$, if they are interchangeable; and they have a direct generalisation relation, noted as $(n, n') \in R_g$, when $n$ is more general than $n'$. We note $A(n)/S(n)$ the ancestors/successors of $n$.

By putting together norms and their relations, we characterise the normative dimension of our decision space.

Def. 2. *A norm net is a structure $\langle N, R \rangle$, where $N$ is a set of norms and $R = \{R_x, R_s, R_g\}$ is the set of exclusive, substitutable, and generalisation relations.*

Likewise [18], henceforth we shall refer to any subset $\Omega \subseteq N$ as a *norm system*. We are interested in a particular type of norm systems: those that contain neither conflicting nor redundant norms. Thus, we characterise norm systems that avoid both conflicts and redundancy as *sound* norm systems.

Def. 3. *Given a norm net $\langle N, R \rangle$, a norm system $\Omega \subseteq N$ is sound iff it is both conflict-free and non-redundant, that is a norm system $\Omega \subseteq N$ is sound if for each $n_i, n_j \in \Omega$, $(n_i, n_j) \notin R_x$; $(n_i, n_j) \notin R_s$; $n_j \notin A(n_i)$; and $\forall n$, such that $|\bar{S}(n)| > 1$, then $\bar{S}(n) \nsubseteq \Omega$, where $\bar{S}(n)$ are the direct successors of $n$ ($\bar{S}(n) = \{n' \in N, (n, n') \in R_g\}$).*

Example 2. *In an airport border we consider the norm net in Figure 1. In this case, $n_1$ permits everyone to cross the border, $n_2$ obliges them to show an id, $n_3$ obliges them to show a passport, $n_4$ obliges all agents to pass a security check, $n_5$ obliges them to fulfil a customs form, and, finally, $n_6$ obliges only foreign travellers to fulfil a customs form while $n_7$ only obliges it to local travellers. Note that $n_6$ and $n_7$ are identical to $n_5$ but only with more particular roles (local/foreigner instead of all), thus $n_5$ generalises them. On the other hand, note that $n_1$ is mutually exclusive with $n_2, n_3, n_4, n_5, n_6$ and $n_7$, as crossing the border freely cannot be enacted while demanding anything to cross it. Finally, since $n_2$ and $n_3$ oblige travellers to show alternative travel documents we consider them substitutable norms.*

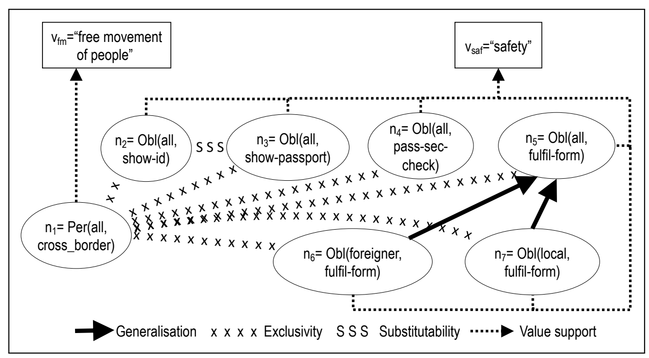


**Figure 1: Example of candidate norms $n_1, \ldots, n_7$ for border control along with their relations and their promotion of the free movement and safety values.**

### 2.2 Rankings

While norms and norm systems are the object of study, we use rankings as the way to define preferences. We use rankings because it is the less restrictive preference structure satisfying totality. This allows us to know the preferences between any pair of elements (unlike e.g. partial orders).

Def. 4. *A ranking $\geq$ of $X$ is a reflexive, transitive and total binary relation, noted as $x \geq y$ ($x, y \in X$). Given $x, y \in X$, if $x \geq y$, we say* $x$ is more preferred than $y$. *If we both have $x \geq y$ and $y \geq x$ then we note it as $x \sim y$ and say* $x$ and $y$ are indifferently preferred. *We note as $\mathcal{R}(X)$ the set of all rankings of $X$.*

Note that, given a set $X$ and a ranking $\succeq \in \mathcal{R}(X)$, we can consider the quotient set $X/\sim$ and its associated quotient order $\succ$, such that $\Sigma_1 \succ \cdots \succ \Sigma_{|N/\sim|}$ where $\Sigma_i$ is the class of equivalence containing the $i$-th more preferred elements of the ranking $\succeq$ (that are indifferently preferred between them).

### 2.3 Value System

Ethical reasoning typically involves a *value system*, that contains a set of moral values, which are principles that society deems valuable. As noted in [5], within a value system, some values are preferred to others, and such preferences over moral values influence decision making. Therefore, the preferences over the moral values of a value system, together with the values themselves, have been identified as a core component for ethical reasoning in [5, 11, 18]. Formally,

Def. 5. *A value system is a pair $\langle V, \succeq \rangle$, where $V$ stands for a non-empty set of values, and $\succeq$ is a ranking of the moral values in $V$.*

In [18], we argued that a norm is related to a given moral value when the norm's goal either promotes or demotes that value. Based on such observation, and taking inspiration in [4], we defined a support rate function that quantifies the support of a norm to each value. Unlike [18] and [4], here we do not assume that such quantitative information is available, and instead we propose that the decision maker is just required to specify whether a norm promotes a value or not. With this aim, we define the following function linking norms and values:

Def. 6. *Given a norm net $\langle N, R \rangle$ and a value system $\langle V, \succeq \rangle$, we define the value promotion function $\sigma_v : N \to \mathcal{P}(V)$ that for each norm returns the subset of values the norm promotes. Because of the norm relations in the norm net, we require this function to fulfil the following syntactic consistency properties:*

- *Exclusive norms cannot share the values they support: If $(n, n') \in R_x$, then $\sigma_v(n) \cap \sigma_v(n') = \emptyset$.*
- *All values promoted by a norm are also promoted by the norms that generalise it: If $(n, n') \in R_g$, then $\sigma_v(n') \subseteq \sigma_v(n)$.*

Furthermore, we also consider the complementary function $\sigma_n : V \to \mathcal{P}(N)$ that yields the norms promoting a particular value:
(i) if $\sigma_v(n) = V'$, where $V' \subseteq V \Rightarrow \forall v \in V', n \in \sigma_n(v)$; and
(ii) if $\sigma_n(v) = N'$, where $N' \subseteq N \Rightarrow \forall\, n \in N', v \in \sigma_v(n)$.
In norm selection, a norm that does not support any value and a value that is not supported by any norm are irrelevant. Henceforth, we suppose that all norms support at least one value, and that all values are supported by at least one norm ($\forall v \in V$, $\sigma_n(v) \neq \emptyset$ and $\forall n \in N$, $\sigma_v(n) \neq \emptyset$

Example 3. *Following Example 2, we observe that $n_1$ promotes free movement ($\sigma_v(n_1) = \{v_{fm}\}$), whereas the rest of norms promote safety ($\sigma_v(n_2) = \sigma_v(n_3) = \ldots = \sigma_v(n_7) = \{v_{saf}\}$), and hence $\sigma_n(v_{fm}) = \{n_1\}$, and $\sigma_n(v_{saf}) = \{n_2, n_3, n_4, n_5, n_6, n_7\}$. Notice that the image of $\sigma_n$ is $\sigma_n(V) = \{\{n_1\}, \{n_2, n_3, n_4, n_5, n_6, n_7\}\}$*

### 2.4 Lex-cel

Given a set of elements $X$ and a ranking $\succeq$ of the subsets of $X$, one interesting topic is how to ground the ranking of subsets to a ranking of elements. In this paper we use the grounding method introduced in [6] called lex-cel. Lex-cel can be viewed as a function $le : \mathcal{R}(\mathcal{P}(X)) \to \mathcal{R}(X)$, such that for a ranking $\succeq \in \mathcal{R}(\mathcal{P}(X))$, $le(\succeq) = \succeq^g$ is a ranking of $X$. Informally we say lex-cel grounds the preferences over subsets to preferences over elements. Although we note it as $le$ the process is performed as follows.

First, we consider the quotient set $\mathcal{P}(X)/\sim$, this way subsets related with indifference relations will fall on the same equivalence class $\Sigma_i \in \mathcal{P}(X)/\sim$. As we have explained in Sec. 2.2, these classes are ordered with a quotient order $\succ$: $\Sigma_1 \succ \cdots \succ \Sigma_{|\mathcal{P}(X)/\sim|}$.

We now define a function $\mu : X \to \mathbb{N}^{|\mathcal{P}(X)/\sim|}$, that for an element $x \in X$ returns its profile vector, a natural vector of dimension $|\mathcal{P}(X)/\sim|$, the number of equivalence classes in the quotient set. The elements in the profile vector represent the number of times $x$ appears in the subsets of each of the equivalence classes. Thus, supposing $\mu(x) = (c_1^x, \ldots, c_{|\mathcal{P}(X)/\sim|}^x)$, $c_i^x$ is the number of times $x$ appears in the subsets of the equivalence class $\Sigma_i$, which is the class containing the $i$-th most preferred subsets of $\mathcal{P}(X)$ in regard to $\succeq$. Formally:

$$\mu(x) = (c_1^x, \ldots c_{|\mathcal{P}(X)/\sim|}^x), \text{ where } c_i^x = |\{S \in \Sigma_i : x \in S\}| \qquad (1)$$

To compare two elements $x, y \in X$, we compare lexicographically $\mu(x)$ and $\mu(y)$. That is:

Def. 7. *We define the lexicographical order of vectors $\geq_L$ such that given two vectors $c = (c_1, \ldots c_m), c' = (c'_1, \ldots c'_m) \in \mathbb{N}^m$, we say $c \geq_L c'$ iff $\exists i$, s.t. $c_1 = c'_1; \ldots; c_{i-1} = c'_{i-1}$ and $c_i > c'_i$.*

We then define the grounded ranking between two elements by comparing their profile vectors. Given $x, y \in X$, we say:

$$\begin{cases} x \succeq^g y \text{ iff } \mu(x) \geq_L \mu(y) \\ x \preceq^g y \text{ iff } \mu(x) \leq_L \mu(y) \\ x \sim^g y \text{ iff } \mu(x) = \mu(y) \end{cases}$$

In [6], the authors prove that grounding preferences with lex-cel satisfies properties that make the grounding fair, namely neutrality, coalitional anonymity, monotonicity and independence of the worst set. Firstly, neutrality ensures that the ranking resulting from applying lex-cel does not depend on the elements' names/identities. Coalitional anonymity states that a ranking between two elements $x, y$ should be independent of the other elements, either if $x, y$ are in coalitions with them or not. Lex-cel also satisfies monotonicity because it breaks possible indifference relations in a consistent way. Furthermore, when grounding preferences lex-cel takes into account higher ranked subsets over lower ranked ones, in fact lex-cel does not take into account the least preferred subsets, thus we say lex-cel is independent of the worst subsets. In [6], the authors not only prove that lex-cel satisfies these axioms but also that is the only grounding function to satisfy them.

## 3 THE MOST VALUE-ALIGNED NORM SYSTEM SELECTION PROBLEM

Given a norm net $\langle N, R \rangle$ (we assume that the norms in $N$ are *beneficial* candidate norms), a value system $\langle V, \succeq \rangle$ and a function $\sigma_n$, our goal is to detail the process to find *the most value-aligned norm system* (MVANS), that is the most preferred sound norm system in terms of the values it promotes based on the value ranking.

To find this desired norm system, in this section we detail how to employ the value ranking in the value system to infer a ranking

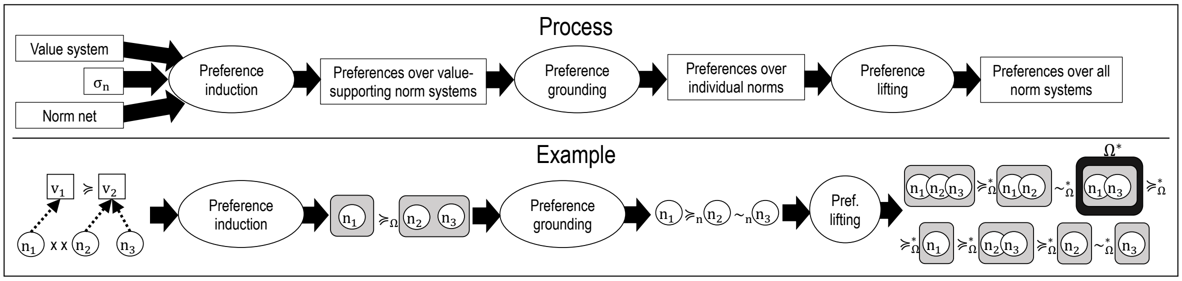


**Figure 2: (Top row) Steps to build a norm system ranking. (Bottom row) Example illustrating the building of a norm system ranking for a norm net and a value system and showing the most value-aligned sound norm system ($\Omega^*$) highlighted in black.**

of norm systems from which we can select the most preferred one that complies with the soundness requirements. Specifically, this section is devoted to detail the process of inferring a ranking of norm systems from a ranking of values. The top row in Figure 2 outlines our process, which involves three steps (preference induction, preference grounding, and preference lifting), which are thoroughly detailed in Sections 3.1, 3.2, and 3.3 respectively. Before that, next we briefly sketch those three steps. For illustrative purposes, notice that the bottom row in Figure 2 displays the building of a ranking of norm systems for: (i) a norm net with three norms $n_1$, $n_2$ and $n_3$, where $n_1$ and $n_2$ are mutually exclusive; (ii) a value system with two values $v_1$ and $v_2$ where $v_1$ is preferred to $v_2$; and (iii) a norm-value support function relating $n_1$ with $v_1$ and $n_2$, $n_3$ with $v_2$.

The first step is preference induction, which is meant to yield a ranking $\succeq_\Omega$ of the value-supporting norm systems (that is, the subsets in $\sigma_n(V)$) from the value ranking $\succeq$. Observe that at this point we can only induce a ranking of the value-supporting norm systems instead of a ranking of all norm systems. In the example in Figure 2, we know that $v_1$ is supported by $n_1$ ($\sigma_n(v_1) = \{n_1\}$) and $v_2$ is supported by $n_2, n_3$ ($\sigma_n(v_2) = \{n_2, n_3\}$). Since $v_1 \succeq v_2$, we induce that $\{n_1\} \succeq_\Omega \{n_2, n_3\}$, namely $\{n_1\}$ is preferred to $\{n_2, n_3\}$.

While at this point the ranking we have is only a ranking of some norm systems, the information to rank all norm systems is implicit in it. We can make it explicit by first applying preference grounding to obtain the preferences over all individual norms and afterwards apply preference lifting to finally achieve a full norm system ranking. This amounts to applying preference grounding to the ranking $\succeq_\Omega$ of value-supporting norm systems to produce a ranking $\succeq_n$ of the norms in $N$. We do so by using the lex-cel method [6] described in Section 2.4 , which satisfies desirable properties for the method to be perceived as fair. Back to our example in Figure 2, the preference grounding of the $\succeq_\Omega$ ranking leads to the norm ranking $n_1 \succeq_n n_2 \sim_n n_3$. In this case we prefer $n_1$ over the other norms because $n_1$ supports $v_1$ which is the most preferred value, while $n_2$ and $n_3$ support $v_2$ which is less preferred.

The last step is preference lifting, which lifts the ranking of norms to a ranking $\succeq^*_\Omega$ of norm systems. In Figure 2, since the norm ranking $n_1 \succeq_n n_2 \sim_n n_3$ tells us that that $n_1$ is the most preferred norm, we will rank all norm systems containing $n_1$ over those norm systems that do not contain it. Additionally, given two norm systems containing $n_1$, we will prefer the one with more norms (as we have assumed all norms in $N$ to be *beneficial* candidate norms). Of the norm systems not containing $n_1$, the more preferred ones will again be those that contain a larger number of norms.

The process of grounding preferences and then lifting them can be viewed as a way of transforming the ranking $\succeq_\Omega$ of value-supporting norm systems into a ranking of all norm systems $\succeq^*_\Omega$.

At this point, finding the MVANS is finding the most preferred sound norm system in $\succeq^*_\Omega$. In Figure 2, since $n_1$ and $n_2$ are mutually exclusive, $\{n_1, n_2, n_3\}$ and $\{n_1, n_2\}$ are not sound norm systems. Therefore, the MVANS is $\{n_1, n_3\}$.

## 3.1 Preference induction over value-supporting norm systems

Let $V = \{v_1, \dots v_r\}$ be a set of values, $\sigma_n(v_i)$ the norm system supporting value $v_i$, and $\sigma_n(V) = \{\sigma_n(v_1), \dots, \sigma_n(v_r)\}$ the set of norm systems supporting each value. Intuitively, if norm $n$ supports value $v$ and norm $n'$ supports value $v'$, knowing that $v \succeq v'$, we can infer that, in terms of values, $n$ is more preferred than $n'$. Since norms support values, the ranking of values in a value system induces a ranking of the norm systems in $\sigma_n(V)$. Thus, we can define a function $ind : \mathcal{R}(V) \to \mathcal{R}(\sigma_n(V))$, such that for a value ranking $\succeq$ we can infer a new ranking $ind(\succeq) = \succeq_\Omega$ of the subsets in $\sigma_n(V)$ as follows:

$$v \succeq v' \Rightarrow \sigma_n(v) \succeq_\Omega \sigma_n(v'), \forall v, v' \in V \qquad (2)$$

Example 4. *In Figure 1, say that we prefer freedom of movement over security ($v_{fm} \succeq v_{saf}$). Thus, in this case, the induced ranking is:* $\{n_1\} \succeq_\Omega \{n_2, n_3, n_4, n_5, n_6, n_7\}$.

## 3.2 Preference grounding over norms

We now tackle grounding the ranking of value-supporting norm systems into a norm ranking $\succeq_n$. We do so by using the lex-cel approach described in [6] and explained in Section 2.4.

At this point we have used the value ranking to induce $\succeq_\Omega$ a ranking of $\sigma_n(V)$, but notice that as explained in Section 2.4, lex-cel has as input a ranking of $\mathcal{P}(N)$. Note though, that since $\sigma_n(V) \subseteq \mathcal{P}(N)$, the ranking $\succeq_\Omega$ can be extended as a ranking of $\mathcal{P}(\mathcal{N})$ with uncertainty over the norm systems in $\mathcal{P}(N) \setminus \sigma_n(V)$. To do so we exploit lex-cel's property of independence of the worst set to extend the ranking $\succeq_\Omega$ to a ranking of $\mathcal{P}(N)$ by considering all missing norm systems in $\mathcal{P}(N) \setminus \sigma_n(V)$ as "the worst set". Effectively, this trick allows us to apply lex-cel to our current ranking over $\sigma_n(V)$ while not affecting its outcome as lex-cel is independent of the uncertainty we have added to the ranking. Therefore, we extend

$\succeq_\Omega \in \mathcal{R}(\sigma_n(V))$ to a ranking in $\mathcal{R}(\mathcal{P}(N))$ noted as $\succeq_\Omega^{ext}$. Such that given two norm systems $\Omega, \Omega' \in \mathcal{P}(N)$, if both $\Omega, \Omega' \in \sigma_n(V)$, then their order in $\succeq_\Omega^{ext}$ is the same as in $\succeq_\Omega$, but otherwise we say:

$$\begin{cases} \Omega \succeq_\Omega^{ext} \Omega' \text{ iff } \Omega \in \sigma_n(V) \text{ and } \Omega' \in \mathcal{P}(N) \setminus \sigma_n(V) \\ \Omega \preceq_\Omega^{ext} \Omega' \text{ iff } \Omega \in \mathcal{P}(N) \setminus \sigma_n(V) \text{ and } \Omega' \in \sigma_n(V) \\ \Omega \sim_\Omega^{ext} \Omega' \text{ iff } \Omega, \Omega' \in \mathcal{P}(N) \setminus \sigma_n(V) \end{cases}$$

We note $ext(\succeq_\Omega) = \succeq_\Omega^{ext}$. We can now apply lex-cel to the extended ranking to obtain a norm ranking, thus $le(\succeq_\Omega^{ext}) = \succeq_n$. The lex-cel process in this case consists of the following $\mu$ function:

$$\mu(n) = (x_1^n, \ldots, x^n_{|\mathcal{P}(N)/\sim_\Omega^{ext}|}), \text{ where } x_i^n = |\{S \in \Sigma_i : n \in S| \quad (3)$$

Example 5. *Following Example 4, and supposing the extended ranking $\succeq_\Omega^{ext}$, we now consider the quotient set $\mathcal{P}(N)/\sim_\Omega^{ext}$, the three equivalence classes are $\Sigma_1 = \{\{n_1\}\}$, $\Sigma_2 = \{\{n_2, n_3, n_4, n_5, n_6, n_7\}\}$ and $\Sigma_3 = \{\Omega | \Omega \in \mathcal{P}(N) \setminus \sigma_n(V)\}$ with the quotient order $\Sigma_1 \succ_\Omega^{ext} \Sigma_2 \succ_\Omega^{ext} \Sigma_3$. We now have the vectors $\mu(n_1) = (1, 0, 63)$, as $n_1$ appears in one (the only) norm system in the most preferred equivalence class $\Sigma_1$, does not appear in the norm system on $\Sigma_2$ and appears 63 times in $\Sigma_3$, because $\mathcal{P}(N)$ has $2^{|N|-1} = 64$ norm systems containing $n_1$, but $n_1$ appears in 1 of the norm systems of $\sigma_n(V)$, therefore only 63 norm systems are in $\mathcal{P}(N) \setminus \sigma_n(V)$. On the other hand, we have $\mu(n_2) = \ldots = \mu(n_7) = (0, 1, 63)$ as $n_2, \ldots, n_7$ do not appear on the norm system in $\Sigma_1$, they appear on one (the only) norm system in $\Sigma_2$, and appear on 63 norm systems of $\Sigma_3$ (as each norm $n_2, \ldots, n_7$ appears in one norm system of $\sigma_n(V)$).*

Now we compare two norms following the same principle as in Section 2.4, thus we say:

$$\begin{cases} n \succeq_n n' \text{ if } \mu(n) \geq_L \mu(n') \\ n \preceq_n n' \text{ if } \mu(n) \leq_L \mu(n') \\ n \sim_n n' \text{ if } \mu(n) = \mu(n') \end{cases}$$

Example 6. *From Example 5, the grounded norm ranking is $n_1 \succeq_n n_2 \sim_n n_3 \sim_n n_4 \sim_n n_5 \sim_n n_6 \sim_n n_7$, since $(1, 0, 63) \geq_L (0, 1, 63)$.*

So far, we have used the initial value ranking to induce a ranking of $\sigma_n(V)$ (the value-supporting norm systems) and subsequently we have obtained a ranking of $N$. We can now use the norm ranking to obtain a norm system ranking. Once we can compare all norm systems, we can find the most value-aligned norm system that satisfies the soundness properties. Therefore, to obtain a ranking of $\mathcal{P}(N)$ (all norm systems of $N$), in the following section we will lift the norm ranking to a ranking of all norms systems.

### 3.3 Preference lifting over all norm systems

Using lex-cel we have grounded the ranking $\succeq_\Omega$ of value-supporting norm systems to a ranking $\succeq_n$ of norms. We now aim at lifting this norm ranking to obtain a ranking $\succeq_\Omega^*$ of norm systems. Such ranking will help us compare any two norm systems in terms of the values each one supports. Since $\succeq_n$ is the norm ranking that we have inferred from value preferences, in fact, the more preferred a norm, the better its alignment with value preferences. Hence, we will target at a lifting function that adheres to the following axioms.

First, the comparison of two norm systems should arise from the differences on their most preferred norms. Thus, our first axiom will require that a norm system containing a single norm $\{n\}$ is more preferred than a norm system containing any norms strictly less preferred than $n$ because $n$ promotes more preferred values.

Axiom 1 (Order dominance). *Given $n \in N$ and the set of its strictly less preferred norms $L_n = \{n' \in N | n \succeq_n n' \text{ and } n \nsim_n n'\}$, $\{n\} \succeq_\Omega^* \Omega'$, where $\Omega' \subseteq L_n$.*

Second, adding further norms to a norm system adds more value alignment whenever the added norms promote some value(s). This is formally captured by our second axiom.

Axiom 2 (Size dominance). *For any norm system $\Omega \in \mathcal{P}(N)$, given $n \in N \Rightarrow \Omega \cup \{n\} \succeq_\Omega^* \Omega$.*

In fact, our axioms are in line with lex-cel. Following lex-cel, element $a$ is more preferred than element $b$ iff either: i) $a$ appears in a more preferred subset, while $b$ only appears in strictly less preferred subsets (similarly to Axiom 1); or (ii) in case both elements appear in equally preferred subsets, the element that appears in more subsets is actually preferred (similarly to Axiom 2).

With our axioms in mind, we base our lifting process on the same ideas as lex-cel, but inverting input and output. Thus, next we introduce the *anti-lexicographic* excellence function (anti-lex-cel), which takes a norm ranking as input and outputs a norm system ranking. And we design this function to satisfy Axioms 1 and 2.

To build anti-lex-cel, we start from the quotient set $N/\sim_n$. Each equivalence class contains a set of indifferently preferred norms. Equivalence classes in $N/\sim_n$ are ordered by the quotient order $\succ_n$. Hence, $\Xi_1 \succ_n \cdots \succ_n \Xi_r$, where $r = |N/\sim_n|$ and $\Xi_i$ is the equivalence class containing the $i$-th most preferred norms. We define a function $\eta : \mathcal{P}(N) \to \mathbb{N}^r$ to count the appearances of the norms in a norm system in each equivalence class. Thus, given a norm system $\Omega \in \mathcal{P}(N)$, $\eta(\Omega)$ is a vector of size $r$ whose $i$-th component stands for the number of norms in $\Omega$ that are found in the equivalence class $\Xi_i$. Formally:

$$\eta(\Omega) = (s_1, \ldots, s_r), \text{ where } s_i = |\Omega \cap \Xi_i| \quad (4)$$

Note that, similarly to $\mu$ in Section 3.2, $\eta(\Omega)$ is a vector whose elements represent how preferred the norms in $\Omega$ are: the larger the first elements of the vector, the more preferred the norms in $\Omega$ are (in terms of $\succeq_n$), and hence we can infer that the more preferred $\Omega$ will be. This again means that ranking norm systems is equivalent to lexicographically ordering their associated vectors as calculated by the $\eta$ function.

With those considerations, we are now ready to tackle the formulation of the anti-lex-cel function $ale$. We define $\succeq_\Omega^*$ as the ranking of norm systems in $\mathcal{P}(N)$ such that given two norm systems $\Omega, \Omega' \in \mathcal{P}(N)$, it orders them according to the following rules:

$$\begin{cases} \Omega \succeq_\Omega^* \Omega' \Leftrightarrow \eta(\Omega) \geq_L \eta(\Omega') \\ \Omega \preceq_\Omega^* \Omega' \Leftrightarrow \eta(\Omega) \leq_L \eta(\Omega') \\ \Omega \sim_\Omega^* \Omega' \Leftrightarrow \eta(\Omega) = \eta(\Omega') \end{cases} \quad (5)$$

After that, we formally define anti-lexicographic-excellence:

Def. 8. *Given a set of norms $N$ and a norm ranking $\succeq_n$, we call anti lexicographic excellence (anti-lex-cel), the lifting function $ale : \mathcal{R}(N) \to \mathcal{R}(\mathcal{P}(N))$, such that $ale(\succeq_n) = \succeq_\Omega^*$.*

As noticed above, the anti-lex-cel function is very similar to lex-cel though it realises the reverse process (from ranking over elements to ranking over subsets of elements). However notice that, since $le : \mathcal{R}(\mathcal{P}(X)) \rightarrow \mathcal{R}(X)$ is not injective (it cannot be as $|\mathcal{R}(\mathcal{P}(X))| > |\mathcal{R}(X)|$), $le$ is not invertible and therefore $ale$ is not the inverse of $le$.

We resort to our running example to illustrate anti-lex-cel lifting:

Example 7. *Consider the norm ranking in Example 6. The quotient set $N/{\sim_n}$ contains the equivalence classes $\Xi_1 = \{n_1\}$, $\Xi_2 = \{n_2, n_3, n_4, n_5, n_6, n_7\}$, with $\Xi_1 \succ_n \Xi_2$. We can now compute $\eta$ for all norm systems to subsequently determine their ranking. Since detailing the whole ranking would be lengthy, we show how to compute the preference between the following norm systems: $\{n_1\}$, $\{n_1, n_2, n_4\}$, $\{n_2, n_4, n_5\}$, $\{n_3, n_4, n_5\}$. We first compute their $\eta$ vectors: $\eta(\{n_1\}) = (1, 0)$; $\eta(\{n_1, n_2, n_4\}) = (1, 2)$; $\eta(\{n_2, n_4, n_5\}) = (0, 3)$; $\eta(\{n_3, n_4, n_5\}) = (0, 3)$. Since $(1, 2) \geq_L (1, 0) \geq_L (0, 3)$, the resulting norm system ranking is: $\{n_1, n_2, n_4\} \succeq^*_\Omega \{n_1\} \succeq^*_\Omega \{n_2, n_4, n_5\} \sim^*_\Omega \{n_3, n_4, n_5\}$.*

At this point, we have formalised a way of transforming the original value ranking in the value system into a ranking over all norm systems. It remains to show that the ranking proudced by anti-lex-cel $ale(\succeq_n) = \succeq^*_\Omega$ does satisfy Axioms 1 and 2.

Theorem 1. *Given a ranking $\succeq_n \in \mathcal{R}(N)$, the ranking $ale(\succeq_n) = \succeq^*_\Omega \in \mathcal{R}(\mathcal{P}(N)$ follows Axioms 1 and 2.*

Proof. 1. *The proof is straightforward considering that $\Omega \succeq^*_\Omega \Omega' \Leftrightarrow (|\Omega \cap \Xi_1|, \dots, |\Omega \cap \Xi_{|N/\sim_n|}|) \geq_L (|\Omega' \cap \Xi_1|, \dots, |\Omega' \cap \Xi_{|N/\sim_n|}|)$. Thus, Axiom 1 follows from $(0, \dots, 0, 1, 0, \dots, 0) \geq_L (0, \dots, 0, 0, |Xi_{i+1}|, \dots, |Xi_{|N/\sim_n|}|)$. Axiom 2 follows from $(|(\Omega \cup \{n\}) \cap \Xi_1|, \dots, |(\Omega \cup \{n\}) \cap \Xi_{|N/\sim_n|}|) \geq_L (|\Omega \cap \Xi_1|, \dots, |\Omega \cap \Xi_{|N/\sim_n|}|)$, since for given position $i$ $|(\Omega \cup \{n\}) \cap \Xi_i| = |\Omega \cap \Xi_i| + 1$.* □

### 3.4 Formalising the problem

At this point, we have learnt how to yield a ranking $\succeq^*_\Omega$ of all norm systems. Therefore, we are ready to provide a formal definition of the problem of finding the most value-aligned norm system:

Problem 1 (Most value-aligned norm system selection problem (MVANSP)). *Given a norm net $\langle N, R\rangle$ a value system $\langle V, \succeq\rangle$ and the value promoting norms function $\sigma_n$, the most valued-aligned norm system selection problem (MVANSP) is that of finding a sound norm system $\Omega^* \subseteq N$ such that $\Omega^* \succeq^*_\Omega \Omega$ for any other sound norm system $\Omega \subseteq N$ and the ranking $ale(le(ext(ind(\succeq)))) = \succeq^*_\Omega$.*

## 4 SOLVING THE MOST VALUE-ALIGNED NORM SYSTEM SELECTION PROBLEM

Solving Problem 1 turns out to be rather costly. There are two alternative ways we might consider. On the one hand, we might compute Eq. 4 for all the norm systems in $N$ in $O(2^{|N|})$ to subsequently order them following Eq. 5, which requires $O(2^{|N|} \cdot log(2^{|N|}))$ in the average case ($O(2^{|N|})$ worst case). On the other hand, we might alternatively consider to only compute Eq. 4 for sound norm systems, but this requires to check the soundness of all norm systems with $O(2^{|N|})$. Therefore, notice that both ways are exponential on the number of norms, hence hindering applicability (just consider solving problems involving hundreds or even thousands of norms!).

In what follows, we will show how to solve the MVANSP while avoiding the cost of building explicitly the $\succeq^*_\Omega$ ranking. In particular, we propose to encode it as a linear program (LP) so that it can be solved with the aid of off-the-shelf LP solvers. Importantly, we prove that the proposed encoding adheres to the $\succeq^*_\Omega$ ranking, and that the solution to our LP is equivalent to that of Problem 1.

Let $\Omega \subseteq N$ be a set of norms, consider $\eta(\Omega) = (c^\Omega_1, \dots, c^\Omega_r)$, where $c^\Omega_i = |\Omega \cap \Xi_i|$ and $r = |N/{\sim_n}|$, we compute the value alignment of norm system $\Omega$ as follows:

$$al(\Omega) = \sum_{i=1}^{r} |\Omega \cap \Xi_i| (\sum_{j=i+1}^{r} al(\Xi_j) + 1), \text{ where } al(\Xi_r) = |\Xi_r|. \quad (6)$$

Notice that by applying equation 6, we can compute the value alignment of class $\Xi_i$: $al(\Xi_i) = |\Xi_i|(\sum_{j=i+1}^{r} al(\Xi_j) + 1)$. Hence, the value alignments of the classes in quotient order $\Xi_1 \succ_n \cdots \succ_n \Xi_r$ can be computed recursively starting from $\Xi_r$. Note also that $al(\Omega) \geq 0$, and $al(\Omega) \in \mathbb{N}$ for any norm system $\Omega$.

Our value alignment function $al$ embodies the norm system ranking $\succeq^*_\Omega$, as we will formally prove later through a theorem. Before that, we need some results that help us in the proof of the theorem. The first lemma tells us how to alternatively compute the value alignment of a norm system.

Lemma 1. $al(\Omega) = \sum_{w=1}^{r} al(\Omega \cap \Xi_w)$

Proof. 2. *By applying equation 6 we obtain the value alignment of an equivalence class $\Xi_w$ as $al(\Omega \cap \Xi_w) = \sum_{i=1}^{r} |\Omega \cap \Xi_w \cap \Xi_i|(\sum_{j=i+1}^{r} al(\Xi_j)+1) = |\Omega \cap \Xi_w|(\sum_{j=w+1}^{r} al(\Xi_j)+1)$. Since all equivalence classes are disjoint, it follows that $|\Omega \cap \Xi_w \cap \Xi_i| = |\emptyset| = 0$ when $i \neq w$ and $|\Omega \cap \Xi_w \cap \Xi_i| = |\Omega \cap \Xi_w|$, when $i = w$. Now $\sum_{w=1}^{r} al(\Omega \cap \Xi_w) = \sum_{w=1}^{r} \sum_{i=1}^{r} |\Omega \cap \Xi_w \cap \Xi_i|(\sum_{j=i+1}^{r} al(\Xi_j) + 1) = \sum_{w=1}^{r} |\Omega \cap \Xi_w|(\sum_{j=w+1}^{r} al(\Xi_j) + 1) = al(\Omega)$.* □

And the second lemma bounds $al(\Omega \cap \Xi_w)$:

Lemma 2. $\forall w, al(\Xi_w) \geq al(\Omega \cap \Xi_w)$

Proof. 3. *Since all equivalence classes are disjoint, from equation 6 we have that $al(\Xi_w) = |\Xi_w|(\sum_{j=w+1}^{r} al(\Xi_j) + 1)$ and $al(\Omega \cap \Xi_w) = |\Omega \cap \Xi_w|(\sum_{j=w+1}^{r} al(\Xi_j) + 1)$. Since $|\Xi_w| \geq |\Omega \cap \Xi_w|$, then $al(\Xi_w) \geq al(\Omega \cap \Xi_w)$.* □

Now we can state the theorem that shows that the value alignment function embodies the $\succeq^*_\Omega$ ranking.

Theorem 2. *Given two norm systems $\Omega, \Omega' \in \mathcal{P}(N)$, $\Omega \succeq^*_\Omega \Omega' \Leftrightarrow al(\Omega) \geq al(\Omega')$.*

Proof. 4. *We divide the proof into three steps. First we prove the necessary and sufficient conditions, and we subsequently show that such proofs suffice to prove the theorem.*

*$\Omega \succ^*_\Omega \Omega' \Rightarrow al(\Omega) > al(\Omega')$: Say that $\Omega \succ^*_\Omega \Omega'$. From Equation 5, we have that $\Omega \succ^*_\Omega \Omega' \Leftrightarrow \eta(\Omega) >_L \eta(\Omega')$. By using the definition of $\eta$ in Equation 4 we can write $\eta(\Omega) >_L \eta(\Omega')$ as $(c^\Omega_1, \dots c^\Omega_r) >_L (c^{\Omega'}_1, \dots c^{\Omega'}_r)$ (where $c^\Omega_i = |\Omega \cap \Xi_i|$ and $c^{\Omega'}_i = |\Omega' \cap \Xi_i|$ $\forall i$). Now, by using the formalisation of the lexicographical order (see Definition 7), we have that $(c^\Omega_1, \dots c^\Omega_r) >_L (c^{\Omega'}_1, \dots c^{\Omega'}_r)$, which implies that $\exists k \in \{1, \dots, r\}$, s.t. $\forall t < k, c^\Omega_t = c^{\Omega'}_t$ and $c^\Omega_k > c^{\Omega'}_k$. In other words, $\exists k \in \{1, \dots, r\}$ s.t. $|\Omega \cap \Xi_k| > |\Omega' \cap \Xi_k|$ and $\forall t < k, |\Omega \cap \Xi_t| = |\Omega' \cap \Xi_t|$ and therefore $al(\Omega \cap \Xi_t) = al(\Omega' \cap \Xi_t)$.*

*Next we prove that $al(\Omega) > al(\Omega')$. First, note that by considering Lemma 1, we have that $al(\Omega) = \sum_{i=1}^{k-1} al(\Omega \cap \Xi_i) + \sum_{i=k}^{r} al(\Omega \cap \Xi_i)$*

$\geq \sum_{i=1}^{k-1} al(\Omega \cap \Xi_i) + al(\Omega \cap \Xi_k)$ *and applying Lemma 1 and Lemma 2 we have that* $al(\Omega') = \sum_{i=1}^{k-1} al(\Omega' \cap \Xi_i) + \sum_{i=k}^{r} al(\Omega' \cap \Xi_i) \leq \sum_{i=1}^{k-1} al(\Omega' \cap \Xi_i) + al(\Omega' \cap \Xi_k) + \sum_{i=k+1}^{r} al(\Xi_i)$. *Therefore, to prove that* $al(\Omega) > al(\Omega')$ *it suffices to prove that* $\sum_{i=1}^{k-1} al(\Omega \cap \Xi_i) + al(\Omega \cap \Xi_k) > \sum_{i=1}^{k-1} al(\Omega' \cap \Xi_i) + al(\Omega' \cap \Xi_k) + \sum_{i=k+1}^{r} al(\Xi_i)$. *This is equivalent to show that* $al(\Omega \cap \Xi_k) - al(\Omega' \cap \Xi_k) - \sum_{i=k+1}^{r} al(\Xi_i) > 0$.

*Now, using Equation 6,* $al(\Omega \cap \Xi_k) - al(\Omega' \cap \Xi_k) - \sum_{i=k+1}^{r} al(\Xi_i) = |\Omega \cap \Xi_k|(\sum_{j=k+1}^{r} al(\Xi_j)+1) - |\Omega' \cap \Xi_k|(\sum_{j=k+1}^{r} al(\Xi_j)+1) - \sum_{i=k+1}^{r} al(\Xi_i) = (|\Omega \cap \Xi_k| - |\Omega' \cap \Xi_k|)(\sum_{j=k+1}^{r} al(\Xi_j)+1) - \sum_{i=k+1}^{r} al(\Xi_i)$.

*As shown above, we know that* $|\Omega \cap \Xi_k| > |\Omega' \cap \Xi_k|$. *From that, and since these sets' cardinalities are natural numbers, we obtain the following lower bound:* $|\Omega \cap \Xi_k| - |\Omega' \cap \Xi_k| \geq 1$. *Therefore,* $(|\Omega \cap \Xi_k| - |\Omega' \cap \Xi_k|)(\sum_{j=k+1}^{r} al(\Xi_j) + 1) - \sum_{i=k+1}^{r} al(\Xi_i)) \geq \sum_{j=k+1}^{r} al(\Xi_j) + 1 - \sum_{i=k+1}^{r} al(\Xi_i)) = 1 > 0$.

*Recall that we assumed that* $\Omega \succ_{\Omega}^{*} \Omega'$. *Since we have managed to prove that* $\Omega \succ_{\Omega}^{*} \Omega'$ *implies that* $al(\Omega \cap \Xi_k) - al(\Omega' \cap \Xi_k) - \sum_{i=k+1}^{r} al(\Xi_i) > 0$, *which in turn implies that* $al(\Omega) > al(\Omega')$, *then it is clear that* $\Omega \succ_{\Omega}^{*} \Omega' \Rightarrow al(\Omega) > al(\Omega')$.

$\Omega \succ_{\Omega}^{*} \Omega' \Leftarrow al(\Omega) > al(\Omega')$: *Assume that* $al(\Omega) > al(\Omega')$ *and* $\Omega \not\succ_{\Omega}^{*} \Omega'$. *If* $\Omega \prec_{\Omega}^{*} \Omega'$, *then we have already shown above that* $al(\Omega) < al(\Omega')$, *which contradicts our initial assumption. If* $\Omega \sim_{\Omega}^{*} \Omega'$, *then* $\eta(\Omega) = \eta(\Omega')$, *which means that* $(c_1^{\Omega}, \dots, c_r^{\Omega}) = (c_1^{\Omega'}, \dots, c_r^{\Omega'})$, *and therefore* $\forall i\ c_i^{\Omega} = c_i^{\Omega'}$. *This means that* $\forall i\ |\Omega \cap \Xi_i| = |\Omega \cap \Xi_i|$, *which implies that* $al(\Omega) = \sum_{i=1}^{r} |\Omega \cap \Xi_i|(\sum_{j=i+1}^{r} al(\Xi_j)+1) = \sum_{i=1}^{r} |\Omega' \cap \Xi_i|(\sum_{j=i+1}^{r} al(\Xi_j)+1) = al(\Omega')$. *The fact that* $al(\Omega) = al(\Omega')$ *also contradicts our initial assumption* $al(\Omega) > al(\Omega')$. *Thus, we conclude that* $al(\Omega) > al(\Omega') \Rightarrow \Omega \succ_{\Omega}^{*} \Omega'$.

$\Omega \succ_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) > al(\Omega')$ *suffices to prove the theorem: Note that we have proved that* $\Omega \succ_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) > al(\Omega')$, *then it trivially follows that* $\Omega \prec_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) < al(\Omega')$. *And these two cases imply that* $\Omega \sim_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) = al(\Omega')$. *Finally,* $\Omega \succ_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) > al(\Omega')$ *and* $\Omega \sim_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) = al(\Omega')$ *imply that* $\Omega \succeq_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega) \geq al(\Omega')$, *which ends the proof of the theorem.* □

Since the value alignment function embodies the $\succeq_{\Omega}^{*}$ ranking, we can cast the most value-aligned norm system selection problem as an optimisation problem to maximise value alignment as follows.

Problem 2. *Given a norm net* $\langle N, R \rangle$, *a value system* $\langle V, \succeq \rangle$, *and the value alignment function al, the problem of finding the sound norm system with maximum value alignment amounts to solving:*

$$\Omega_{max} = \underset{\Omega \in \mathcal{P}(N),\ \Omega\ sound}{\arg\max}\ al(\Omega) \quad (7)$$

In order to solve the optimisation problem defined by Problem 2, we will encode it as a linear program (LP). The first step is to build the objective function of the LP. The challenge here is to compactly represent norm systems. Notice that for $N = \{n_1, n_2, n_3\}$, norm system $\Omega = \{n_1, n_2\}$ can be represented as $\{n_1, n_2, \neg n_3\}$, or as a binary vector $(1, 1, 0)$. In general, any norm system $\Omega$ can be encoded as a vector $(x_1, \dots, x_{|N|})$, where $x_i \in \{0, 1\}$ is the decision variable for norm $n_i \in N$: if $x_i = 1$ means that norm $n_i$ is in $\Omega$, while $x_i = 0$ means $n_i$ is not in $\Omega$. Using the $(x_1, \dots, x_{|N|})$ encoding for norm systems and following equation 6, in general we can obtain the value alignment of a norm system as $\sum_{i=1}^{r}(\sum_{n_w \in \Xi_i} x_w)(\sum_{j=i+1}^{r} al(\Xi_j) + 1)$, making use of the fact that $|\Omega \cap \Xi_i| = \sum_{n_w \in \Xi_i} x_w$. Therefore, solving Problem 2 amounts to finding the assignment of variables $(x_1, \dots, x_{|N|})$ representing a sound norm system with maximum value alignment. For that, we propose to solve the following LP:

$$\max \sum_{i=1}^{r} (\sum_{n_w \in \Xi_i} x_w)(\sum_{j=i+1}^{r} al(\Xi_j) + 1) \quad (8)$$

We require that the selected norm system complies with the sound norm system properties. Thus, we translate the requirements in Definition 3 into the following constraints:

- Mutually exclusive (incompatible) norms cannot be selected at once:

$$x_i + x_j \leq 1 \text{ for each } (n_i, n_j) \in R_x \quad (9)$$

- Substitutable norms cannot be jointly selected as these norms are interchangable:

$$x_i + x_j \leq 1 \text{ for each } (n_i, n_j) \in R_s \quad (10)$$

- A norm cannot be simultaneously selected with any of its ancestors:

$$x_i + x_k \leq 1 \text{ for each } n_k \in A(n_i) \ \ 1 \leq i \leq |N| \quad (11)$$

- If a norm has more than one direct successor (we note $\bar{S}(n) = \{n' \in N, (n, n') \in R_g\}$), these direct successors cannot be simultaneously selected:

$$\text{If } |\bar{S}(n)| > 1 \text{ then } \sum_{n_j \in \bar{S}(n)} x_j < |\bar{S}(n)| \quad \text{for each } n \in N \quad (12)$$

Observe that our LP employs $|N|$ decision variables and avoids the expensive, explicit computation of the norm system ranking required by the preference lifting process.

Note that in this section we have posed our norm selection problem as a (quantitative) optimisation problem, while in Section 3 we defined the most value-aligned norm system selection problem as a qualitative optimisation problem. Hence, now we have to prove that both approaches yield equivalent solutions.

Theorem 3. *Given a norm net* $\langle N, R \rangle$ *and a value system* $\langle V, \succeq \rangle$, *the solution to the most value-aligned norm system selection problem* $(\Omega^*)$ *and the solution to the problem of finding the norm system with maximum value alignment* $(\Omega_{max})$ *are the same, namely* $\Omega^* = \Omega_{max}$. *In case the solution is not unique, then* $\Omega^* \sim_{\Omega}^{*} \Omega_{max}$.

Proof. 5. *From Theorem 2, we know that* $\forall \Omega, \Omega' \in \mathcal{P}(N)$, $\Omega' \succeq_{\Omega}^{*} \Omega \Leftrightarrow al(\Omega^*) \geq al(\Omega)$. *Since* $\Omega^*$ *is the most value-aligned sound norm system, we have that* $\forall \Omega'$ *sound,* $\Omega^* \succeq_{\Omega}^{*} \Omega' \Leftrightarrow al(\Omega^*) \geq al(\Omega')$. *Hence,* $\Omega^*$ *is a sound norm system of maximum value alignment, which is exactly the definition of* $\Omega_{max}$ *(Equation 7). Thus, if the solution is unique* $\Omega^* = \Omega_{max}$, *and if it is not unique we have that* $al(\Omega) = al(\Omega') \Leftrightarrow \Omega^* \sim_{\Omega}^{*} \Omega_{max}$ *as a consequence of Theorem 2.* □

Having shown that the solution of Problem 1 and Problem 2 are equivalent we now tackle the resolution of our running example:

Example 8. *Examples 2, 3, 4, 5, 6 and 7 detail how to build a norm system ranking. Note though, that even in this case with only 7 norms, the number of norm systems to rank is* $2^7 = 128$. *Next we* *provide the LP encoding for our running example. First, we will build our objective function. Since the quotient order is $\Xi_1 \succ_n \Xi_2$, we first compute $al(\Xi_2) = |\Xi_2| = 6$ (because $\Xi_2 = \{n_2, n_3, n_4, n_5, n_6, n_7\}$); and from that $al(\Xi_1) = |\Xi_2| \cdot (al(\Xi_2) + 1) = 7$ (because $\Xi_1 = \{n_1\}$). Therefore, the objective function (following Equation 8) is* $\max 7x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7$. *Since the norms of our running example have some relations between them, as shown in Figure 1, we consider the following constraints:*

- *Regarding exclusive norms:* $x_1+x_2 \leq 1, x_1+x_3 \leq 1, x_1+x_4 \leq 1, x_1 + x_5 \leq 1, x_1 + x_6 \leq 1$ *and* $x_1 + x_7 \leq 1$.
- *Regarding substitutable norms:* $x_2 + x_3 \leq 1$.
- *Regarding the relation between successors and ancestors:* $x_5 + x_6 \leq 1$ *and* $x_5 + x_7 \leq 1$.
- *Finally, for direct successors we have to constrain:* $x_6 + x_7 < 2$.

*With this encoding the solution to the LP, the most value-aligned norm system, is $\{n_1\}$. Note though that Example 7 showed that $\{n_1, n_2, n_4\} \succeq^*_\Omega \{n_1\}$, but since $\{n_1, n_2, n_4\}$ is not sound (because $n_1$ and $n_2$ (or $n_4$) are mutually exclusive), it is not a feasible solution.*

## 5 DISCUSSION

Different value rankings may vary the selection of the most value-aligned norm system. In previous examples we solved the airport problem in Figure 1 by considering the value ranking $v_{fm} \succeq v_{saf}$. This section explores how the solution changes for alternative value rankings ($v_{fm} \preceq v_{saf}$, $v_{fm} \sim v_{saf}$). We will also compare the approach above to the utilitarian optimisation approach in [18].

### 5.1 Safety preferred to free movement

If $v_{fm} \preceq v_{saf}$, preference induction produces the ranking $\{n_2, \ldots, n_7\} \succeq_\Omega \{n_1\}$. Grounding this ranking, we obtain the norm ranking $n_2 \sim_n \ldots \sim_n n_7 \succeq_n n_1$. As in Example 7, we depict how to compute the preference between a few illustrative norm systems: $\{n_1\}$, $\{n_1, n_2, n_4\}$, $\{n_2, n_4, n_5\}$, $\{n_3, n_4, n_5\}$. In this case we have $\Xi_1 = \{n_2, n_3, n_4, n_5, n_6, n_7\}$ and $\Xi_2 = \{n_1\}$. Therefore, we have $\eta(\{n_1\}) = (0, 1)$; $\eta(\{n_1, n_2, n_4\}) = (2, 1)$; $\eta(\{n_2, n_4, n_5\}) = (3, 0)$; $\eta(\{n_3, n_4, n_5\}) = (3, 0)$.

Since $(3, 0) \geq_L (2, 1) \geq_L (0, 1)$, we have the following ranking over norm systems: $\{n_3, n_4, n_5\} \sim^*_\Omega \{n_2, n_4, n_5\} \succeq^*_\Omega \{n_1, n_2, n_4\} \succeq^*_\Omega \{n_1\}$. After encoding as an LP and subsequently solving the norm selection problem considering $v_{saf} \succeq v_{fm}$, we would either obtain $\{n_2, n_4, n_5\}$ or $\{n_3, n_4, n_5\}$, since they are equally value-aligned.

### 5.2 Safety and free movement equally valued

With the value ranking $v_{fm} \sim v_{saf}$, namely $v_{fm}$ and $v_{saf}$ are equally preferred, the ranking resulting from preference induction would be: $\{n_1\} \sim_\Omega \{n_2, \ldots, n_7\}$. Grounding this ranking, we obtain the norm ranking: $n_1 \sim_n \ldots \sim_n n_7$. As before, we show how to compute a ranking over a few norm systems, namely $\{n_1\}$, $\{n_1, n_2, n_4\}$, $\{n_2, n_4, n_5\}$, $\{n_3, n_4, n_5\}$. In this case, since all norms are indifferently preferred, we only have one equivalence class, $\Xi_1 = \{n_1, \ldots, n_7\}$, which leads us to define, for any norm system $\Omega$, $\eta(\Omega) = |\Omega|$. Hence, it follows that: $\eta(\{n_1\}) = 1$ and $\eta(\{n_1, n_2, n_4\}) = \eta(\{n_2, n_4, n_5\}) = \eta(\{n_3, n_4, n_5\}) = 3$.

Since $3 \geq_L 1$, we would have the preferences: $\{n_3, n_4, n_5\} \sim^*_\Omega \{n_2, n_4, n_5\} \sim^*_\Omega \{n_1, n_2, n_4\} \succeq^*_\Omega \{n_1\}$. After encoding as an LP and subsequently solving the norm selection problem considering $v_{saf} \sim v_{fm}$, we either obtain $\{n_2, n_4, n_5\}$ or $\{n_3, n_4, n_5\}$.

### 5.3 Differences between qualitative and utility-based norm selection

Next we compare the qualitative norm selection process with the quantitative approach described in [18]. In broad terms, the method in [18] assigns some utility to each value. The utility of a value is obtained by adding up the utilities of its immediately less preferred values in the value system plus 1. For instance, considering the case in Figure 1, following [18], and $v_{fm} \succeq v_{saf}$ as preferences over values in the value system, would result in the following utilities: $u(v_{saf}) = 1$, $u(v_{fm}) = u(v_{saf}) + 1 = 2$. From this, the utility of a norm is computed as the sum of the utilities of the values it supports, and the utility of a norm system is the sum of the utilities of the norms it contains. As to the case in Figure 1, we would obtain that $u(n_1) = 2$, and the rest of norms would get utility 1. The norm system $\{n_3, n_4, n_5\}$ would get utility 3 while $\{n_1\}$ would get utility 2. Thus, by design, the method in [18] would select norm system $\{n_3, n_4, n_5\}$ supporting value $v_{saf}$. But this is against the preferences in the value system, since we have specified that $v_{fm} \succeq v_{saf}$. Instead, following the method in this paper (shown in Example 8), the solution would be $\{n_1\}$, which supports $v_{fm}$ instead. Therefore, we conclude that the quantitative approach in [18] does not necessarily lead to the most value-aligned norm system since it tends to bias the selection towards values being supported by a large number of norms. This is a consequence of the ad-hoc nature of the utility formulas of [18], whereas our encoding is derived from the ranking $\succeq^*_\Omega$ (Theorem 2), which we have obtained through a process grounded on social choice literature and clear axioms.

## 6 CONCLUSIONS

While there has been previous work on selecting norms in regard to the moral values they support, this paper explores norm selection following a qualitative approach. In [17, 18] we focused on translating the available qualitative information into quantitative one and then we selected those norms that maximised an ad-hoc utility formula. In this work, we take full potential of the qualitative preferences in hand, by transforming them from value preferences to norm preferences and finally to norm system preferences. Additionally, we base these transformations in a novel method presented in recent literature [6]. Although theoretically, obtaining a ranking of all norm systems is possible, we noticed that as the number of norms considered grows, this task becomes exponentially complex. Therefore, we provide an encoding that allows us to find the most value-aligned norm system without the burden of building the whole ranking. We prove that the outputted norm system of the encoding is indeed the most value-aligned norm system in the whole ranking satisfying soundness.

We divide our future work in two fronts. First, in terms of norm selection, here we have assumed that norms only promote values, and hence considering demotion remains future work. Second, in terms of ranking theory, we can further research the generalised subset selection problem and also study anti-lex-cel's properties as we think anti-lex-cel's lifted rankings follow a rule more general than leximax [3, 14] but more specific than $K$-leximax [10].

## REFERENCES


[1] Robert Axelrod. 1986. An evolutionary approach to norms. *American political science review* 80, 4 (1986), 1095–1111.

[2] Ofer H Azar. 2004. What sustains social norms and how they evolve?: The case of tipping. *Journal of Economic Behavior & Organization* 54, 1 (2004), 49–64.

[3] Salvador Barberà, Walter Bossert, and Prasanta K Pattanaik. 2004. Ranking sets of objects. In *Handbook of utility theory*. Springer, 893–977.

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

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

[6] Giulia Bernardi, Roberto Lucchetti, and Stefano Moretti. 2019. Ranking objects from a preference relation over their subsets. *Social Choice and Welfare* 52, 4 (01 Apr 2019), 589–606. https://doi.org/10.1007/s00355-018-1161-1

[7] Sven Ove Hansson. 2001. *The structure of values and norms*. Cambridge University Press.

[8] Adrian Haret, Hossein Khani, Stefano Moretti, and Meltem rk. 2018. Ceteris paribus majority for social ranking. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*. International Joint Conferences on Artificial Intelligence Organization, 303–309. https://doi.org/10.24963/ijcai.2018/42

[9] Hossein Khani, Stefano Moretti, and Meltem rk. 2019. An Ordinal Banzhaf Index for Social Ranking. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 378–384. https://doi.org/10.24963/ijcai.2019/54

[10] Takashi Kurihara. 2017. Leximax and leximin extension rules for ranking sets as final outcomes with null alternatives. *Available at SSRN 3009073* (2017).

[11] 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.

[12] 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.

[13] Javier Morales, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, Michael Wooldridge, and Wamberto Vasconcelos. 2013. Automated synthesis of normative systems. In *AAMAS 2013*. 483–490.

[14] Prasanta K Pattanaik and Bezalel Peleg. 1984. An axiomatic characterization of the lexicographic maximin extension of an ordering over a set to the power set. *Social Choice and Welfare* 1, 2 (1984), 113–122.

[15] Ganesh Ram Santhanam. 2016. Qualitative optimization in software engineering: A short survey. *Journal of Systems and Software* 111 (2016), 149–156.

[16] Sandip Sen and Stéphane Airiau. 2007. Emergence of Norms through Social Learning. In *IJCAI*. 1507–1512.

[17] 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. In *Proceedings of the 1st Conference on artificial intelligence, ethics and society (AIES'18)*. 1–7. https://doi.org/10.1145/3278721.3278735

[18] 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 *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'18)*. International Foundation for Autonomous Agents and Multiagent Systems, 1294–1302.

[19] Yoav Shoham and Moshe Tennenholtz. 1997. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence* 94, 1-2 (1997), 139–166.

[20] Toshiharu Sugawara. 2011. Emergence and Stability of Social Conventions in Conflict Situations. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, Toby Walsh (Ed.). IJCAI/AAAI, 371–378. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-071