Runtime Revision of Sanctions in Normative Multi-Agent Systems

Davide Dell'Anna ✉ · **Mehdi Dastani** ·
Fabiano Dalpiaz

Abstract To achieve system-level properties of a multiagent system, the behavior of individual agents should be controlled and coordinated. One way to control agents without limiting their autonomy is to enforce norms by means of sanctions. The dynamicity and unpredictability of the agents' interactions in uncertain environments, however, make it hard for designers to specify norms that will guarantee the achievement of the system-level objectives in every operating context. In this paper, we propose a runtime mechanism for the automated revision of norms by altering their sanctions. We use a Bayesian Network to learn, from system execution data, the relationship between the obedience/violation of the norms and the achievement of the system-level objectives. By combining the knowledge acquired at runtime with an estimation of the preferences of rational agents, we devise heuristic strategies that automatically revise the sanctions of the enforced norms. We evaluate our heuristics using a traffic simulator and we show that our mechanism is able to quickly identify optimal revisions of the initially enforced norms.

1 Introduction

Multiagent systems (MASs) comprise autonomous agents that interact in a shared environment [57]. To achieve the system-level objectives of a MAS, the behavior of the autonomous agents should be controlled and coordinated [12]. For example, a smart traffic system is a MAS that includes autonomous agents like cars, traffic lights, etc. The objectives of the system include avoiding the occurrence of traffic jams as well as minimizing the number of accidents.

D. Dell'Anna, M. Dastani, F.Dalpiaz
Utrecht University
Utrecht, The Netherlands
Email: {d.dellanna, m.m.dastani, f.dalpiaz}@uu.nl

One way to control the behavior of the agents in a MAS without limiting their autonomy is norm enforcement [1, 49]. Norm enforcement via sanctions is traditionally contrasted with norm regimentation; the latter alternative prevents the agents from reaching certain states of affairs. For example, in a smart traffic system, a regimentation strategy is to close a road to prevent cars from entering that road, while a sanctioning strategy is to impose sanctions on cars that drive through the road.

Due to the dynamicity and unpredictability of the behaviours of interacting agents in uncertain environments, it is difficult for the designers who engineer a MAS to specify norms that, when enforced, will guarantee the achievement of system-level objectives in every operating context. To cope with this issue, the enforced norms need to be revised at runtime. Existing research has investigated the offline revision of the enforced norms [3], proposed logics that support norm change [4, 34, 35], and examined the legal effects of norm change [29].

In [24], we proposed a framework for engineering normative MASs that, using observed data from MAS execution, revises the norms in the MAS at runtime to maximize the achievement of the system objectives. In that work, we made the simplistic assumption that norms are regimented and we introduced algorithms for switching among alternative predefined norms. In [25] we extended the framework to support the revision of norm enforced via *sanctioning*. In addition to observed data from MAS execution, we used an estimation of the preferences of the agents to guide the runtime norm revision. However, we considered MASs where only one norm at a time was enforced.

In this paper, we significantly extend our previous work by supporting MASs where multiple norms are enforced. We formalize different types of rational agents that behave according to their preferences and we disccus their properties. We use Bayesian Networks to learn the norm effectiveness from data observed from MAS execution and to inform the runtime norm revision mechanism that revises the sanctions of multiple norms.

The contributions of this paper are as follows:

— We provide a formal definition of different types of rational preferences of agents, specified in terms of desired states of affairs and the maximum payment that the agent is willing to make to achieve such states of affairs. We prove that such preferences satisfy the basic rationality requirements [38].
— We build on and extend the general architecture proposed in [24, 25], and study in detail the relationships between estimated agents' preferences, sanctions, and system-level objectives. We use a framework where the normative MAS is flanked by a norm monitoring and enforcement component, and we introduce a norm revision component that uses observed data from MAS execution and an estimation of agents' preferences to modify norm sanctions at runtime.
— We propose six heuristic strategies for the revision of multiple norms that leverage probabilistic information learned from observed data from MAS execution and an estimation of the preferences of agents.

– We report on an evaluation through a traffic simulator that shows the effectiveness and efficiency of our revision strategies in identifying optimal sanctions for multiple norms.

Organization. Section 2 reports on related work. Section 3 presents our framework to characterize norms and agents' preferences. Section 4 explains the overall approach for the supervision of normative MAS based on probabilistic reasoning over norm effectiveness and agents' preferences. Section 5 introduces six strategies for revising norms by combining agents' preferences with the achievement of the system-level objectives. Section 6 evaluates our work through simulation experiments. Section 7 discusses the results and the assumptions, limitations and future directions of our work. Section 8 presents our conclusions.

2 Related Work

In the MAS literature, norms have been proposed as a way to regulate the behavior of the agents in order to achieve system-level properties without limiting the autonomy of the agents [1, 49, 52].

Many approaches focus on the design-time construction of robust normative MASs. Several techniques enable proving the correctness of normative systems through the model checking of formulas that describe liveness or safety properties [2, 23, 33]. These works are useful for the initial design of a MAS, but they cannot cope with the runtime unpredictability of the system that stems from the autonomy and heterogeneity of the agents.

In order to successfully supervise and regulate dynamic MASs, researchers have studied the revision of norms. Some frameworks formalize norm dynamics thereby allowing the assessment of the impact of norms on the specification of a MAS, i.e., whether the designed MAS will be norm compliant. Aucher *et al.* [4] introduce a dynamic context logic to describe the operations of contraction and expansion of theories that occur when removing or adding new norms. Governatori *et al.* [29] investigate how the application of theory revision leads to legal abrogations and annulments. Knobbout *et al.* [35] propose a dynamic logic to characterize the dynamics of state-based and action-based norms. Both in Knobbout's work [34, 35] and in Alechina *et al.*'s approach [2], norm change is restricted to norm addition. This family of approaches focus on the impact of revising a norm on an existing normative system. In this paper, instead we study the relationship at runtime between the enforced norms and the achievement of system-level objectives, and suggest mechanisms to determine *how* to revise the (sanctions of the) current norms.

Jiang *et al.* [31] discuss the contextualization of norms. They explicitly represent the context of application of a norm and they use such context to organize norms during the design of a MAS. In our work, we also enforce different norms in different contexts. Unlike them, however, we determine the most appropriate context for different norm sets at runtime and based on observed data from MAS execution.

Miralles *et al.* [39] present a framework for the adaptation of MAS regulations at runtime. Their approach is complementary to ours. They represent conditional norms via norm patterns and describe an adaptation mechanism based on case-based reasoning. Adaptation is performed at runtime individually by a number of assistant agents and then, via a voting mechanism, a final adaptation is approved. The decision on how to adapt norms is taken based on similar previously seen cases. In their work, however, they do not consider sanctions. In our work, we focus on the revision of sanctions, we perform norm revision through a centralized component, and we make use of an estimation of agents' preferences to guide norm revision.

Cardoso *et al.* [13] present a framework for the runtime adaptation of sanctions associated with obligations. In their work, they assume that norm violations are bad for the system-level objectives. In our work, we relax such assumption, as agents ability to violate norms can be useful [14]. We evaluate the effectiveness of a norm at runtime based on observed data from MAS execution. Furthermore, they assume that the strength of a sanction should be directly proportional to its application frequency, and they constantly try to lower sanctions in order to give agents maximum autonomy. In our work, we base the revision of norms on an estimation of the preferences of the agents, and we determine the appropriate value of their sanctions based on the relationship between obedience of norms and achievement of system-level objectives determined at runtime.

In MASs, agents' preferences have been mainly used as a way to choose at runtime between different plans or actions to execute [21, 32, 42, 53]. Preferences are usually interpreted as constraints that, if satisfied by a certain plan (or action), increase the desirability of executing such plan (or action). Formal languages have been proposed and used for expressing preferences (e.g., $\mathcal{LPP}$ [6, 10] or *LTL* [12]). In this work we focus on strategies for sanctions' revision. For this reason, we make use of a high-level representation of preferences, without restricting ourselves to, but supporting, any specific language. In particular, we consider preferences that satisfy the basic rationality requirements [38] and order different alternative states of affairs that agents may desire to achieve. Our agents are rational and norm-aware [43], in the sense that they always try to aim at the most preferred state of affairs for which they have enough budget, taking also into account the possible sanctions they would incur when violating some of the enforced norms. Furthermore, our agents are autonomous, in the sense that they are able to make decisions without the intervention of human users but in line with their preferences [5, 22]. As we aim to investigate the process of norm revision, we assume that we have an accurate estimation of the agents' preferences. In future work, we can relax this assumption and investigate norm revision based on inaccurate estimations of the agents' preferences.

Chopra *et al.* [17] study how agents' preferences—expressed in terms of goals—interact with norms—represented as commitments. In particular, they propose a framework for the agents to adapt their behavior. We take an orthogonal approach, for we study how to change the norms without altering the

agent construction. In particular, we study how to alter the sanctions used to enforce the norms on the agents, so to guarantee at runtime the system-level objectives. Our proposed mechanisms, therefore, relate also to the idea of adjustable autonomy [40]. The proposed runtime mechanism of revision of the sanctions of the norms can be seen as an automated mechanism to adjust the decisions' options of the agents (thus their degree of autonomy) so to maximize the objectives of the system and its operators.

Cranefield *et al.* [19] present a Bayesian approach to norm identification. They show that agents can internalize norms that exist in an environment, by learning from the behavior that complies with or violates certain norms. This work is a valuable addition to ours, for it shows that it is possible for agents to learn norms even when they are not explicitly communicated to them.

Tumer *et al.* [50] use multi-agent reinforcement learning in a smart traffic simulation to determine the behavior of the car agents that maximizes the utility of the city designer and of the individual agents. Their interesting work focuses on regimentation; instead, we focus on enforcement that does not violate agents' autonomy.

3 Normative Multiagent Systems

This section presents a generic framework for specifying normative multiagent systems in which the agents behave in line with their preferences while norms are enforced on them via sanctions. This framework allows us to analyze the interplay between norms and agents' preferences in normative multiagent systems.

3.1 Illustrative example

Consider the two-lanes ring road depicted in Fig. 1. In a ring road, a population of vehicles moves continuously in a circle. Every vehicle is autonomous and acts according to its own preferences. For example, vehicles have preferences about, among other things, their speed, based on which they determine their willingness to risk sanctions for violating traffic norms. Such preferences and their corresponding willingness to risk sanctions allow the vehicles to autonomously decide when and how to accelerate or decelerate or to change lane. If a fast vehicle is using the outer line and a slower vehicle blocks its way, the fast vehicle may move to the inner line to overtake the slow vehicle. Since all vehicles share the same environment, their local decisions have an effect on the (emergent) system-level behavior of the vehicles driving on the ring road [48]. For example, based on contextual factors such as the density of vehicles on the ring road, the vehicles' behavior may provoke traffic jams and the average speed may vary, as well as the average time to complete a loop of the ring road. The ring road is a simple example of a MAS. Although far from realistic traffic situations, the ring road illustrates the fundamental
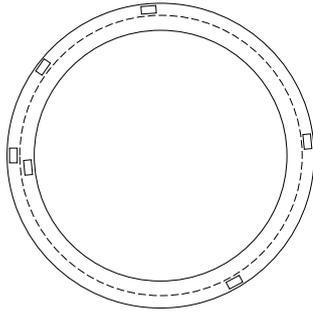
Fig. 1: Two lanes ring road. Rectangles are vehicles, moving in counter clockwise direction.

phenomena of emergent system-level properties, caused by the local decisions of individual agents, and the importance of mechanisms to control and steer such system-level behaviors.

We assume that the main stakeholder of the ring road (the city council) has two system-level objectives: *to minimize the average time to complete a loop of the ring road* and *to minimize the number of halted cars*. Despite interdependence, the stakeholder desires to evaluate the two objectives independently due to their distinct nature. We consider two contextual variables that may influence the achievement of the system-level objectives, together with the vehicles' behavior: the *density of vehicles* and the *presence of an obstacle* on the ring road. The higher is the density of the vehicles on the ring road, the higher is the risk of breaking waves and slowdowns. The presence of an obstacle may force vehicles to halt and wait for an adequate moment to take over the obstacle. If the density of vehicles on the ring road is high enough, this may also cause queues after the obstacle. To achieve the objectives, the behavior of the agents is regulated by enforcing norms concerning (i) the speed limit, such as the norm *every vehicle on the ring road shall not exceed a speed of 50km/h, otherwise it will receive a sanction of 100€*, and (ii) the minimum safety distance between cars, such as the norm *every vehicle on the ring road shall keep a minimum distance of 2m, otherwise it will receive a sanction of 20€*. Regulating speed and safety distance of the cars on the ring road is expected to help achieving the system-level objectives in the traffic contexts represented by the contextual variables. A car that keeps a sufficient safety distance from the car ahead, is less likely affected by sudden deceleration of the car ahead. More space between cars may also favour surpasses of slow cars when necessary. An opportune safety distance, together with opportune cars speed, may reduce jams in the presence of obstacles or the effect of breaking waves.

The ring road above described is a normative MAS. Vehicles are autonomous agents, each acting according to their own preferences. Each agent belongs to an agent type that can be characterized by the agent's preferences. For instance a *cautious* agent is a type of agent that prefers to go slow rather than

fast on the ring road and prefers to maintain the appropriate safety distance. A *brave* agent is a type of agent that prefers to go fast rather than slow, and to approach cars closer than the minimum safety distance, even if it has to pay some money to do so.

3.2 Norms

The focus of this paper is the runtime revision of the sanctions of the norms enforced in the MAS. In order to focus on this aspect, we propose a simple but extensible language for norms. Consider a set of propositional atoms $L = \{p_1, \ldots, p_k\}$, each representing a fact that can hold or not in a system state[1] (e.g., propositional atom sp_{100} indicates that *the speed of a vehicle on the ring road is $\leqslant 100$ km/h*).

Let $AL = (L_1, ..., L_n)$ be an ordered list of n disjoint subsets of L, s.t. L_i contains atoms related to an aspect i of the system[2] (e.g., $L_i = \{sp_{100}, sp_{50}\}$, in the ring road scenario, contains atoms related to the *speed of the cars*).

We consider a norm as a pair $N = (p, s)$, where $p \in L$ and $s \in \mathbb{N}$, indicating that p should hold in the current system state for all agents, otherwise sanction s will incur. For instance, a norm $N = (sp_{50}, 100)$ indicates that *every vehicle on the ring road shall not exceed a speed of 50km/h, otherwise it will receive a sanction of 100€*.

In the following we consider an ordered set of norms $\mathcal{N} = \langle N_1, N_2, \ldots, N_n \rangle$ and assume that (i) norms are non-conflicting, i.e., obeying a norm N_i does not prevent an agent from obeying or violating any other norm in $\mathcal{N}$; and (ii) each norm regulates a different aspect of the system, so that the i-th norm $N_i = (p, s)$ in $\mathcal{N}$ is a pair where $p \in L_i$ (with L_i i-th set in AL) and $s \in \mathbb{N}$. For instance, if $AL = (L_1, L_2)$, $L_1 = \{sp_{50}, sp_{100}\}$ and $L_2 = \{dist_1, dist_2\}$, then $N_1 = (sp_{50}, 100)$ is a norm concerning the speed limit and $N_2 = (dist_2, 100)$ is a norm concerning the minimum safety distance.

Note that, despite these assumptions, norms can still influence each others by means of the behavior that they cause on the agents. For instance, if the density of vehicles on the ring road is high, in order to obey a norm concerning the minimum safety distance from the car ahead, an agent may need to decrease its speed, therefore obeying also a norm concerning the maximum speed limit. We distinguish, however, such influence from the concept of conflict, in the sense that the norm concerning the minimum safety distance does not prevent, *a priori*, an agent to either obey or violate the norm concerning the maximum speed limit, and vice-versa.

---

[1] A system state is assumed to consist of the state of individual cars (e.g., speed and position of the cars) as well as the state of the environment (e.g., density of vehicles in the ring road).

[2] We use the term *aspect* to indicate any particular characteristics of the behavior of the agents, such as the speed of the cars, that is both monitorable by an organization that enforces norms in the MAS, and over which agents have control.

3.3 Rational Agents and Their Preferences

In MASs, agents are often assumed to be autonomous and possibly hetero-geneous. Moreover, it is common to assume that the internal states of the agents such as their beliefs, preferences, and decision making mechanisms are unknown or partly known to other agents or to the institutions that regulate their behaviour. In line with the theory of economic rationality [38], in this paper we consider rational agents that behave according to their rational pref-erences, which determine an ordering between different alternative states of affairs (simply *alternatives* in the following). A rational agent aims to achieve its most preferred states of affairs: when a rational agent believes it is possi-ble to achieve a certain state of affairs s, the agent will never aim to achieve states of affairs that are less preferred than s. For example, a *cautious* agent that prefers to go slow on the ring road and maintain appropriate safety dis-tance, may be less prone to surpass other cars or to change lane, and may exhibit more moderate acceleration or deceleration than less cautious agents. The behavior of such a cautious agent, however, can vary significantly, based on contextual conditions. For example, a sudden break from the car ahead may force also the cautious agent to brusquely decelerate.

In this work, we assume we have an estimation of the preferences of the agents concerning the n different aspects of the system that we aim to regu-late by a norm, as per Sec. 3.2. In the rest of the paper, when we refer to the preferences of the agents, we refer therefore to such an estimation of their pref-erences. We do not assume access to the agents' internals such as their beliefs or their preferences regarding other aspects of the system (e.g., information about fuel reserve or the preference on road types). Having an estimation of the preferences of the agents should not be seen as a violation of the auton-omy of agents or access to their internals. Having some knowledge of agents' preferences is realistic in most MAS settings. For example, in some coopera-tive settings, agents may be requested to declare their true preferences prior entering the system and agents can autonomously decide whether to join or not, while in other settings the preference of agents can be learned from their behaviors [8]. Note that we do not focus on the process of preference elicita-tion, which is essential for deriving and formulating agents' preferences, but beyond the scope of this paper. Several techniques for the elicitation of pref-erences have been proposed in literature, including both automated methods and methods that directly involve the end-user (see for example [11, 16, 46]). Here, we rely on such techniques and we just assume that some relevant part of agents' preferences is already given or estimated.

We represent the alternatives over which the agents have preferences as lists of pairs such as $(\langle p_1, b_1 \rangle, \ldots, \langle p_n, b_n \rangle)$, indicating that for a state of affairs where $p_1, ...,$ and p_n hold, the agent is willing to spend, if necessary, a budget b_1 to achieve p_1, a budget b_2 to achieve p_2, etc. We focus on finite preferences, therefore we constrain the budgets expressed in the alternatives to be member of a budget set $\mathcal{B} \subset \mathbb{N}$.

We denote by $Pref(a) = (A, \geq)$ the preference of an agent $a \in Ag$, where $Ag = \{a_1, \ldots, a_n\}$ is a set of agents, A is a set of alternatives defined as per Def. 1, and $\geq$ is a partial order on A. We write $x \geq y$ to denote the fact that the agent either prefers alternative x to alternative y or is indifferent between x and y.

Definition 1 (Preference Alternatives) Let $AL = (L_1, \ldots, L_n)$ be a list as per Sec. 3.2. Given a set of budget lists $BL \subseteq \mathcal{B}^n$ (with $\mathcal{B}^n$ the n-ary Cartesian power of $\mathcal{B}$), the set of alternatives A is the set $\{ (\langle p_1, b_1 \rangle, \ldots, \langle p_n, b_n \rangle) \mid p_i \in L_i \ \& \ (b_1, \ldots, b_n) \in BL \}$.

Notation. Before continuing, we provide here a summary of the notation that we will use in the rest of the paper in the context of preferences. Given a preference $Pref(a) = (A, \geq)$, an alternative $x = (\langle p_1, b_1 \rangle, \ldots, \langle p_n, b_n \rangle) \in A$, and a set of budget lists $BL \subseteq \mathcal{B}^n$, we call:

- $prop(x) = (p_1, ..., p_n)$, the list of propositional atoms in x
- $bud(x) = (b_1, ..., b_n) \in BL$, the list of budgets associated to each propositional atom in x.
- $req\_bud(x) = \sum_{b \in bud(x)} b$ the budget required by alternative x (*required budget*, in the following), i.e., the sum of all budgets in x.
- $x[B']$ a new alternative $x' = (\langle p_1, b_1' \rangle, ..., \langle p_n, b_n' \rangle)$ with same propositional atoms as x, but using budgets $B' = (b_1', \ldots, b_n') \in BL$ instead of budgets $B = (b_1, \ldots, b_n)$.

Furthermore, in the rest of the paper, unless specified otherwise, when we provide an example concerning preferences or norms, we make use of L defined as the set $\{sp_{50}, sp_{100}, dist_1, dist_2\}$ with $AL = (L_1, L_2)$ and $L_1 = \{sp_{50}, sp_{100}\}$ and $L_2 = \{dist_1, dist_2\}$ so that $\mathcal{N} = \langle N_1, N_2 \rangle$ with L_1 related to N_1 (norm concerning speed limit) and L_2 related to N_2 (norm concerning safety distance), and we use n to indicate the number of norms in $\mathcal{N}$.

In the following we define the types of preferences that we consider in this paper. We first define two basic types of preferences. Then, after providing some examples of such preferences, we define more complex preferences that combine the two basic types.

3.3.1 Basic Preferences

We define here two types of basic preferences. The first kind of preference orders the alternatives based on their budgets, while the second type orders the alternatives based on the propositional atoms (i.e., states).

Definition 2 (Basic Preference) Given a list $AL = (L_1, ..., L_n)$ and a set $BL \subseteq \mathcal{B}^n$, an agent is said to have a *basic preference* $(A, \geq)$ when for all alternatives x and y in A, the partial order $\geq$ satisfies one of the following two clauses:

a. $x \geq y$ iff
   $req\_bud(x) \leqslant req\_bud(y)$ &
   $\forall v, w \in A, \forall B, B' \in BL : v[B] > w[B] \Rightarrow v[B'] > w[B']$

b. $x \geq y$ iff
   if $prop(x) = prop(y)$ then $req\_bud(x) \leqslant req\_bud(y)$
   else $\forall B, B' \in BL : x[B] \geq y[B']$

In the rest of the paper, we write $x \sim y$ when $x \geq y$ and $y \geq x$. We write $x > y$ when $x \geq y$ but not $y \geq x$.

If an agent's preference adheres to Def. 2a, then the required budget determines the order of the alternatives. In particular, Def. 2a determines a preference where alternatives that require a lower budget are preferred to alternatives that require higher budget (first condition of Def. 2a) and the relative order between two alternatives with different propositional atoms is the same for all possible budgets (second condition of Def. 2a). Note that in a basic preference that adheres to Def. 2a, two alternatives x and y such that $req\_bud(x) > req\_bud(y)$ cannot be equally preferred. In fact, if $x \sim y$ we have that $req\_bud(x) \leqslant req\_bud(y)$ and $req\_bud(y) \leqslant req\_bud(x)$. As a consequence, all alternatives with required budget 0 are strictly preferred to all the other alternatives, and all alternatives with same required budget are equally preferred.

If an agent's preference adheres to Def. 2b, then the propositional atoms determine the order of the pairs. If a set of propositional atoms is preferred to another, then it is preferred regardless of the required budget. In a preference that adheres to Def. 2b though, the alternatives with required budget 0 are strictly preferred to all the other alternatives with same propositional atoms.

We would like to emphasize that the basic preferences as we defined here are different than lexicographic ordering [28]. An agent's preference, as per Def. 2, satisfies, instead, the basic rationality requirements [38], as per Prop. 1.

Proposition 1. *A basic preference $Pref(a) = (A, \geq)$ for an agent $a \in Ag$ is*

- *transitive: $\forall x, y, z \in A$ if $x \geq y$ and $y \geq z$ then $x \geq z$; and*
- *complete: $\forall x, y \in A$ either $x \geq y$ or $y \geq x$ or $x \sim y$.*

Proof. See Appendix A. $\square$

3.3.2 Examples of Basic Preferences

Given $\mathcal{B} = \{0, 1\}$ and $BL = \mathcal{B}^2$, an example of basic preference defined according to Def. 2a is the following.

$$
\begin{aligned}
(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle) &\succeq (\langle sp_{100}, 0 \rangle, \langle dist_2, 0 \rangle) \succeq \\
(\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle) &\succeq (\langle sp_{50}, 0 \rangle, \langle dist_2, 0 \rangle) \succ \\
(\langle sp_{100}, 0 \rangle, \langle dist_1, 1 \rangle) &\succeq (\langle sp_{100}, 1 \rangle, \langle dist_1, 0 \rangle) \succeq \\
(\langle sp_{100}, 0 \rangle, \langle dist_2, 1 \rangle) &\succeq (\langle sp_{100}, 1 \rangle, \langle dist_2, 0 \rangle) \succeq \\
(\langle sp_{50}, 0 \rangle, \langle dist_1, 1 \rangle) &\succeq (\langle sp_{50}, 1 \rangle, \langle dist_1, 0 \rangle) \succeq \\
(\langle sp_{50}, 0 \rangle, \langle dist_2, 1 \rangle) &\succeq (\langle sp_{50}, 1 \rangle, \langle dist_2, 0 \rangle) \succ \\
(\langle sp_{100}, 1 \rangle, \langle dist_1, 1 \rangle) &\succeq (\langle sp_{100}, 1 \rangle, \langle dist_2, 1 \rangle) \succeq \\
(\langle sp_{50}, 1 \rangle, \langle dist_1, 1 \rangle) &\succeq (\langle sp_{50}, 1 \rangle, \langle dist_2, 1 \rangle)
\end{aligned}
\tag{1}
$$

Note that in preference (1), alternatives with lower required budget are preferred over alternatives with higher required budget and the agents' prefers sp_{100} over sp_{50} for every safety distance, and $dist_1$ over $dist_2$ for every speed.

Given $\mathcal{B} = \{0, 1\}$ and $BL = \mathcal{B}^2$, an example of basic preference defined according to Def. 2b is the following.

$$
\begin{aligned}
(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{100}, 0 \rangle, \langle dist_1, 1 \rangle) \succeq \\
(\langle sp_{100}, 1 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{100}, 1 \rangle, \langle dist_1, 1 \rangle) \succ \\
(\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{50}, 0 \rangle, \langle dist_1, 1 \rangle) \succeq \\
(\langle sp_{50}, 1 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{50}, 1 \rangle, \langle dist_1, 1 \rangle) \succ \\
(\langle sp_{100}, 0 \rangle, \langle dist_2, 0 \rangle) &> (\langle sp_{100}, 0 \rangle, \langle dist_2, 1 \rangle) \succeq \\
(\langle sp_{100}, 1 \rangle, \langle dist_2, 0 \rangle) &> (\langle sp_{100}, 1 \rangle, \langle dist_2, 1 \rangle) \succ \\
(\langle sp_{50}, 0 \rangle, \langle dist_2, 0 \rangle) &> (\langle sp_{50}, 0 \rangle, \langle dist_2, 1 \rangle) \succeq \\
(\langle sp_{50}, 1 \rangle, \langle dist_2, 0 \rangle) &> (\langle sp_{50}, 1 \rangle, \langle dist_2, 1 \rangle)
\end{aligned}
\tag{2}
$$

Notice that in preference (2) states of affairs where sp_{100} and $dist_1$ hold are preferred over states of affairs where sp_{50} and $dist_1$ hold, regardless of the budget. Analogously, regardless of the budget, states of affairs where sp_{50} and $dist_1$ hold are preferred over states of affair where sp_{100} and $dist_2$ hold, which, in turn, are preferred over states of affair where sp_{50} and $dist_2$ hold. Such preference describes an agent type that prefers to drive fast rather than slow and that prefers to have a short safety distance rather than high, for whom maximizing speed and minimizing safety distance have priority over minimizing the budget to be spent, and, finally, who gives more importance to having a short safety distance rather than driving fast.

Finally, an example of a preference that does not satisfy Def. 2 is $(\langle sp_{50}, 1 \rangle, \langle dist_1, 1 \rangle) > (\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle) > \ldots$. This is because the first two alternatives share the same propositional atoms but the alternative with higher required budget is preferred to the alternative with lower required budget.

3.3.3 Preferences

The basic preference as defined in Def. 2 may not be expressive enough to capture some realistic cases. In order to cover more cases and to make our approach applicable to model more realistic scenarios, we consider more complex types of agents' preferences that combines the two basic types of preferences (defined in Def. 2a and Def. 2b).

Intuitively, a rational agent may exhibit different preferences when the required budget increases. For example, consider a brave agent that prefers to drive fast and to keep a short safety distance rather than long, e.g., as per preference (2). Suppose, however, that such an agent is ready to pay only up to 1€ for driving fast and for keeping short safety distance. In such case, the agent would prefer to drive fast and to keep a short safety distance, compared to other alternatives (e.g., to drive slow and keep a long safety distance), if the required budget is lower than 1€. For example, in preference (2), ordered according to Def. 2b, we have $(\langle sp_{100}, 1 \rangle, \langle dist_1, 1 \rangle) > (\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle)$. If the required budget for either driving fast or keeping a short safety distance is higher than 1, however, the agent may instead give priority to spending the least possible. For example, $(\langle sp_{50}, 0 \rangle, \langle dist_1, 2 \rangle)$, not reported in preference (2), would be preferred to $(\langle sp_{100}, 1 \rangle, \langle dist_1, 2 \rangle)$, adhering to Def. 2a instead of Def. 2b. In other words, a rational agent may use different criteria to order the alternatives in a preference depending on the required budget.

We formalize this intuition by defining a type of preference $(A, \succeq)$ that is a sequence of k basic preferences, with $1 \leqslant k \leqslant |\mathcal{B}|$. We call such a complex preference simply *preference*. Each of the k basic preferences adhere to either Def. 2a or Def. 2b, and the alternatives in the different basic preferences have *increasing* budgets. In particular, the set of possible budget lists $BL_i \subseteq \mathcal{B}^n$ for an alternative in the i-th basic preference $(A_i, \succeq_i)$, for $i \leqslant k$, is determined as per Def. 3.

Definition 3 (Budget Lists of the i-th Basic Preference) Consider a set $\mathcal{B} \subset \mathbb{N}$, and k disjoint subsets of $\mathcal{B}$, i.e., $\mathcal{B}_1, \ldots, \mathcal{B}_k$, such that each element of $\mathcal{B}_i$ is bigger than each element of $\mathcal{B}_j$, for $j < i \leqslant k$. In a preference composed by k basic preferences, the set of possible budget lists for the alternatives in the i-th basic preference $(A_i, \succeq_i)$, for $i \leqslant k$, is $BL_i = (\bigcup_{j \leqslant i} \mathcal{B}_j)^n \setminus \bigcup_{j < i} BL_j$

For instance, given the set $\mathcal{B} = \{0, 1, 2\}$ and $k = 2$ two possible subsets of $\mathcal{B}$ as per Def. 3 are $\mathcal{B}_1 = \{0, 1\}$ and $\mathcal{B}_2 = \{2\}$. The possible budget lists for the alternatives of 2 basic preferences are therefore $BL_1 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ and $BL_2 = \{(0, 2), (1, 2), (2, 0), (2, 1), (2, 2)\}$. In other words, the budgets in the alternatives of the i-th basic preference are always lower or equal to $max(\mathcal{B}_i)$. This means that the required budget of every alternative in A_i is always *lower* or equal to $n \cdot max(\mathcal{B}_i)$, while the required budget of every alternative in A_{i+1} is always *higher* or equal to $n \cdot max(\mathcal{B}_i)$.

Definition 4 (Preference) Let $(A_1, \succeq_1), \ldots, (A_k, \succeq_k)$ be k basic preferences as per Def. 2, such that alternatives in A_i are defined with respect to a set of

budget lists BL_i as per Def. 3. An agent is said to have a *preference* $(A, \geq)$, iff $A = \bigcup_{i=1}^{k} A_i$ and $\geq = \bigcup_{i=1}^{k} \geq_i \cup \{ (x, y) \mid x \in A_j \;\&\; y \in A_i \;\&\; 1 \leq j < i \leq k \}$.

Note that a preference $(A, \geq)$ that is composed by only one basic preference $(A_1, \geq_1)$ so that $A = A_1$ for $BL_1 \subseteq \mathcal{B}^n$, and $\geq = \geq_1$, is a basic preference. If a preference is composed by more than one basic preference, every basic preference $(A_i, \geq_i)$ composing the preference adheres to either Def. 2a or Def. 2b, and for every pair of alternatives $x, y \in A$ such that $x \in A_i, y \in A_j$ and $i < j$, it holds that $req\_bud(x) \leq req\_bud(y)$. Furthermore, notice that the sets $A_1, \ldots, A_k$ of alternatives of the k basic preferences composing a preference $(A, \geq)$ are disjoint subsets of A, since the possible budget lists of the k basic preferences are disjoint subsets of $\mathcal{B}^n$.

Again, we note that a preference as per Def. 4 is transitive and complete.

Proposition 2. *A preference $Pref(a) = (A, \geq)$ for an agent $a \in Ag$ is*

- *transitive: $\forall x, y, z \in A$ if $x \geq y$ and $y \geq z$ then $x \geq z$; and*
- *complete: $\forall x, y \in A$ either $x \geq y$ or $y \geq x$ or $x \sim y$.*

Proof. See Appendix A. $\square$

3.3.4 Examples of Preferences

An example of a preference composed by two basic preferences $(A_1, \geq_1)$ and $(A_2, \geq_2)$ is given in Eq. 3, given $\mathcal{B} = \{0, 1, 2\}$.

$$
\begin{aligned}
&\#\text{from here ordered by Def. 2b} \\
(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{100}, 0 \rangle, \langle dist_1, 1 \rangle) \geq \\
(\langle sp_{100}, 1 \rangle, \langle dist_1, 0 \rangle) &> (\langle sp_{100}, 1 \rangle, \langle dist_1, 1 \rangle) > \\
(\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle) > \ldots &> (\langle sp_{50}, 1 \rangle, \langle dist_1, 1 \rangle) > \ldots > \\
&\quad (\langle sp_{50}, 1 \rangle, \langle dist_2, 1 \rangle) > \\
&\#\text{from here ordered by Def. 2a} \\
(\langle sp_{100}, 0 \rangle, \langle dist_1, 2 \rangle) &\geq (\langle sp_{100}, 2 \rangle, \langle dist_1, 0 \rangle) \geq \ldots > \\
(\langle sp_{100}, 2 \rangle, \langle dist_1, 2 \rangle) &> (\langle sp_{50}, 0 \rangle, \langle dist_1, 2 \rangle) \geq \ldots \geq \\
&\quad (\langle sp_{50}, 2 \rangle, \langle dist_2, 2 \rangle)
\end{aligned}
\tag{3}
$$

In such preference, the budget lists of the alternatives in A_1 are elements of $BL_1 = \{(0,0), (0,1), (1,0), (1,1)\}$ for $\mathcal{B}_1 = \{0, 1\}$, and the alternatives are ordered by Def. 2b. The budget lists of the alternatives in A_2, instead, are elements of $BL_2 = \{(0,2), (1,2), (2,0), (2,1), (2,2)\}$, for $\mathcal{B}_2 = \{2\}$, and they are ordered by Def. 2a. The required budget of every alternative in A_1 is lower or equal to 2, while the required budget of every alternative in A_2 is higher or equal to 2 and lower or equal to 4.

3.3.5 Consistent Preferences

The preferences above described allow to express a multitude of possible order-
ings between different states of affairs. In the following we define an additional
property that a preference can exhibit. We call such property *consistency* [30].

Intuitively a preference is consistent if when a state of affairs where a propo-
sitional atom p holds is preferred to a state of affair where q holds, then states
of affairs where p holds are preferred to states of affairs where q holds also when
a third atom r is considered. For instance, if $(\langle p, b_1 \rangle, \langle x, b_2 \rangle) \succeq (\langle q, b_1 \rangle, \langle x, b_2 \rangle)$,
then in a consistent preference this holds for every propositional atom x.

Notice that preferences as per Def. 4 are not necessarily consistent. An
example of a preference that is not consistent (i.e., does no exhibit the consis-
tency property) is the following:

$$(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle) > (\langle sp_{80}, 0 \rangle, \langle dist_1, 0 \rangle) > (\langle sp_{80}, 0 \rangle, \langle dist_2, 0 \rangle) >$$
$$(\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle) > (\langle sp_{50}, 0 \rangle, \langle dist_2, 0 \rangle) > (\langle sp_{100}, 0 \rangle, \langle dist_2, 0 \rangle) > ...$$

Notice that, given $dist_1$, sp_{100} is preferred to sp_{80}, but given $dist_2$, sp_{80} is
preferred to sp_{100}.

We define consistent preferences by means of an *enumeration condition*
over the propositional atoms of the alternatives. In particular, if two alterna-
tives x and y with same budget lists differ exactly for one propositional atom,
then if x is preferred to y, this has to hold also for all other pairs of alternatives
with same budget lists differing exactly for the same propositional atoms as x
and y. Intuitively the *enumeration condition* imposes an ordering on the alter-
natives that corresponds to an ordering that can be obtained by systematically
enumerating the possible combinations of propositional atoms. For instance,
if, given $dist_1$, the proposition sp_{100} from the set $\{sp_{100}, sp_{50}\}$ is enumerated
before proposition sp_{50} (i.e, $(\langle dist_1, b_1 \rangle, \langle sp_{100}, b_2 \rangle) > (\langle dist_1, b_1 \rangle, \langle sp_{50}, b_2 \rangle)$),
then in a consistent preference sp_{100} is enumerated before sp_{50} also given $dist_2$
(i.e, $(\langle dist_2, b_1 \rangle, \langle sp_{100}, b_2 \rangle) > (\langle dist_2, b_1 \rangle, \langle sp_{50}, b_2 \rangle)$).

Definition 5 A preference $Pref(a) = (A, \succeq)$ is *consistent* if and only if for all
alternatives x, y in A s.t. their lists of propositional atoms differ exactly for
one element, the following *enumeration condition* holds.
Let $\diamond \in \{>, \sim\}$.
$x \diamond y \Rightarrow$
$\qquad \forall v, w \in A \mid prop(v) = (p_1, .., p_n) \ \& \ prop(w) = (p'_1, .., p'_n)$
$\qquad\qquad$ if $p_i \neq p'_i \ \& \ \forall_{k \in \{1, ..., n\} | k \neq i} p_k = p'_k \ \& \ bud(v) = bud(w)$
$\qquad\qquad$ then $v \diamond w$

3.4 Norms and Agents' Preferences

As above mentioned, in this paper we assume that norms and agents' pref-
erences are comparable. Consider $AL = (L_1, \ldots, L_n)$ and a norm set $\mathcal{N} =
\langle N_1, \ldots, N_n \rangle$ as per Sec. 3.2. Given an alternative $(\langle p_1, b_1 \rangle, ..., \langle p_n, b_n \rangle)$ in an
agent's preference, we have that both the proposition p_i of i-th pair $\langle p_i, b_i \rangle$

and the proposition p of the i-th norm $N_i = (p, s)$ in $\mathcal{N}$ belong to L_i. Furthermore, since both the sanctions of the norms and the agents' budgets of agent's preferences are natural numbers, they also are commensurable. This makes it possible to analyze an agent's preference in the context of a norm to determine whether the preference motivates an agent to comply with a norm or to violate it.

Intuitively, in the context of a set of enforced norms, an agent that follows its preference aims at realizing a state of affairs that can be compliant with some of the enforced norms and violating other norms for which he is willing to pay the corresponding sanctions.

Given a set $\mathcal{N}$ of n norms and a preference $(A, \succeq)$, we say that an alternative $x \in A$ such that $x = (\langle p_1, b_1 \rangle, \ldots, \langle p_i, b_i \rangle, \ldots, \langle p_n, b_n \rangle)$ is a *violating alternative* w.r.t. the i-th norm $N_i = (p, s)$ in $\mathcal{N}$, and we write $viol(x, N_i)$, if and only if p_i (e.g., sp_{100}) excludes[3] p (e.g., sp_{50}); otherwise x is said to be a *complying alternative* w.r.t. norm N_i. An alternative that is compliant w.r.t. all norms in $\mathcal{N}$ is said *fully compliant*. Note that any rational preference, due to its completeness property as per Prop. 2, always contains at least one fully compliant alternative. This means that agents always have a choice to aim at a state of affairs that does not violate any norm.

Definition 6 (Most Preferred Alternatives to Act Upon) Given a preference $(A, \succeq)$ and a set $\mathcal{N}$ of n norms, a subset $A' \subseteq A$ of alternatives is called the *set of most preferred alternatives to act upon* in the context of $\mathcal{N}$ if and only if for all alternatives $x \in A \backslash A'$ it holds that for all alternatives $y \in A'$ either $y > x$ or $x \succeq y$ and there exists a norm $N_j = (p, s)$ in $\mathcal{N}$ s.t. $viol(x, N_j)$ & $b_j < s$ (with b_j budget of the j-th pair in x).

The set of most preferred alternatives to act upon in the context of $\mathcal{N}$ is the set of alternatives $A' \subseteq A$ such that every other alternative $x \in A \backslash A'$ is either strictly less preferred (i.e., $y > x \ \forall y \in A'$), or is an alternative that violates at least a norm N_j but the budget is not enough to pay the sanction (i.e., $viol(x, N_j)$ & $b_j < s$). This means that the alternatives in A' are either fully compliant or they violate some norms and the budget is enough to pay the sanction, and there is no other alternative that satisfies such conditions that is strictly preferred to them.

A rational agent always acts upon one of its most preferred alternatives. We say that an agent a has *a reason to violate* a norm N whenever the agent's preference $Pref(a)$ is so that, among the set of most preferred alternatives, there is at least one alternative x such that $viol(x, N)$. When different alternatives are equally preferred by an agent, the agent can freely choose to aim at any of them. This means that an agent that has a reason to violate a norm will not necessary aim to violate it: if another alternative is equally preferred

---

[3] In this paper we assume that information about exclusion between propositional atoms (e.g., in the sense of material implication) is given as background knowledge. A formal definition of violation of a norm depends on the specific language used to specify the norms and is out of the scope of the paper.

to the violating state of affairs, the agent may decide to aim to the obeying state of affairs, despite it has a reason to violate the norm. Consider for example an agent type characterized by the preference in Eq. 1 and a norm $N = (sp_{50}, 0)$ that prohibits agents to drive faster than $50 km/h$. Given N, the agents' most preferred alternatives to act upon are $(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle)$, $(\langle sp_{100}, 0 \rangle, \langle dist_2, 0 \rangle)$, $(\langle sp_{50}, 0 \rangle, \langle dist_1, 0 \rangle)$ and $(\langle sp_{50}, 0 \rangle, \langle dist_2, 0 \rangle)$. Some of these alternatives violate the norm N (e.g., $(\langle sp_{100}, 0 \rangle, \langle dist_1, 0 \rangle)$), therefore the agent has a reason to violate N. However, some of the other most preferred alternatives are compliant with the norm (e.g., $(\langle sp_{50}, 0 \rangle, \langle dist_2, 0 \rangle)$). Since all most preferred alternatives are equally preferred, the agent may rationally decide to aim at any of them.

We introduce the notion of *maximum budget for norm violation* as the maximal payment that an agent is willing to pay for violating a given norm according to its preference. Let $N_i = (p, s)$ be the i-th norm in $\mathcal{N}$, and let $Pref(a) = (A, \succeq)$ be the preference of agent a. Let $x \in A$ be the agent's most preferred *fully compliant* alternative, and $A' = \{y \in A \mid y \succeq x\}$ be the set of alternatives in A that are (equally) preferred to x. The maximum budget that a is willing to pay for the violation of N_i, denoted as $maxB(a, N_i)$, is the highest budget b that occurs in the i-th pair of the alternatives in A'. Note that if the maximum budget for violating a norm is lower than the sanction of norm N_i, then the most preferred alternatives to act upon are necessarily alternatives compliant w.r.t. N_i. For instance if $N = (sp_{50}, 3)$ and an agent a has $maxB(a, N) = 2$, then all alternatives x in the set of most preferred alternatives are compliant to N, i.e., $viol(x, N)$ does not hold, and it does not exists a pair $\langle p, b \rangle \in x$ with $b \geqslant 3$, since $b \leqslant maxB(a, N) < 3$.

Finally, it is worth noting that in case of preference composed by more than one basic preference as per Def. 4, it is always the case that if the first basic preference is strictly preferred to the remaining ones then the set of most preferred alternatives to act upon in the context of $\mathcal{N}$ never contains any alternatives from any basic preference apart from the first one. This is because the first basic preference necessarily contains an alternative that is fully compliant (due to completeness of every basic preference $(A_i, \succeq_i)$ w.r.t. AL and BL_i for $1 \leqslant i \leqslant k$ and k number of basic preferences composing the preference), and such alternative is strictly preferred to any other alternative that belongs to the remaining basic preferences.

4 Norm-based Supervision

In this section we present the key concepts of a norm-based supervision of a multi-agent system. We build on the runtime norm-based supervision mechanism for multiagent systems as proposed in [24] and sketched in Fig. 2. Such mechanism corresponds to a control loop that continuously monitors the behavior of a multiagent system, evaluates the enforcement of the norms w.r.t. the system-level objectives, and, when needed, intervenes by revising the norms.
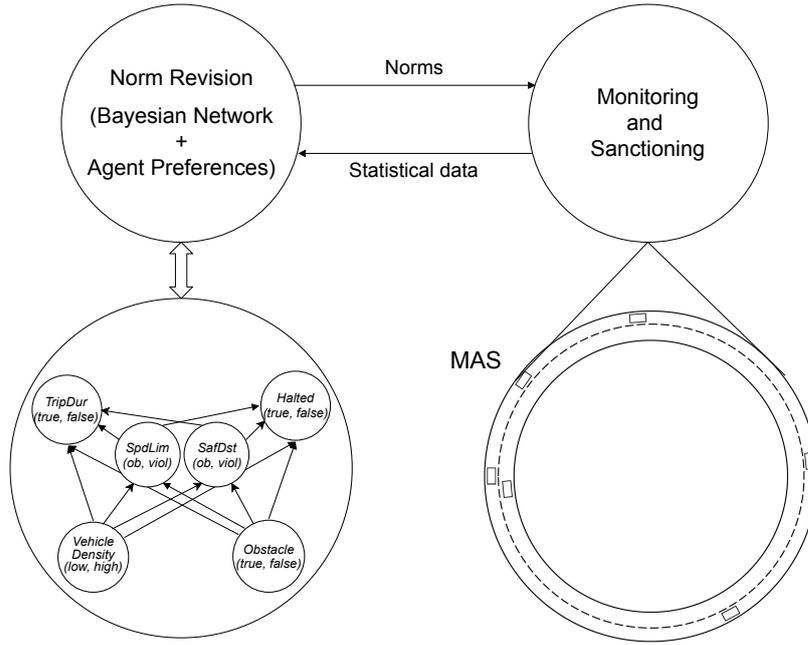
Fig. 2: Illustration of the MAS supervision mechanism.

Consider an ordered set $\mathcal{N} = \langle N_1, ..., N_n \rangle$ of norms and a set $\mathcal{C}$ of all possible operating contexts of the multiagent system (e.g, a context $c \in \mathcal{C}$ in the ring road scenario could be "low vehicle density and no obstacle"). We call *system configuration* an assignment of a sanction $s \in \mathbb{N}$ to each norms in $\mathcal{N}$ in each of the MAS operating contexts.

For example, given two possible operating contexts c_1 and c_2, and given a norm set $\mathcal{N} = \{N_1, N_2\}$, a possible system configuration is $\{(c_1, (N_1, 1), (N_2, 0)),$ $(c_2, (N_1, 0), (N_2, 1))\}$, meaning that in context c_1 norms N_1 and N_2 are enforced respectively with sanctions 1 and 0, while in context c_2 they are enforced respectively with sanctions 0 and 1.

The control loop of the supervision mechanism sketched in Fig. 2 starts with an initial system configuration. A *Monitoring and Sanctioning* component collects, at runtime, perfect information about the obedience or violation of the norms in the contexts in which they are evaluated and sanctions agents that violate the norms. Such component also provides a Boolean evaluation of the system-level objectives (e.g., whether the number of halted cars is below a certain threshold or not, in the ring road scenario).

The collected information is used to automatically train a Bayesian Network called *Norm Bayesian Network* (described in Sec. 4.1) that is used to learn and reason at runtime about the correlation between norm obedience or violation and the achievement of the system-level objectives. For example, the *Norm Bayesian Network* helps answering questions like how well, and in

which contexts, does the norm $(sp_{50}, 100)$ help achieve the objective of avoiding halted cars?

A *Norm Revision* component makes use of the learned knowledge, encoded in the Bayesian Network, to determine whether some norms should be revised and how. Revising a norm $N = (p, s)$ means modifying either the proposition p or the sanction s, or both. In this paper we focus on the revision of the sanctions of the norms. The norm revision process generates as output a (possibly) new system configuration, replacing the current one.

In previous work [24], we proposed an implementation of the control loop above described as a variation of the hill climbing optimization technique. In this paper we follow the same approach. We consider the system configurations as possible solutions to explore in order to find an optimal one. The quality of a solution is determined, by means of the observed data from MAS execution, as the probability of achieving the system-level objectives. Instead of terminating the exploration of the space when a local optimum is found, as in traditional hill climbing, we use as stopping criterion a constraint defined by the system designer that determines whether or not the current solution is acceptable. In particular we use, as stopping criterion, a minimum desired value of the probability of achieving the system-level objectives. We call such value t_{oa}. We use the *Norm Revision* component to determine the next solution to try, when the current one is not acceptable.

In [24] we proposed heuristic algorithms for suggesting norm revisions that alter the *regimented norms*. In this paper, differently from the earlier work, we make use of some additional information concerning the preferences of the agents in order to determine how to revise the norms, and we focus on the revision of sanctions. In [25], we used the same framework of [24] to revise the way one norm is *enforced* by modifying its sanction. In this paper, we significantly extend our previous work by devising several new strategies for the revision of the sanctions of *multiple norms* enforced at the same time.

In the rest of the section we first provide some background concerning the *Norm Bayesian Network*, then we analyze some properties of the relationships between norms, agents' preferences and system-level objectives.

4.1 Norm Bayesian Network

Consider some monitorable environmental properties such as the density of vehicles or the presence of an obstacle on the ring road. Each of these properties is called *contextual variable*, and is associated to a domain of values. For example, *Vehicles density* can be either *low* or *high*, while *Obstacle* can be *true* or *false*. Given a set of contextual variables, a *context* assigns a value to each contextual variable. For instance, given *Vehicles density* and *Obstacle*, four possible contexts exist: *high-true, high-false, low-true, low-false.*

A *Norm Bayesian Network* $\mathcal{NBN} = (\mathcal{X}, \mathcal{A}, \mathcal{P})$ [24] is a Bayesian Network where:

- $\mathcal{X} = \mathbf{N} \cup \mathbf{O} \cup \mathbf{C}$ are nodes that represent random variables in probability theory. $\mathbf{N}$, $\mathbf{O}$ and $\mathbf{C}$ are disjoint sets. $\mathbf{N}$ consists of *norm nodes*; each node $N \in \mathbf{N}$ corrensponds to a norm and has a discrete domain of 3 possible values: *obeyed*, *violated* and *disabled*. $\mathbf{O}$ consists of *objective nodes*; each node $O \in \mathbf{O}$ corresponds to a Boolean objective and has a discrete domain of 2 values: *true* and *false*. Finally, $\mathbf{C}$ consists of *context nodes*; each node $C \in \mathbf{C}$ corresponds to a contextual variable and can have a discrete or continuous domain of values.
- $\mathcal{A} \subseteq (\mathbf{C} \times \mathbf{N}) \cup (\mathbf{C} \times \mathbf{O}) \cup (\mathbf{N} \times \mathbf{O})$ is the set of arrows that connect pairs of nodes. If there is an arrow from node X to node Y, X is called parent of Y.
- $\mathcal{P}$ is a set of conditional probability distributions. These are encoded into conditional probability tables (CPTs), each one associated with a node in $\mathcal{X}$ and quantifying the effect of the parents on the node. The conditional probability values in the CPT of a node are the *parameters* of the network. These parameters are automatically learned from observed data from MAS execution through classic Bayesian learning.

Notation. In the rest of the paper, we use the following notation for Bayesian Networks. Italic uppercase $(X, Y, \ldots)$ for random variables; bold uppercase $(\mathbf{X}, \mathbf{Y}, \ldots)$ for sets of random variables; italic lowercase $(v_1, v_2, \ldots)$ for values in the domain of a random variable; N_v abbreviates $(N = v)$, i.e., an assignment of value v to a norm variable N; $\mathbf{O}_v$ denotes an assignment of value v to all nodes in $\mathbf{O}$; P denotes a single probability. An evidence $\mathbf{e}$ is an observed assignment of values for some or all of the random variables in the network. An evidence $\mathbf{c}$ for all the context nodes $\mathbf{C}$ is an observation for a certain context; for example, *Vehicles density* has value *low* and *Obstacle* has value *false*. For simplicity, we use the term *context* also to refer to the associated evidence in the Bayesian Network.

Fig. 3 reports an example of a *Norm Bayesian Network* for the running example of the ring road.

Since we focus on revising the sanctions that enforce norms, norms are never disabled, therefore in the following we ignore the *disabled* value of the nodes in the Bayesian Network. Despite we do not explicitly disable a norm, we consider enforcing a norm with a sanction of 0 as equivalent to disabling the norm, assuming that an agent that violates a norm with sanction of 0 does not incur in any other kind of sanctions (e.g., consequences in the relation between the individual and the other agents due to shared (moral) values [7]).

Finally, the construction and training of the *Norm Bayesian Network* is a fully automated process. In particular, the structure of the network can be trivially obtained from the the definition of $\mathcal{X}$ and $\mathcal{A}$. The conditional probability distributions $\mathcal{P}$ (i.e., the parameters of the network), instead, are automatically learned through classical Bayesian learning using data collected from MAS execution. Without going into the details of the *Monitoring and Sanctioning* component, which are out of the scope of this paper, Table 1 reports a sample dataset that can be obtained from monitoring norms and
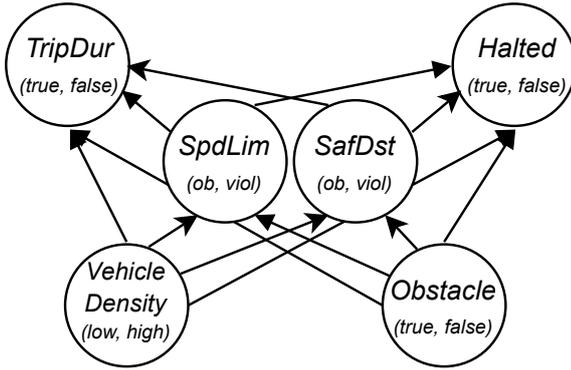
Fig. 3: A *Norm Bayesian Network* for the ring road.

objectives for the running example of the ring road. The values that each of the variables assumes belongs to its domain as above specified (e.g., *obeyed*, *violated*, for norm nodes, *true* or *false* for objective nodes). Such dataset can be used to automatically train the *Norm Bayesian Network* of Fig. 3 and learn the set of conditional probability distributions $\mathcal{P}$. As in this work we assume that the population of agents do not change over time and that the behavior of agents is consistent over time, the CPTs of the *Norm Bayesian Network* stabilize after receiving a sufficient number of evidences.

Table 1: Example of part of a dataset used to train the *Norm Bayesian Network* of Fig. 3 and obtained from monitoring the execution of the MAS.

VehicleDensity	Obstacle	SpdLim	SafDst	TripDur	Halted
low	true	viol	ob	true	false
low	false	ob	viol	false	false
high	true	viol	ob	true	false
high	false	ob	ob	true	true
...					

4.2 Norms, Agents' Preferences and System-level Objectives in MAS

Consider a set of agent types $\mathcal{T} = \{t_1, \ldots, t_k\}$, each type corresponding to a preference as per Sec. 3. In order to focus on the revision of the norms' sanctions, we assume that we possess a correct estimation of the preferences of agents concerning the aspects of the system we aim to regulate. Additionally, we assume that the agents' preferences do not change in different contexts. As we will see in the following, an accurate estimation of agents' preferences is helpful for improving the effectiveness of our heuristics. Our technique, however can be extended to support partial or inaccurate estimations of the agents'

preferences. In Section 7.1, we sketch some directions for future work to support these aspects.

Take a set of agents $Ag = \{a_1, \ldots, a_n\}$, each with a specific type from $\mathcal{T}$. We use $Pref(a) \in \mathcal{T}$ to indicate that agent $a \in Ag$ behaves according to a type from $\mathcal{T}$. For simplicity we assume that the behaviors exhibited in the multiagent system are uniformly distributed over all the agents: at every time instant every agent either violates or obeys each of the enforced norms.

Given these assumptions and a set of norms $\mathcal{N}$, we say that a norm N in $\mathcal{N}$ is *well defined* in the context of $\mathcal{N}$ (simply *well defined*, for brevity) if the probability that N is violated, denoted as $P(N_{viol})$, is never higher than the percentage of agents in the MAS with a reason to violate N in the context of $\mathcal{N}$.[4] In other words, the *upper bound* of the probability $P(N_{viol})$ in the context of $\mathcal{N}$ (denoted as $UB(N_{viol}, \mathcal{N})$) is the percentage of the agents with a reason to violate N in the context of $\mathcal{N}$.

Let N be a norm in $\mathcal{N}$, and let $\delta = (d_1, ..., d_k)$ be a distribution over the agent types $\mathcal{T} = \{t_1, \ldots, t_k\}$, where $d_i \in [0, 1]$ is the percentage of population of agents of type t_i, with $\sum_{i=1}^{k} d_i = 1$. The percentage of agents with a reason to violate N (as per Sec. 3.4) in the context of $\mathcal{N}$ is $\sum_{i=1}^{k} (d_i \cdot hasReason(i, N, \mathcal{N}))$, with $hasReason(i, N, \mathcal{N}) = 1$ if agent type t_i has a reason to violate N in the context of $\mathcal{N}$, *0* otherwise.

Consider, as an example, a norm set $\mathcal{N} = \langle N_1, N_2 \rangle$, with $N_1 = (sp_{50}, s_1)$ and $N_2 = (dist_2, s_2)$ and $\mathcal{B} = \{0, 1\}$. Consider the two types of agents t_1 and t_2 as per Eq. 1 and Eq. 2, respectively. Assuming a uniform distribution of agents between the two types, Fig. 4 reports the upper bound of the probability of violating N_1 and N_2 for this example with different sanctions (i.e., different values of s_1 and s_2).

The upper bound of $P(N_{viol})$ describes a *worst-case* hypothetical situation where all agents behave according to their preferences, and if they have reason to violate a norm they are assumed to violate it, no contextual factor influences agent behavior, and interactions among agents do not prevent them to act according to their preferences. This would happen, for example, when a single car drives on an empty highway with perfect road and car conditions. Note, however, that the actual probability to violate a norm is affected by the agents' decisions, their interactions and by the MAS environment, and it is assumed to be unknown *a priori*. Even if all agents have a reason to violate a norm, due to their interaction or to environmental circumstances (e.g., large number of cars on the ring-road), none of them may end up violating it. Furthermore, as explained in Sec. 3.4, if an agent equally prefers two states of affairs, one

---

[4] Consider a norm N = *every vehicle on the ring road shall always exceed 70km/h*, a type of norm employed in our society, for instance, to prevent vehicles to have negative impact on road throughput and safety. Our framework supports such type of norm if it is well-defined. Suppose that in our running example no agent has reason to violate N. If N was well-defined we would expect $P(N_{viol}) = 0$. However, *in our running example*, such norm is not well-defined, for in case of high density, for example, the agents may be forced to slow down below the minimum speed, therefore violating the norm and exhibiting $P(N_{viol}) > 0$. A well-defined norm guarantees agents that have no reason to violate the norm (i.e., their preferred alternatives are compliant with the norm) to be able to obey such norm.
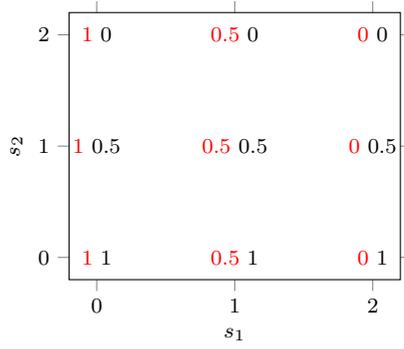
Fig. 4: Upper bound of the probability of violating norms $N_1 = (sp_{50}, s_1)$ (in red) and $N_2 = (dist_2, s_2)$ (in black) with the two types of agents t_1 and t_2 as per Eq. 1 and Eq. 2, respectively, uniformly distributed.

violating a norm, and another obeying the norm, the agent, since autonomous, may decide to obey the norm even if it has a reason to violate it. We call, therefore, the monitored probability of violating (obeying) a norm *exhibited norm violation* (*obedience*). We do not assume any prior knowledge about such probability.

Note that, since we consider agent types with rational preferences as per Sec. 3.3, increasing the sanction s of a norm $N = (p, s)$, without changing the sanctions of other norms, does not increase the percentage of agents with a reason to violate N. Therefore, given k agent types and $maxB(\mathcal{T}, N)$ as the maximum budget among all agent types to violate a well-defined norm $N = (p, s)$, the percentage of agents with a reason to violate a well-defined norm $N' = (p, maxB(\mathcal{T}, N) + 1)$ in the context of $\mathcal{N}$ is 0. This is to say that increasing the sanction of a norm above the maximum budget that any agent is willing to pay causes all agents to comply with the norm. Consequently, given two well-defined norms $N = (p, s_1)$ and $N' = (p, s_2)$ such that $s_2 > s_1$, and assuming no change in other norms of $\mathcal{N}$, the upper bound of the probability $P(N'_{viol})$ is never bigger than the upper bound of the probability $P(N_{viol})$.

Furthermore, it is possible to prove that, if all agents in the MAS have a consistent preference (as per Def. 5), then given a set of norms $\mathcal{N} = \langle N_1, ..., N_n \rangle$, increasing the sanction of a norm N_j in $\mathcal{N}$ without changing the sanctions of other norms, does not increase the upper bound of the probability $P(N_{viol})$ for every N in $\mathcal{N}$.

Proposition 3. *Given an ordered set of norms $\mathcal{N} = \langle N_1, ..., N_n \rangle$, and a set of t agent types $\mathcal{T}$, each type corresponding to a consistent preference (as per Def. 5), increasing the sanction of a norm N_j in $\mathcal{N}$ without changing the sanctions of other norms, does not increase the upper bound of the probability $P(N_{viol})$, i.e., $UB(N_{viol}, \mathcal{N})$, for all N in $\mathcal{N}$.*

Proof. See Appendix A.   □

The concept of well-defined norm as described above, concerns the relationship between a norm and the preferences of the agents. In a multiagent system, norms are enforced in order to achieve some system-level objectives. Although setting the sanction of all norms in $\mathcal{N}$ above $max(\mathcal{B})$ makes all the agents fully compliant (i.e., $P(N_{viol}) = 0$ and $P(N_{ob}) = 1$ for all $N \in \mathbf{N}$), this does not necessarily guarantee the achievement of the system-level objectives, as norms can be ineffective, or even harmful, when obeyed by all agents [24]. Having an estimation of the agents' preferences on its own is therefore not sufficient for an effective supervision of a MAS.

We describe here two properties that, instead, relate a norm with the system-level objectives: the concept of *synergy* between a norm and the system-level objectives, and the concept of *effectiveness* of a norm set.

We say that there is a *positive synergy* between a norm and the system-level objectives if it is more likely to achieve the system-level objectives when the norm is obeyed than when it is violated. A *positive synergy* between a norm N and a set of Boolean objectives $\mathbf{O}$ exists if $P(\mathbf{O}_{true}|N_{ob}) > P(\mathbf{O}_{true}|N_{viol})$. We say that there is a *negative synergy* between N and $\mathbf{O}$ if $P(\mathbf{O}_{true}|N_{ob}) < P(\mathbf{O}_{true}|N_{viol})$. Finally, we say that there is *no synergy* between N and $\mathbf{O}$ if $P(\mathbf{O}_{true}|N_{ob}) = P(\mathbf{O}_{true}|N_{viol})$.

We say, instead, that a norm set $\mathcal{N}$ is *effective* if, when norms in $\mathcal{N}$ are enforced, $\mathcal{N}$ guarantees the desired achievement level t_{oa} of the system-level objectives, i.e., when $P(\mathbf{O}_{true}) \geqslant t_{oa}$. Conversely, if, when enforcing a norm set $\mathcal{N}$, we have that $P(\mathbf{O}_{true}) < t_{oa}$, we say that $\mathcal{N}$ is *ineffective.*

Information such as the exhibited norm obedience, the synergy and the effectiveness described above, are hard to determine while designing a MAS. This is due to several factors, including the complexity of the system, the interaction between autonomous agents, the lack of complete knowledge of the agents' internals, and the uncertainty of the environment. However, they can be learned at runtime by monitoring the MAS execution. In this paper, we learn such properties by means of the *Norm Bayesian Network* and, in Sec. 5, we propose different strategies to combine these properties with the agents' preferences, in order to revise the sanctions of an ineffective norm set $\mathcal{N}$.

5 Norm Revision

In this section we propose different heuristic strategies for the revision of the sanctions of a set of norms whose enforcement is currently ineffective (as per Sec. 4.2). Opportune sanctioning of agents is a well-known mechanism to achieve the system-level objectives in MASs [12,13]. Our strategies leverage the knowledge learned at runtime about norm effectiveness and an estimation of the preferences of the agents in the system, and determine a new set of sanctions to use to enforce the norms.

Take the *Norm Bayesian Network* in Fig. 3. By analyzing the CPTs of the objectives nodes $\mathbf{O} = \{TripDur, Halted\}$, we can determine whether a norm set

$\mathcal{N}$ is effective or not in a context $\mathbf{c}$. If $\mathcal{N}$ is not effective (i.e., $P(\mathbf{O}_{true}|\mathbf{c}) < t_{oa}$), a norm revision process is triggered. In such case, in this paper we aim to revise the sanctions of the norms in $\mathcal{N}$. For example, if the two norms $(sp_{50}, 1)$ and $(dist_1, 1)$ are ineffective when on the ring road there is an obstacle and high vehicle density, we aim to identify another set of values for their sanctions. Given a norm set $\mathcal{N}$ consisting of n norms, a set of agent types $\mathcal{T}$ and the maximum possible budget $max(\mathcal{B})$ among all agent types in $\mathcal{T}$, the possible sets of sanctions that can be used to enforce norms in $\mathcal{N}$ is $\mathcal{S} = \times_{i=1}^{n}\{s \in \mathbb{N} \mid s \leqslant max(\mathcal{B}) + 1\}$. When a norm is enforced with a sanction 0, agent's decisions are not affected by the norm, since every agent can always afford to violate (if preferred) a norm with sanction 0. When a norm is enforced with a sanction $max(\mathcal{B}) + 1$, instead, no agent can violate such norm, since no agent can afford to pay such sanction, for the maximum possible budget among all agent types is $max(\mathcal{B})$.

The set $\mathcal{S}$ is the search space within which our heuristic strategies for norm revision search for new sanctions.

In Sec. 5.1, we describe six strategies for the suggestion of a revision of the sanctions of a norm set. Such strategies extend and adapt heuristics presented in previous work [24, 25] by supporting the revision of sanctions of *multiple norms*. Each strategy suggests how the behavior of agents w.r.t. the aspects of the system regulated by norms should change in order to improve the probability of achieving the system-level objectives. For example, given two norms, one strategy could suggest to reduce the violations of one norm and to increase the violations of the second norm. Based on the upper bound of the violation of norms obtained from agents' preferences (Sec. 4.2), we provide then in Sec. 5.2 an algorithm to explore the search space $\mathcal{S}$ in order to identify a new set of sanctions that satisfies (as much as possible) the suggestions provided by the revision strategies.

It is worth noting that we do not claim that modifying sanctions is always enough in order to achieve the system's objectives. As shown in previous work [24], sometimes the enforced norms (and not their enforcement) need to be revised. In this paper, however, we focus on mechanisms for the revision of the sanctions associated to the norms (i.e., the way norms are enforced). The combination of the mechanisms proposed here with the revision of the content of the norms is left for future work.

5.1 Norm Revision Strategies

We propose six strategies for the suggestion of norm revisions. Each strategy determines a list of n suggestions (one per each norm in $\mathcal{N}$). We present three types of strategies: *synergy-based strategies*, *sensitivity-based strategies*, and *category-based strategies*.

Each strategy is applied to a context $\mathbf{mpc}$ that, in our framework, corresponds to the most problematic context in which the objectives are not achieved. In particular, $\mathbf{mpc} = argmax_{\mathbf{c} \in all(\mathbf{c})} P(\mathbf{O}_{false} \mid \mathbf{c})$, where $all(\mathbf{c})$ is

the set of all possible contexts (assignments of a value to each of the context nodes in $\mathcal{NBN}$). For simplicity, in the rest of the section, we call such context simply $\mathbf{c}$.

5.1.1 Synergy-based strategies

Synergy-based strategies are based on the concept of norm-objectives synergy described in Sec. 4.2. The idea is that, if there is a positive synergy between a norm N and the objectives $\mathbf{O}$ in $\mathbf{c}$, the objectives $\mathbf{O}$ are more likely to be achieved when N is obeyed. In this case, by reducing the violations of N, we expect to increase $P(\mathbf{O}_{true}|\mathbf{c})$. If there is a negative synergy between N and $\mathbf{O}$ in $\mathbf{c}$, instead, we expect that increasing the violations of N, and $P(\mathbf{O}_{true}|\mathbf{c})$ would increase. We present two strategies of this type (*Naive synergy* and *Combined synergy*), which differ in the way they determine the synergy between norms and objectives.

Naive synergy. Consider, *for each norm $N \in \mathbf{N}$*, its synergy with the objectives $\mathbf{O}$:

$$argmax_{v \in \{ob, viol\}} P(\mathbf{O}_{true}|N_v \wedge \mathbf{c}) \tag{4}$$

For instance, for a norm node *SpdLim* in the Bayesian Network of Fig. 3, where $\mathbf{O} = \{TripDur, Halted\}$, we have that

$$P(\mathbf{O}_{true} \mid N_v \wedge \mathbf{c}) = P(TripDur_{true}, Halted_{true} \mid$$
$$SpdLim_v \wedge \mathbf{c} \wedge SafDst_{ob}) \cdot P(SafDst_{ob} \mid \mathbf{c}) + P(TripDur_{true}, Halted_{true} \mid$$
$$SpdLim_v \wedge \mathbf{c} \wedge SafDst_{viol}) \cdot P(SafDst_{viol} \mid \mathbf{c})$$

To determine the *argmax* of Eq. 4 means therefore to determine if $SpdLim_{ob}$ is better than $SpdLim_{viol}$ for the achievement of the objectives *TripDur* and *Halted*.

Naive synergy calculates such *argmax* for each norm node and suggests to decrease violations of norms such that $v = ob$ in Eq. 4, and to increase violations of norms where $v = viol$ in Eq. 4. For instance, given $\mathcal{N} = \langle N1, N2 \rangle$, if $v = ob$ for $N1$ and $v = viol$ for $N2$, then *naive synergy* suggests to decrease violations of norm $N1$ and to increase violations of norm $N2$.

Combined synergy. Determine which combination of values *obeyed* and *violated* for each norm is the best for the achievement of the objectives $\mathbf{O}$.
Let $\mathbf{ov}$ be the set of all possible assignments of values in the set $\{ob, viol\}$ to all norm nodes in $\mathbf{N}$ (e.g., given $\mathbf{N} = \{N1, N2\}$, then $\mathbf{ov} = \{\{N1_{ob}, N2_{ob}\}, \{N1_{ob}, N2_{viol}\}, \{N1_{viol}, N2_{ob}\}, \{N1_{viol}, N2_{viol}\}\}$). Determine:

$$\mathbf{n}_d = argmax_{\mathbf{n} \in \mathbf{ov}} P(\mathbf{O}_{true}|\mathbf{n} \wedge \mathbf{c}) \tag{5}$$

This strategy suggests to decrease violations of norms with value ob in $\mathbf{n}_d$ and to increase violations of norms with value $viol$ in $\mathbf{n}_d$. For instance if $\mathbf{n}_d = \{N1_{ob}, N2_{viol}\}$, then *combined synergy* suggests to increase violations

of norm $N1$ and to decrease violations of norm $N2$.

It is worth noting that *Combined synergy* purely determines the best combination of values for the norms, according to the observed data from MAS execution, without considering the prior probability of observing those values (in practice, *Combined synergy* only compares, one by one, the rows of the CPT of the objective nodes). *Naive synergy*, instead, when comparing different combinations of values for the norms, takes also into account the probability to observe those values (*Naive synergy* compares sums of different rows of the CPT of the objectives nodes, multiplied by the prior probability of observing the corresponding values for the norm nodes). Adopting the *Naive synergy* strategy may have the advantage of providing more precise suggestion w.r.t. the data acquired so far during the system execution. Considering only the CPT of the objective nodes, as per *Combined synergy*, may help instead determining the actual best combination of values of obedience of the norms for the system-level objectives, without being biased by the current probabilities of violating the norms, which will be modified after the sanctions revision.

5.1.2 Sensitivity-based strategies

Sensitivity-based strategies are based on the sensitivity analysis technique from probabilistic reasoning [15]. Such strategies do not only determine the direction of the revision—i.e., increasing or decreasing the probability of violating a norm, as in the case of synergy-based strategies—, but also estimate the required change in such probability in order to make the entire norm set effective in context $\mathbf{c}$. In particular, given a norm node N, the probability $P(N_{viol}|\mathbf{c})$ is a parameter $\theta_{N_{viol|\mathbf{c}}}$ of the *Norm Bayesian Network*. Sensitivity-based strategies try to identify possible changes to the parameter $\theta_{N_{viol|\mathbf{c}}}$ that can ensure the satisfaction of the constraint $P(\mathbf{O}_{true}|\mathbf{c}) \geqslant t_{oa}$. We call *required revision strength* (*RRS*) for a norm set $\mathcal{N} = \langle N1,...Nn \rangle$, the set of desired changes $\{\Delta\theta_{N1_{viol|\mathbf{c}}},...,\Delta\theta_{Nn_{viol|\mathbf{c}}}\}$ in the parameters $\theta_{N_{viol|\mathbf{c}}}$ of each N in $\mathcal{N}$ that ensure the satisfaction of the constraint $P(\mathbf{O}_{true}|\mathbf{c}) \geqslant t_{oa}$. We present two strategies of this type (*Naive sensitivity analysis* and *n-CPT sensitivity analysis*), which differ in the way they determine such set of desired changes for each norm in $\mathcal{N}$.

Naive sensitivity analysis. Determine, *for each norm N*, the required revision strength (*RRS*) $\Delta\theta_{N_{viol}}$ by solving equation 6.

$$P(\mathbf{O}_{true}|\mathbf{c}) + \frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N_{viol|\mathbf{c}}}} \cdot \Delta\theta_{N_{viol|\mathbf{c}}} \geqslant t_{oa} \tag{6}$$

Consider the topology of a *Norm Bayesian Network*. Following Chan *et al.* [15], the derivative $\frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N_{viol|\mathbf{c}}}}$ for a norm node N in $\mathcal{N}$ can be computed as follows.

$$\frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N_{viol|\mathbf{c}}}} = \frac{P(\mathbf{O}_{true}, N_{viol}|\mathbf{c})}{P(N_{viol}|\mathbf{c})} - P(\mathbf{O}_{true}|N_{ob}, \mathbf{c}) \tag{7}$$

For instance, for a norm node $SpdLim$ in the Bayesian Network of Fig. 3, where $\mathbf{O} = \{TripDur, Halted\}$, the left member of the difference in Eq. 7 is

$$\frac{P(\mathbf{O}_{true}, N_{viol}|\mathbf{c})}{P(N_{viol}|\mathbf{c})} = \frac{P(TripDur_{true}, Halted_{true}, SpdLim_{viol}|\mathbf{c})}{P(SpdLim_{viol}|\mathbf{c})} =$$
$$= \frac{P(TripDur_{true}, Halted_{true}|SpdLim_{viol}, \mathbf{c}) \cdot P(SpdLim_{viol}|\mathbf{c})}{P(SpdLim_{viol}|\mathbf{c})} =$$
$$P(TripDur_{true}|SpdLim_{viol}, \mathbf{c}) \cdot P(Halted_{true}|SpdLim_{viol}, \mathbf{c}) =$$
$$= P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) \cdot P(SafDst_{viol} \mid \mathbf{c}) +$$
$$P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) \cdot P(SafDst_{ob} \mid \mathbf{c})$$

while the right member of the difference in Eq. 7 is

$$P(\mathbf{O}_{true}|N_{ob}, \mathbf{c}) = P(TripDur_{true}, Halted_{true}|SpdLim_{ob}, \mathbf{c}) =$$
$$P(TripDur_{true} \mid SpdLim_{ob}, SafDst_{viol}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{ob}, SafDst_{viol}, \mathbf{c}) \cdot P(SafDst_{viol} \mid \mathbf{c}) +$$
$$P(TripDur_{true} \mid SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) \cdot P(SafDst_{ob} \mid \mathbf{c})$$

Therefore the derivative of Eq. 7 for a norm node $SpdLim$ in the Bayesian Network of Fig. 3 can be computed as:

$$P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) \cdot P(SafDst_{viol} \mid \mathbf{c}) +$$
$$P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) \cdot P(SafDst_{ob} \mid \mathbf{c}) -$$
$$P(TripDur_{true} \mid SpdLim_{ob}, SafDst_{viol}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{ob}, SafDst_{viol}, \mathbf{c}) \cdot P(SafDst_{viol} \mid \mathbf{c}) -$$
$$P(TripDur_{true} \mid SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid$$
$$SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) \cdot P(SafDst_{ob} \mid \mathbf{c})$$

The RRS for a norm N determines the change in $P(N_{viol}|\mathbf{c})$ that is estimated, based on observed data from MAS execution, to be required in order to make the norm set $\mathcal{N}$ effective.

Naive sensitivity analysis suggests to change (increase or decrease) the violations of norms of the amount determined by the corresponding RRSs. The sign of the required revision strength determines whether it is required to reduce (negative RRS) or to increase (positive RRS) violations of a norm, i.e., it determines the direction of the required revision. The value of the RRS determines the intensity of the required change. For instance if $\Delta\theta_{N1_{viol|\mathbf{c}}} = +0.2$ and $\Delta\theta_{N2_{viol|\mathbf{c}}} = -0.5$, then the suggestion is to increase $P(N1_{viol})$ of 0.2 and to decrease $P(N2_{viol})$ of 0.5.

This strategy computes the RRS for a norm, without considering that a change could be applied, at the same time, also to other norms. In other words, the RRS for a norm N is computed as if no change in the probability of violating any other norm could happen (from this the term *naive*). However, when determining the RRS for a norm, *Naive sensitiity analysis* considers all possible values of the other norms. Therefore, this strategy may result robust to unexpected changes in the probability of violating other norms when changing the sanctions.

n-CPT sensitivity analysis. Determine the required revision strength for all norms together, by solving, following Chan et al. [15], equation 8 for the n parameters $\Delta\theta_{N1_{viol}}, ..., \Delta\theta_{Nn_{viol}}$. Let $co(\mathbf{N}, i)$ be the set of all possible combinations of i norm nodes from the set $\mathbf{N}$, and, given a set $\mathbf{M} = \{N1, ..., Nm\} \subseteq \mathbf{N}$ of norm nodes, let $\frac{\delta^i}{\delta\theta\mathbf{M}_{viol|\mathbf{c}}}$ be the Leibniz's notation for the i-th partial derivative $\frac{\delta^i}{\delta\theta N1_{viol|\mathbf{c}}...\delta\theta Nm_{viol|\mathbf{c}}}$ for $N_j \in \mathbf{M}$.

$$P(\mathbf{O}_{true}|\mathbf{c}) + \sum_{i=1}^{n}\left[\sum_{\mathbf{M}\in co(\mathbf{N},i)}\left(\frac{\delta^i P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta\mathbf{M}_{viol|\mathbf{c}}} \cdot \prod_{N\in\mathbf{M}}\Delta\theta_{N_{viol|\mathbf{c}}}\right)\right] \geqslant t_{oa} \quad (8)$$

Solving equation 8 means to determine a list of n values $\Delta\theta_{N_{viol}}$, one for each norm node $N \in \mathbf{N}$. To do so, first of all it is required to compute: the n first partial derivatives $\frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N_{viol|\mathbf{c}}}}$ (one for each norm $N \in \mathbf{N}$); the second partial derivatives for the $\binom{n}{2}$ possible combinations of two norm nodes from $\mathbf{N}$; the third partial derivatives for the $\binom{n}{3}$ possible combinations of three norm nodes from $\mathbf{N}$; and so on until the n-th partial derivative $\frac{\delta^n P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N1_{viol|\mathbf{c}}}...\delta\theta_{Nn_{viol|\mathbf{c}}}}$.

For instance, in the case of $\mathbf{N} = \{N1, N2\}$, we have that $n = 2$, $co(\mathbf{N}, 1) = \{\{N1\}, \{N2\}\}$, and $co(\mathbf{N}, 2) = \{\{N1, N2\}\}$, and inequality 8 corresponds to inequality 9.

$$P(\mathbf{O}_{true}|\mathbf{c}) +$$
$$\frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N1_{viol|\mathbf{c}}}} \cdot \Delta\theta_{N1_{viol|\mathbf{c}}} + \frac{\delta P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N2_{viol|\mathbf{c}}}} \cdot \Delta\theta_{N2_{viol|\mathbf{c}}} +$$
$$\frac{\delta^2 P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N1_{viol|\mathbf{c}}}\delta\theta_{N2_{viol|\mathbf{c}}}} \cdot \Delta\theta_{N1_{viol|\mathbf{c}}}\Delta\theta_{N2_{viol|\mathbf{c}}} \geqslant t_{oa} \quad (9)$$

The first partial derivatives in Eq. 8 can be computed as per Eq. 7, while the second partial derivative, in the case of two norms (as it is in Eq. 9), can be computed as per Eq 10.

$$\frac{\delta^2 P(\mathbf{O}_{true}|\mathbf{c})}{\delta\theta_{N1_{viol|\mathbf{c}}}\delta\theta_{N2_{viol|\mathbf{c}}}} =$$
$$P(\mathbf{O}_{true}|N1_{viol}, N2_{viol}, \mathbf{c}) + P(\mathbf{O}_{true}|N1_{ob}, N2_{ob}, \mathbf{c}) -$$
$$P(\mathbf{O}_{true}|N1_{viol}, N2_{ob}, \mathbf{c}) - P(\mathbf{O}_{true}|N1_{ob}, N2_{viol}, \mathbf{c}) \quad (10)$$

If we consider the running example from Fig. 3, the derivative in Eq. 10 can be computed as follows.

$$P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) \cdot P(Halted_{true} \mid SpdLim_{viol}, SafDst_{viol}, \mathbf{c}) +$$
$$P(TripDur_{true} \mid SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid SpdLim_{ob}, SafDst_{ob}, \mathbf{c}) -$$
$$P(TripDur_{true} \mid SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) \cdot P(Halted_{true} \mid SpdLim_{viol}, SafDst_{ob}, \mathbf{c}) -$$

$$P(\mathit{TripDur}_{true} \mid \mathit{SpdLim}_{ob}, \mathit{SafDst}_{viol}, \mathbf{c}) \; \cdot \; P(\mathit{Halted}_{true} \mid$$
$$\mathit{SpdLim}_{ob}, \mathit{SafDst}_{viol}, \mathbf{c}).$$

After determining the values of the opportune derivatives, as above reported, inequality 8 can be solved by solving the following optimization problem.

$$\begin{aligned} &\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} \; f(\mathbf{x}) \\ &\text{subject to: } t_{oa} - f(\mathbf{x}) \leqslant 0 \end{aligned} \qquad (11)$$

where $\mathbf{x} = (x_1, ..., x_n)$ is a vector of real values, such that x_i is a possible value for $\Delta\theta_{Ni_{viol\|\mathbf{c}}}$ and $f(\mathbf{x})$ is the left member of inequality 8. Notice that the constraint to which the optimization problem is subject to corresponds to the canonical form of Eq. 8. Solving the optimization problem 11 means to determine the minimum values for the n parameters $\Delta\theta_{N1_{viol\|\mathbf{c}}}, ..., \Delta\theta_{Nn_{viol\|\mathbf{c}}}$ that satisfy the desired constraint of inequality 8 (i.e., the probability of achieving the objectives, after applying the required change in the probability of violating the enforced norms, is above the desired threshold t_{oa}).

Analogously to *naive sensitivity analysis*, *n-CPT sensitivity analysis* suggests to change (increase or decrease) the violations of norms of the amount of the corresponding *RRS*s determined by solving inequality 8. For instance, in the case of two norms, if $\Delta\theta_{N1_{viol\|\mathbf{c}}} = +0.2$ and $\Delta\theta_{N2_{viol\|\mathbf{c}}} = -0.5$, then the suggestion is to increase $P(N1_{viol})$ of 0.2 and to decrease $P(N2_{viol})$ of 0.5. Differently from the previous strategy, however, such values are obtained taking into account the change applied at the same time to the probability of violating all norms (instead of applying a change only one norm at a time).

5.1.3 Category-based strategies

Category-based strategies classify norms into different categories, based on their exhibited norm violation and on their relationship with the system-level objectives discovered at runtime, and determine an adequate revision for each norm based on their category. We present two strategies of this type (*Synergy+MLE* and *State-based*), based on two heuristic strategies presented in [24] and used to suggest a revision of regimented norms. In this paper we adapt them to support the revision of sanctions.

Synergy+MLE. This strategy is based on the PUREBN strategy presented in [24]. We distinguish between norms that are *more useful when obeyed* (*useful-ob* for brevity) or *more useful when violated* (*useful-viol*). Furthermore, norms can also be either *most likely obeyed* when the objectives are not achieved (*likely-ob* for brevity) or *most likely violated* (*likely-viol)*. In order to distinguish between *useful-ob* and *useful-viol* we calculate the combined synergy $\mathbf{n}_d$ (as per Eq. 5). Norms with value *ob* in $\mathbf{n}_d$ are *useful-ob*, norms with value *viol* in $\mathbf{n}_d$ are *useful-viol*. In order to distinguish between *likely-ob* and *likely-viol*,

instead, we determine the *most likely explanation* [37] **mle** for $\mathbf{O}_{false}$ in context **c**, as follows (with **ov** defined as per Eq. 5).

$$\mathbf{mle} = argmax_{\mathbf{n} \in \mathbf{ov}} \ P(\mathbf{n} \ |\mathbf{O}_{false} \wedge \mathbf{c}) \tag{12}$$

Norms with value *ob* in **mle** are *likely-ob*, norms with value *viol* in **mle** are *likely-viol*.

Synergy+MLE suggests to increase violations of norms belonging to category *useful-viol* (*more useful when violated*); to reduce violations of norms belonging to both categories *useful-ob* and *likely-viol* (norms that are *more useful when obeyed*, but most likely *violated* when the objectives are not achieved); and to do nothing for, or reduce violations of, norms belonging to both categories *useful-ob* and *likely-ob* (norms that are *more useful when obeyed*, and most likely *obeyed* when the objectives are not achieved).

The original PUREBN strategy [24] included the concept of *harmful norm*: a norm that is better when disabled. The suggestion of PUREBN for harmful norms is to disable them. In this paper we only consider active norms and we focus on the sanction revision, thereby omitting specific suggestions for harmful norms. However, a suggestion of increasing violation of a norm *N*, may lead to enforce *N* with a sanction equals to 0. In this paper, enforcing a norm *N* with a sanction of 0 corresponds to disabling *N*.

Finally, note that *Synergy+MLE* is a refinement of *Combined synergy* strategy. In addition to the combined synergy, this strategy also takes into account the most likely explanation for the objectives being not achieved, in terms of obedience or violation of norms.

State-based. This strategy, based on the STATEBASED strategy presented in [24], considers, in addition to the classification of norms described for strategy *Synergy+MLE*, information about the system state in context **c**. In particular, as illustrated in Fig. 5, the system can be in four states with respect to the average norm obedience, calculated as the mean $ns = mean_{N \in \mathbf{N}} P(N_{ob}|\mathbf{c})$, and the objectives achievement probability $oa = P(\mathbf{O}_{true} \mid \mathbf{c})$.

- In state A, norms are sufficiently obeyed, but this does not lead to sufficient objectives achievement (i.e., $ns \geqslant t_{ns}$ and $oa < t_{oa}$ for some given t_{ns} and t_{oa}).
- In state B, norms are not sufficiently obeyed and also objectives are not achieved (i.e., $ns < t_{ns}$ and $oa < t_{oa}$).
- In state C, the objectives are achieved even though the norms are not obeyed (i.e., $ns < t_{ns}$ and $oa \geqslant t_{oa}$).
- In state D, (the desired state of the system) the norms are satisfied and the objectives are achieved (i.e., $ns \geqslant t_{ns}$ and $oa \geqslant t_{oa}$).

If the system is in state A, *State-based* suggests to increase violations of norms belonging to both categories *useful-viol* and *likely-ob*, i.e., norms that are more useful when violated but most likely obeyed when the objectives are not achieved, if any. Otherwise, *State-based* suggests to do nothing for (or to
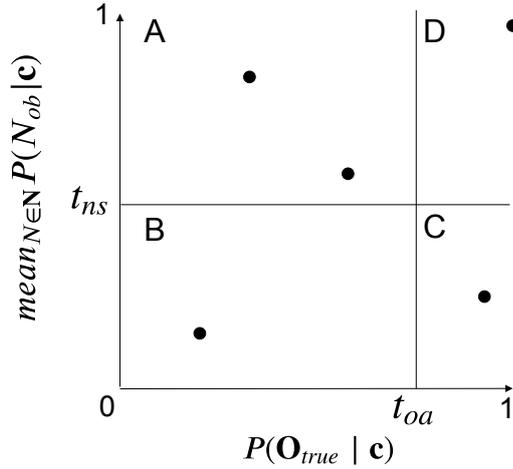
Fig. 5: System states (points) in four states (A-D) w.r.t. average norm obedience and objectives achievement probability.

reduce violations of) the current norm set. In this case, there is probably some aspect of the system that has not been considered during its design, for the current norms are mostly obeyed and they are most useful when obeyed, but the system-level objectives are not achieved as desired. If the system is in state B, *State-based* suggests to reduce violations of norms belonging to both categories *useful-ob* and *likely-viol*, i.e., norms that are more useful when obeyed but most likely violated when the objectives are not achieved. It also suggests to increase violations of norms belonging to category *useful-viol*, i.e., norms that are more useful when violated. If the system is in state C, finally, *State-based* suggests to increase violations of norms belonging to both categories *useful-viol* and *likely-viol*, if any. Otherwise it suggests to decrease violations of norms belonging to both categories *useful-ob* and *likely-viol*.

While *Synergy+MLE* suggests for all the norms in $\mathcal{N}$ the most adequate revision to perform, *State-based* considers the global state of the system and suggests to revise only a specific category of norms at every iteration (for the norms that do not belong to the category above mentioned it is suggested to do nothing). In case of high number of norms enforced, this strategy may significantly reduce the number of revisions that need to be performed at every step.

5.2 Sanctions Revision

Consider a norm $N = (p, s)$ and a revision of it $N' = (p, s')$, with $s' \neq s$. Let $P(N_{viol}|\mathbf{c})$ be N's exhibited norm violation (i.e., the probability of violating N monitored during system's execution) in context $\mathbf{c}$. We call *applied revision strength* the difference $UB(N'_{viol}, \mathcal{N}) - P(N_{viol}|\mathbf{c})$ between the upper bound

UB for violation of N' (as per Sec. 4.2) and the N's exhibited norm violation. For instance, in the example reported in Fig. 4, supposing that when enforcing $N_1 = (sp_{50}, 1)$ and $N_2 = (dist_2, 0)$ the exhibited norm violation of N_1 is 0.3, the applied revision strength when revising N_1 into $N'_1 = (sp_{50}, 2)$ is $0 - 0.3 = -0.3$, while the applied revision strength when revising N_1 into $N'_1 = (sp_{50}, 0)$ is $1 - 0.3 = 0.7$.

The strategies described in Sec. 5.1, provide, for each norm N in $\mathcal{N}$, a suggestion such as *reduce/increase* violations of N, do *nothing* with N, *reduce/increase* violations of N of a certain *amount RRS* (as per Sec. 5.1.2). Given these suggestions, and all possible sets of sanctions $\mathcal{S}$ that can be used to enforce norms in $\mathcal{N}$, we need to find a new set of sanctions such that the *applied revision strength* satisfies (as much as possible) the given suggestions.

A trivial solution is to systematically go through all elements in $\mathcal{S}$ until the desired sanction set (if it exists) is found. Such solution is however computationally expensive, as the number of possible sanction sets is $(max(\mathcal{B}) + 2)^n$, with $max(\mathcal{B}) + 2$ maximum budget among all agent types ($max(\mathcal{B})$) plus sanction 0 and sanction $max(\mathcal{B}) + 1$, and n number of norms.

In the following, we propose a simple alternative way to explore the search space $\mathcal{S}$ that can be used in case of a population of consistent agent types as per Def. 5. With a population of consistent agent types, according to Prop. 3, the upper bound of the probability of violating norms decreases monotonically when any sanction increases. This means that given a sanction set, and the exhibited norm violation for each enforced norm, if we desire to apply a negative revision strength, we need to move towards higher values of sanctions. To apply a positive revision strength, instead, we could change in any way the sanctions (even though typically we should move towards lower values of sanctions), since the currently exhibited norm violation could be lower than the upper bound of norm violation with an higher sanction.

Under the assumption of consistent agent types, we can reduce therefore the exploration of $\mathcal{S}$ by directing the search towards the desired values of sanctions. For instance, suppose to have two norms $N1 = (p_1, s_1)$ and $N2 = (p_2, s_2)$, and a list of suggestions $sugg = (reduce, increase)$ for a context $\mathbf{c}$ (i.e., it is suggested to reduce violations of norm $N1$ and to increase violations of norm $N2$ in context $\mathbf{c}$). Given $sugg$, we need to look for a new sanction set $\{s'_1, s'_2\}$ such that $UB(N1'_{viol}, \mathcal{N}') < P(N1_{viol}|\mathbf{c})$ and $UB(N2'_{viol}, \mathcal{N}) > P(N2_{viol}|\mathbf{c})$, with $\mathcal{N}' = \langle N1', N2' \rangle$. We can therefore reduce the search space to the subset of $\mathcal{S}$ such that $s'_1 \geqslant s_1$ and $s'_2 \neq s_2$.

Algorithm 1 reports the pseudo-code of a procedure to perform such search.

Notice that if preferences are not consistent, we have no guarantees that by moving towards higher values of sanctions we will not increase violations of norms, since Prop. 3 does not hold in the general case (i.e., for preferences that are not consistent). Despite this, one may still heuristically explore $\mathcal{S}$ by using Algorithm 1, also when not all preferences are consistent.

Algorithm 1 is invoked when a suggestion of norm revision has been determined with one of the strategies of Sec. 5.1 after a norm revision is triggered, and there is at least one sanction set that has not been tried previously in

context **c**. If a sanction set has already been tried, we know it is not effective (otherwise no further norm revision would have been triggered). If all possible sanction sets have been already tried (omitted from Algorithm 1), then the sanction set that, when enforced, maximizes $P(\mathbf{O}_{true} \mid \mathbf{c})$ is selected.

Algorithm 1 The algorithm for the selection of a new sanction set

---

 Input: cs the current sanction set,
E the norms' exhibited violation,
$sugg$ the n suggestions,
UB the upper bound matrix
RRS a (possibly empty) list of required revision strengths
c the context
 Output: a new sanction set $sugg$ to enforce in **c**

1: **function** GETSANCTIONSET(cs, E, $sugg$, UB, RRS, **c**)
2:     $op \leftarrow \{reduce : [0,1], increase : [0,-1,1], nothing : [0,1]\}$
3:     $comb \leftarrow \times_{i=1}^{n} op[sugg[i]]$
4:     **for all** $ch \in [1, max(\mathcal{B})]$ **do**
5:         **for all** $o \in comb$ **do**
6:             $ns \leftarrow [cs[i] + o[i] \cdot ch \mid \forall i \in [1,n]]$
7:             **if** ISNEWSANCTIONSET(ns,**c**) **then**
8:                 **if** SUGGSAT(cs, ns, $sugg$, UB, E, RRS, **c**) **then**
9:                     **return** ns
    return cs

The algorithm takes as input: the list of currently enforced sanctions cs; the exhibited violation of the enforced norms E; the list $sugg$ of suggestions obtained with one of the strategies of Sec. 5.1 (a value *reduce* (or *increase*, or *nothing*) in $sugg[i]$ corresponds to a suggestion to reduce (increase, or do nothing with) violations of the i-th norm); a matrix UB containing the upper bounds for norms violations as per Fig. 4; a list RRS of required revision strengths (empty if no sensitivity-based strategy is used); and the context **c**. As output, Algorithm 1 returns a (possibly new) list of sanctions to use to enforce norms in context **c**.

The algorithm explores the possible sanction sets starting from the current sanction set cs. The first step is to determine the subset of $\mathcal{S}$ to explore. Notice that, if a *reduce* suggestion has been given for norm i (i.e., $sugg[i] = reduce$), the new sanction for norm i must be greater of equal than the current one (i.e., $ns[i] \geqslant cs[i]$). This means that $ns[i]$ has to be equal to $cs[i] + ch$, with $ch \geqslant 0$ amount of change. Conversely, if $sugg[i] = increase$, $ns[i]$ has to be equal to either $cs[i] + ch$ or $cs[i] - ch$. If we put these cases together (line 6 of Algorithm 1), $ns[i]$ has to be equal to $cs[i] + o[i] \cdot ch$, with $o[i] \in \{0,1\}$ if $sugg[i] = reduce$, and $o[i] \in \{0,-1,1\}$ if $sugg[i] = increase$, and $ch > 0$ amount of change.

Variable $comb$ (line 3) is a set of all possible combinations of operators $o[i]$ for each norm i, obtained from their suggestion $sugg[i]$ (see $op[sugg[i]]$, which retrieves from the labeled set op declared at line 2, the opportune list of operators given suggestion $sugg[i]$). For instance, supposing to have two norms $N1 =$

(p_1, s_1) and $N2 = (p_2, s_2)$, and a list of suggestions $sugg = (reduce, increase)$, we have that $op[sugg[1]] = [0, 1]$ and $op[sugg[2]] = [0, -1, 1]$ and $comb$ is the set of all possible combinations of operators in $op[sugg[1]]$ and $op[sugg[21]]$, i.e., $\{(0, 0), (0, -1), (0, 1), (1, 0), (1, -1), (1, 1)\}$, such that given a certain element $o \in comb$, $o[i]$ is the operator to apply to the change of sanction $cs[i]$.

The algorithm iterates through all possible changes that can be applied to sanctions (line 4). For each possible change, the algorithm iterates through all possible new sanction sets that can be obtained with the combinations of operators in $comb$ (lines 5-6). Notice that, by iteratively increasing the change, we explore the search space at increasing distance from the current sanction set. This means that if the algorithm finds a new (function isNewSanction-Set at line 7) sanction set ns that satisfies the given suggestions (function suggSat at line 8), such sanction set is also the closest possible to the current one.

Finally, if no new sanction set satisfying the suggestions is found, the current sanction set is returned. In this case, in our framework a random sanction set never tried before is enforced in context $\mathbf{c}$.

Notice also that function suggSat (line 8), whose purpose is to verify that a proposed sanction set satisfies the given suggestions, does not need to require that *all* suggestions are perfectly satisfied. In particular, especially when suggestions include also a required revision strength (i.e., when using sensitivity-based strategies), it may be more useful to search for a good-enough sanction set. For our experiments, described in Sec. 6, when list RRS is not empty, we keep track of the best new sanction set found so far (if not all suggestions are satisfied) and for every new sanction set tested we require at least 80% of suggestions to be satisfied. Furthermore in case of suggestion *nothing*, since unlikely, in our experiments, that the exhibited probability of a norm exactly corresponds to a value on its upper bound, we accept also a reduction of the probability of violating the norm of a small ϵ (we used $\epsilon = 0.1$).

After enforcing the new norm set $\mathcal{N}'$, obtained by revising the sanctions of norms in $\mathcal{N}$ according to the new sanction set obtained from Algorithm 1, we monitor the new behavior of the agents and detect the new exhibited norm violation $P(N'_{viol}|\mathbf{c})$, for each norm $N' \in \mathcal{N}'$. We call *actual revision strength* the difference $P(N'_{viol}|\mathbf{c}) - P(N_{viol}|\mathbf{c})$ between the exhibited norm violation of N' and N, with $N' = (p, s')$ and $N = (p, s)$.

6 Experimentation

We report on an experiment that investigates the process through which the norm-based supervision mechanism of Sec. 4 identifies an optimal system configuration. The object of our study consists of the strategies for norm revision proposed in Sec. 5. In particular, we study the process through which the norm-based supervision mechanism identifies an optimal system configuration when employing each of the six proposed strategies as possible *informed heuristics*

for defining the neighborhood of a configuration, i.e., the configurations where the sanctions of the enforced norms are revised as suggested by the heuristics.

We compare the results in terms of *convergence speed*. The convergence speed measures the number of steps (i.e., revisions of the sanctions of norms triggered) required by the heuristic strategies to make the norms effective in achieving the system-level objectives. This allows us to study the time efficiency of the norm revision strategies in refining sub-optimal norms at runtime.

6.1 Experimental Setting

Our experiment is run through a simulation[5] of the ring road scenario described in Sec. 3. Our implementation of the norm-based supervision mechanism of Sec. 4, as a modified version of hill climbing, is called SASS (Supervisor of Autonomous Software Systems)[6]. The supervisor performs a local search and stops when either (i) all the system configurations have been tried; or (ii) a local optimum (system configuration) is found that has objectives achievement probability $oa = P(\mathbf{O}_{true})$ above the desired threshold t_{oa}. The objectives achievement probability of a certain system configuration is *not* known to SASS before the configuration is actually enforced. Such probability is determined at runtime from simulation data, given the chosen system configuration. In this experimental setting, the last system configuration that is selected before stopping is called *optimal*, since either the objectives achievement is above the desired threshold or there is no other better configuration.

In the ring road scenario, we consider the two contextual variables *Vehicle density*, which can be *low* (40 cars on the ring road) or *high* (80 cars)[7]; and *Obstacle*, which is *true* when an obstacle is placed on the outer lane of the ring road. Each car in the simulation is an agent that acts according to its specific characteristics, beliefs and preferences. At each simulation step, every agent also deliberates about a number of things, including its desired speed and the minimum safety distance, whether and how much to accelerate or decelerate, whether to change lane to surpass or to move back to the outer lane, whether to activate the turn signals. Agents' decisions are based on their own internals, which are specific for each agent and unknown to the norm revision mechanism. In our simulations, when an agent equally prefers two alternatives x and y concerning the speed and safety distance (i.e., $x \sim y$), the agent applies a deterministic choice to determine what state of affair to pursue (i.e., simply the first one in the representation of the alternatives), instead of random choice.

In order to define norms and agents' preferences, we consider the set of propositional atoms $L = \{sp_{15}, sp_8, sp_3, dist_{0.5}, dist_1, sp_2\}$, with $AL =$

---

[5] For our experiment, we used the SUMO traffic simulator [36] and CrowdNav+RTX [45]

[6] The source code and the material for experiments' replication can be found in [26].

[7] The density values have been determined empirically based on the size of the ringroad used for the experiments.

(L_1, L_2) and $L_1 = \{sp_{15}, sp_8, sp_3,\}$ and $L_2 = \{dist_{0.5}, dist_1, dist_2\}$. Each element in L_1 represents a speed in m/s and each element in L_2 represents a safety distance in meters. Furthermore, we consider a language $\mathcal{B} = \{0, 1, 2\}$ for defining budgets and a language $\mathcal{S} = \{0, 1, 2, 3\}$ for defining sanctions of norms.

6.1.1 Agent Types

We experiment with four types of rational agents with consistent preferences (as per Def. 5). In the following we briefly describe such types, and we report in Appendix B the full preferences.

- *BraveRich* is a consistent basic preference that adheres to Def. 2b, i.e., where alternatives are ordered by propositional atom. It describes an agent type with a maximum budget of 4, that prefers to drive fast and to keep a short safety distance, and that gives priority to the short safety distance rather than to driving fast.
- *BraveMiddleClass* is a consistent preference composed by two basic preferences. The first basic preference $(A_1, \succeq_1)$ adheres to Def. 2b. The alternatives in A_1 are such that $A_1 = \{(\langle p_1, b_1 \rangle, \ldots, \langle p_n, b_n \rangle) \mid p_i \in L_i \ \& \ (b_1, \ldots, b_n) \in BL_1\}$, with $BL_1 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. The second basic preference $(A_2, \succeq_2)$ adheres to Def. 2a. The alternatives in A_2 are such that $A_2 = \{(\langle p_1, b_1 \rangle, \ldots, \langle p_n, b_n \rangle) \mid p_i \in L_i \ \& \ (b_1, \ldots, b_n) \in BL_2\}$, with $BL_2 = \{(0, 2), (1, 2), (2, 2), (2, 1), (2, 0)\}$. *BraveMiddleClass* describes an agent type similar to *BraveRich*, but that is willing to pay no more than 2 for a certain state of affairs. The alternatives in A_2 are ordered by required budget and, for consistency, they maintain the same relative order as in A_1.
- *BravePoor* is a consistent basic preference ordered by required budget, as per Def. 2a. It describes an agent type that equally prefers to drive fast or slow and to keep a short or long safety distance, but is not willing to pay anything to reach any state of affairs.
- *Cautious* is a consistent basic preference ordered by required budget, as per Def. 2a. It describes an agent type that equally prefers to drive slow or fast and to keep a long or short safety distance, and is not willing to pay anything to reach any state of affairs. Notice that this preference is equivalent to *BravePoor*, however due to the deterministic mechanism of choice of an alternative that our agents employ (i.e., the first one in the representation of the alternatives), these two agent types will exhibit different behaviors at runtime. For instance, even though states of affairs where sp_{15} and $dist_{0.5}$ hold are equally preferred to state of affairs where sp_3 and $dist_{0.5}$ hold, in both preferences, and they could both be chosen in the case of random choice, in our simulation, given enough budget, *BravePoor* will aim at a state of affair where sp_{15} and $dist_{0.5}$ hold, while *Cautious* will aim at a state of affair where sp_3 and $dist_{0.5}$ hold.

We consider three distributions of types of agents:

- *uniform*, the entire population of agents is uniformly distributed across the four types above described.
- *mostly compliant*, 75% of agents belongs to type *Cautious* and the rest is uniformly distributed across the remaining types.
- *mostly violating*, 75% of agents belongs to type *BraveRich* and the rest is uniformly distributed across the remaining types.

Note that despite our estimation of the preferences of the agents concerning speed and safety distance, we do not have any control on the exact speed or safety distance of the agents, which is internally and opaquely set by the agents, together with the rest of their behaviors.

6.1.2 Norms

We consider four ordered norm sets: $\mathcal{N}_{31} = \langle SpdLim_3, SafDst_1 \rangle$, $\mathcal{N}_{32} = \langle SpdLim_3, SafDst_2 \rangle$, $\mathcal{N}_{81} = \langle SpdLim_8, SafDst_1 \rangle$, $\mathcal{N}_{82} = \langle SpdLim_8, SafDst_2 \rangle$, with $SpdLim_x = (sp_x, s_1)$, $SafDst_y = (dist_y, s_2)$, $x \in \{3, 8\}$, $y \in \{1, 2\}$ and s_1 and s_2 sanctions in $\mathcal{S}$.

Fig. 6 illustrates the upper bounds of the probability of violating the two norms $SpdLim_x$ and $SafDst_y$ above defined (as per Sec. 4.2) for the three agent type distributions. Notice that the reported upper bounds hold for all combinations of the values x and y above defined (i.e., values of speed limit and minimum safety distance). This is due to the types of agents that we considered for our experiments. *BraveRich* prefers to keep a speed of 15 m/s by maintaining a short safety distance and it is willing to pay a sanction of 2 for each of these aspects. When the sanction of a norm is above 2 this agent is compliant with the norm, *regardless of the value of the speed limit*, because the agent has no budget for violating the norm. *BraveMiddleClass* is analogous to *BraveRich* but with a maximum budget of 1 for the violation of a norm: up to sanction 1 *BraveMiddleClass* has reason to violate a norm (it also prefers to go at a speed of 15 m/s by maintaining a short safety distance), while when the sanction of a norm is above 1 *BraveMiddleClass* is compliant. Finally *BravePoor* and *Cautious* have reason to violate the norms only when their sanctions is 0. With higher sanctions, these agent types are compliant.

Furthermore notice that, since all the agent types that we considered are consistent as per Def. 5, the upper bounds reported in Fig. 6 satisfy Prop. 3: when increasing the sanction of only one norm the upper bound of violating the other norm never increases. This allows us to take advantage, in our experiments, of Algorithm 1 for the selection of a new sanction set.

6.2 Experiments

By combining the three distributions of agents of Sec. 6.1.1 with the four norm sets of Sec. 6.1.2, we derived 12 different experiments. We ran a simulation of the ring road for each of the 12 experiments and we collected data about norm obedience and objective achievement in the four different operating contexts

(a) *uniform*

(b) *mostly compliant*

(c) *mostly violating*
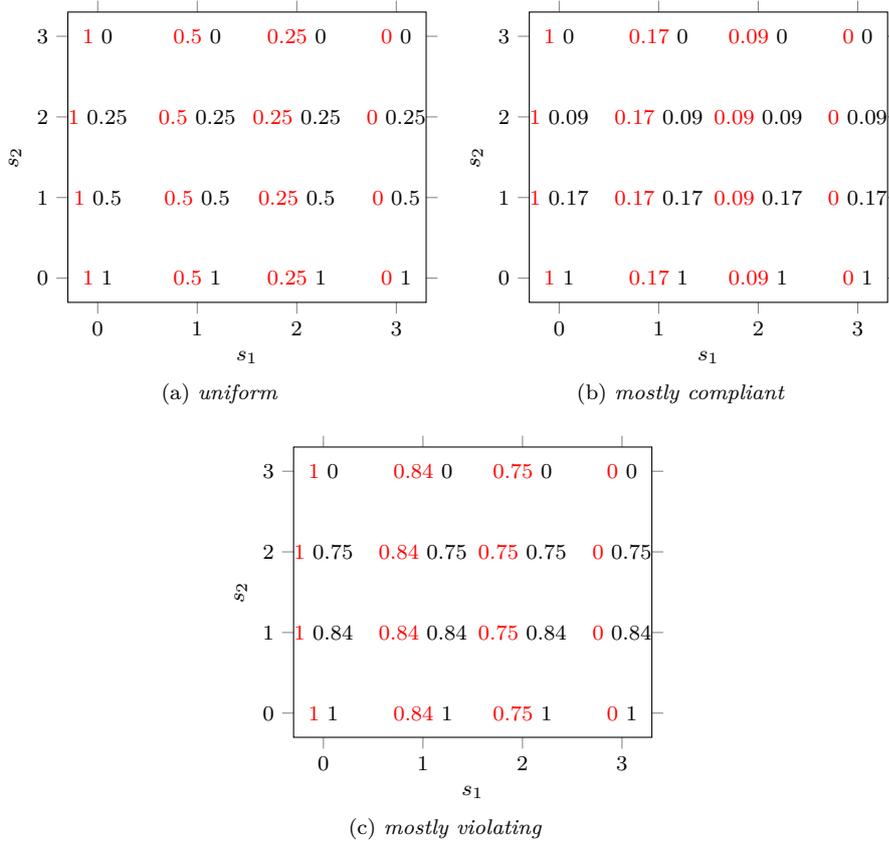
Fig. 6: Upper bound of the probability of violating norms $SpdLim_x = (sp_x, s_1)$ (red) and $SafDst_y = (dist_y, s_2)$ (black) with different agent type distributions. In each subfigure the x-axis represents the sanction s_1 of norm $SpdLim_x$, while the y-axis represents the sanction s_2 of norm $SafDst_y$.

c1 $= VehicleDensity_{low} \wedge Obstacle_{false}$, **c2** $= VehicleDensity_{low} \wedge Obstacle_{true}$, **c3** $= VehicleDensity_{high} \wedge Obstacle_{false}$, **c4** $= VehicleDensity_{high} \wedge Obstacle_{true}$. This means that during a simulation, the contexts in which the cars on the ring road operate changes three times (for a total of fours different operating contexts in each simulation). During the simulations, we monitored the behavior of the cars and sanctioned each car that violated one of the enforced norms. A car sanctioned for the violation of a norm N was not sanctioned anymore for violations of norm N until it completed a full loop of the ring road. The Boolean value of the system-level objectives was measured every 25 simulation steps. The objective *TripDur* was considered achieved if, on average in the 25 steps, the cars on the ring road took less than 2.5 times the theoretical

average trip time[8] to complete a loop of the ring road. The objective *Halted* was considered achieved if, on average in the 25 steps, less than $x\%$ of cars were halted on the ring road, with $x = 25$ if the density of vehicles on the ring road is high, and $x = 5$ if the density of vehicles is low[9]. A car in SUMO is considered halted if its speed is below $0.1m/s$. Cars could be halted on the ring road for several reasons. For example, the presence of an obstacle may force them to stop and wait for the right moment to surpass the obstacle or breaking waves may force cars to temporary slow down significantly to avoid collisions.

In every experiment that we perform, the system has n^m possible configurations, with n possible sanction sets and m different operating contexts. Since the speed of convergence to an optimal solution depends on the initial system configuration (i.e., a different amount of revisions may be required starting from different initial configurations), we execute each strategy starting from each possible configuration and we calculate statistics information (i.e., median, maximum, mean and standard deviation) concerning the convergence speed in the different executions. To keep our experimentation's time manageable, in our experiments we considered only 2 of the 4 operating contexts: **c2** and **c3**. This allowed us to reduce the number of possible configurations from 16^4 to $16^2 = 256$: 16 possible sanction sets for the enforced norms in any of the 2 contexts. Fig. 7 shows the probability $P(\mathbf{O}_{true})$ obtained with the 256 configurations in each of the 12 experiments and highlights the optimal configurations (the configurations s.t. $P(\mathbf{O}_{true}) \geqslant t_{oa}$). Every dot in Fig. 7 represents the probability of achieving the objectives during a simulation with a certain system configuration (i.e., $P(\mathbf{O}_{true})$), considering both the contexts **c2** and **c3**. In each sub-figure (one per experiment) we see therefore 256 dots, one per system configuration. Notice that in the 12 experiments, the distribution of the 256 configurations w.r.t. the probability of achieving the system-level objectives is different. In other words, a certain system configuration c (i.e., enforcing norms with certain sanctions in the two contexts **c2** and **c3**) can be effective in an experiment but ineffective in another experiment. This makes the 12 experiments independent, thereby increasing the generality of our results.

For each of the 12 experiments, we defined a different t_{oa} as indicated in Table 2, which summarizes the entire experimental setting. The different thresholds allow us to test our strategies with different degrees of difficulty (i.e., number of optimal configurations to be found).

---

[8] The theoretical trip time is $\sum_{t_i \in \mathcal{T}} d_i \times t_{i,\mathcal{N}}$, with $\mathcal{T}$ being the set of agent types, d_i being the percentage of agents of type t_i, and $t_{i,\mathcal{N}}$ being the theoretical time needed by t_i to complete a loop in case of free ring road when norm set $\mathcal{N}$ is enforced.

[9] Values for the evaluation of the objectives were determined based on some preliminary experimentation with the ring road simulation in order to retrieve a variegate set of experiments.
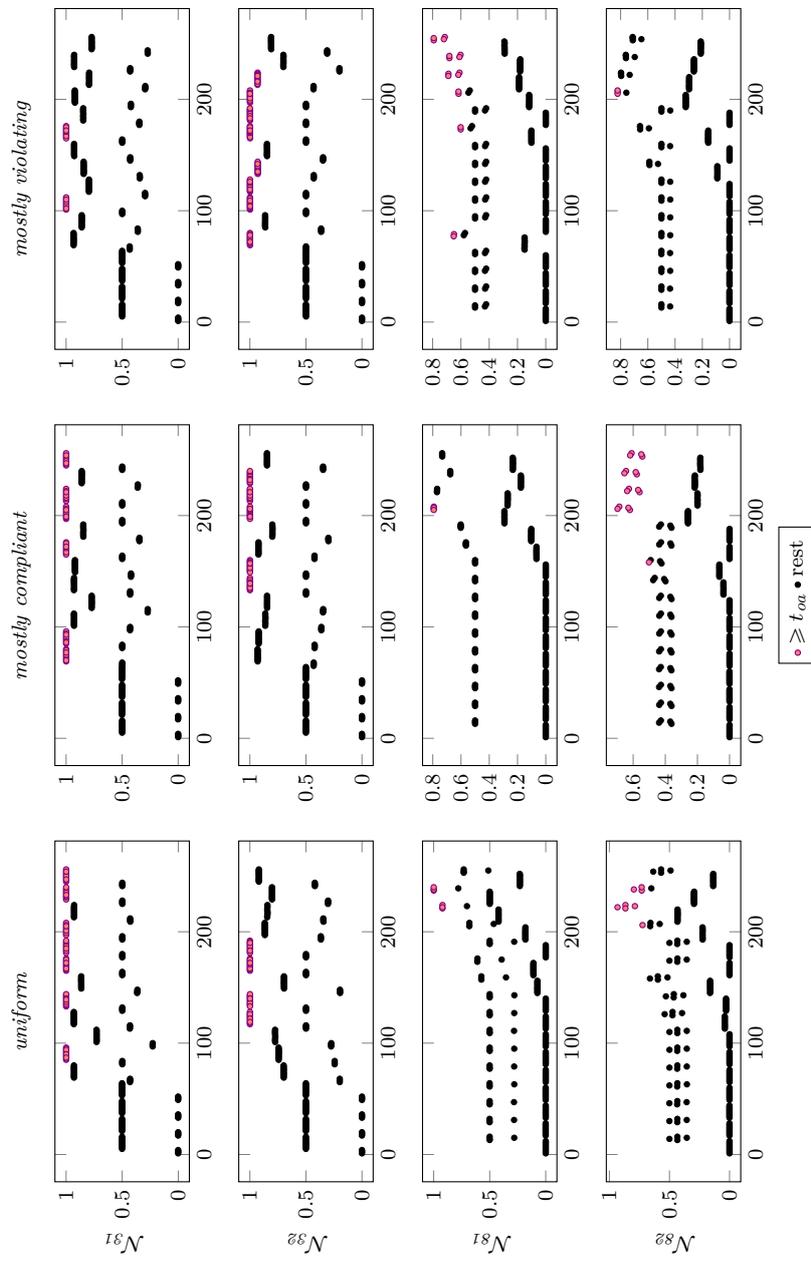
Fig. 7: Probability of objectives achievement (y-axis) for the 256 tried configurations (x-axis) in the 12 experiments.

Exp.	ID	Agents distribution	Norm set	Contexts	t_{oa}	Optimal
1	*DUN31*	*uniform*	$\mathcal{N}_{31}$	**c2, c3**	0.99	84/256 (32.8%)
2	*DUN32*	*uniform*	$\mathcal{N}_{32}$	**c2, c3**	0.99	48/256 (18.7%)
3	*DUN81*	*uniform*	$\mathcal{N}_{81}$	**c2, c3**	0.9	6/256 (2.3%)
4	*DUN82*	*uniform*	$\mathcal{N}_{82}$	**c2, c3**	0.7	8/256 (3.1%)
5	*DCN31*	*mostly compliant*	$\mathcal{N}_{31}$	**c2, c3**	0.99	72/256 (28.1%)
6	*DCN32*	*mostly compliant*	$\mathcal{N}_{32}$	**c2, c3**	0.99	60/256 (23.4%)
7	*DCN81*	*mostly compliant*	$\mathcal{N}_{81}$	**c2, c3**	0.79	4/256 (1.6%)
8	*DCN82*	*mostly compliant*	$\mathcal{N}_{82}$	**c2, c3**	0.5	17/256 (6.6%)
9	*DVN31*	*mostly violating*	$\mathcal{N}_{31}$	**c2, c3**	0.99	24/256 (9.4%)
10	*DVN32*	*mostly violating*	$\mathcal{N}_{32}$	**c2, c3**	0.9	96/256 (37.5%)
11	*DVN81*	*mostly violating*	$\mathcal{N}_{81}$	**c2, c3**	0.6	18/256 (7%)
12	*DVN82*	*mostly violating*	$\mathcal{N}_{82}$	**c2, c3**	0.8	3/256 (1.2%)

Table 2: The setting of the 12 experiments.

6.3 Analysis of the Results

Table 3 reports the results concerning the steps required by the supervision mechanism to find an optimal configuration in the 12 experiments when employing each of the six proposed revision strategies. In particular, we report the median, the maximum, the average, and the standard deviation of the number of steps. We highlight the best performing strategies in each experiment.

On average, all the strategies required a limited number of steps to find an optimal configuration in almost all experiments. In the 12 experiments, while the number of optimal configurations to be found ranges from 3 to 96 out of 256 configurations, on average the strategies never required more than 52 steps to find one of those configurations (see columns Avg (σ) in Table 3, where σ is the standard deviation), with a minimum of 0 for all strategies (trivially in the cases the initial configuration is optimal, not reported in Table 3), a maximum of 218 in the most difficult scenario (see columns Max of experiment *DVN82*), and a median value never above 35 steps.

If we look at the average values, the strategy that performed less well in the 12 experiments is *Naive sensitivity analysis*, which, in order to find an optimal configuration among the 256 possible configurations, required an average number of steps between 1 and 52. The strategy that, on average, performed best, instead, is *n-CPT sensitivity analysis*, requiring an average number of steps between 2 and 12. In particular, these results show that when using *n-CPT sensitivity analysis*, on average, about 6 norm revisions were triggered by the norm-based supervision mechanism before finding a configuration where the system-level objectives were achieved as desired.

Despite *n-CPT sensitivity analysis* performed, on average, better than the other strategies in the 12 experiments, the results show that using that strategy was mostly advantageous when very few configurations were optimal among all the possible ones. In particular, *n-CPT sensitivity analysis* appeared to be more effective than the other strategies when the number of optimal configurations was lower than 2% of all the configurations. For instance, in experiment *DCN81* (1.6% of configuration are optimal), never more than 13 steps were

Table 3: Comparison of the strategies in terms of number of steps required to find an optimal solution.

Exp.	Naive synergy			Combined synergy			Naive sensitivity			n-CPT sensitivity			Synergy+MLE			State-based		
	Median	Max	Avg (σ)	Median	Max	Avg (σ)	Median	Max	Avg (σ)	Median	Max	Avg (σ)	Median	Max	Avg (σ)	Median	Max	Avg (σ)
DUN31	1.00	4	1.13 (1.03)	1.00	4	1.13 (1.03)	1.00	6	1.43 (1.40)	2.00	8	2.20 (1.98)	1.00	4	1.13 (1.03)	1.00	4	1.13 (1.03)
DUN32	3.00	14	3.15 (2.79)	3.00	14	3.12 (2.68)	2.00	11	2.48 (2.34)	4.00	8	3.40 (2.26)	2.00	13	2.79 (2.57)	2.00	11	3.12 (2.82)
DUN81	5.00	59	5.51 (5.63)	7.00	93	9.12 (9.46)	15.00	160	16.50 (15.45)	14.00	19	12.41 (4.21)	6.00	98	7.08 (7.50)	5.00	152	5.81 (10.10)
DUN82	6.00	23	6.05 (4.15)	6.00	61	6.91 (5.91)	7.00	42	6.54 (4.48)	13.00	19	11.72 (4.44)	7.00	22	6.75 (4.13)	6.00	25	6.43 (3.94)
DCN31	1.00	4	1.19 (1.00)	1.00	4	1.19 (1.00)	1.00	4	1.38 (1.14)	2.00	6	1.56 (1.34)	1.00	4	1.19 (1.00)	1.00	4	1.25 (1.03)
DCN32	2.00	11	2.10 (2.06)	2.00	9	2.09 (1.96)	3.00	10	3.29 (2.69)	4.00	10	3.82 (2.76)	2.00	8	2.11 (1.88)	1.00	9	1.86 (1.82)
DCN81	22.00	216	33.44 (36.92)	21.00	149	28.00 (24.79)	24.00	198	36.64 (33.58)	9.00	13	7.97 (2.64)	22.00	178	32.49 (30.12)	8.50	104	10.44 (10.36)
DCN82	3.00	8	3.39 (2.19)	7.00	14	5.95 (3.78)	9.00	12	6.68 (4.31)	8.00	12	6.79 (2.99)	7.00	15	6.26 (3.82)	5.00	11	4.57 (2.84)
DVN31	7.00	18	6.89 (4.72)	7.00	20	7.14 (4.90)	5.00	22	5.70 (4.27)	6.00	10	4.62 (2.94)	7.00	19	7.23 (4.95)	7.00	20	7.01 (4.66)
DVN32	1.00	7	1.44 (1.55)	1.00	6	1.43 (1.52)	1.00	8	1.38 (1.62)	1.00	6	1.36 (1.37)	1.00	8	1.23 (1.44)	1.00	5	1.09 (1.14)
DVN81	4.00	12	5.03 (3.56)	7.00	16	6.70 (4.74)	3.00	6	3.11 (1.71)	6.00	13	5.78 (2.74)	7.00	16	6.61 (4.63)	7.00	15	6.71 (4.70)
DVN82	8.00	218	20.63 (33.75)	20.00	204	38.43 (44.11)	35.50	159	52.32 (42.90)	9.00	162	12.20 (18.37)	19.00	217	35.97 (41.52)	19.50	191	32.75 (38.35)

required to find an optimal configuration when employing *n-CPT sensitivity analysis*, while the other strategies required a maximum number of steps between 104 and 216. Furthermore, while the median number of step is 9 with *n-CPT sensitivity analysis*, the median number of steps with the other strategies is more than twice. One exception is *State-based*, which in such experiment required an average number of steps similar to *n-CPT sensitivity analysis* and an even lower median. *State-based*, however, exhibited an higher variance, requiring in some executions up to 104 steps. In experiment *DVN82* (1.2% of configuration are optimal), while all other strategies (including *State-based*) required an average number of steps between 20 and 52, *n-CPT sensitivity analysis* was able to find on average an optimal configuration in about 12 steps.

If we consider, instead, simpler experiments (e.g., *DUN31* or *DCN31*), *n-CPT sensitivity analysis* did not outperform significantly the other strategies. In fact, if we consider the average number of steps, among all the strategies, *Naive synergy* outperformed (even though by few steps) all the others in 5 experiments, requiring in all of them less then 6 steps to find an optimal configuration. Furthermore in 8 experiments the average number of steps required by *Naive synergy* was below the average between the different algorithms. *State-based* had similar performances to *Naive synergy* and, even though it was the absolute best strategy in only 3 experiments in terms of average number of steps, in 8 experiments out of 12 it exhibited the lowest median value.

Fig. 8 plots the percentage of configurations explored in the 12 experiments by the six strategies before finding an optimal one. In most experiments, all algorithms required to explore less than 10% of all configurations. The only cases that required to explore more than 10% of configurations were experiments *DCN81* and *DVN82*, where the number of optimal configuration to be found was less than 2%. Fig. 8 emphasizes that all proposed strategies performed similarly, with the exception of *n-CPT sensitivity analysis*, which did not show a degradation in the cases of very few optimal configurations and required to explore a significantly lower number of configurations.

The values in Table 3 and in Fig. 8 concern the absolute number of steps required, and configurations explored, to find one of the optimal configurations among the total amount of 256 configurations. They provide an overview of the behaviour of the strategies proposed in this paper in problems of different difficulty with a search space of 256 possible solutions. Fig. 9 compares the percentage of explored configurations by the different strategies with the percentage of optimal configurations to be found.

Note that, in problems with more than 6% of optimal configuration, the strategies did not exhibit significant differences. In more difficult problems (less than 3% of optimal configurations), the number of configurations to explore increased up to 20% with *Synergy+MLE*, *Combined synergy* and in particular with *Naive sensitivity analysis*. *Naive synergy* and *State-based*, instead, as reported above, exhibited a similar behavior in most of the cases. In problems with less than 2% of optimal configurations, however, they also required to explore a higher number (up to ∼15%) of configurations. Finally, the figure
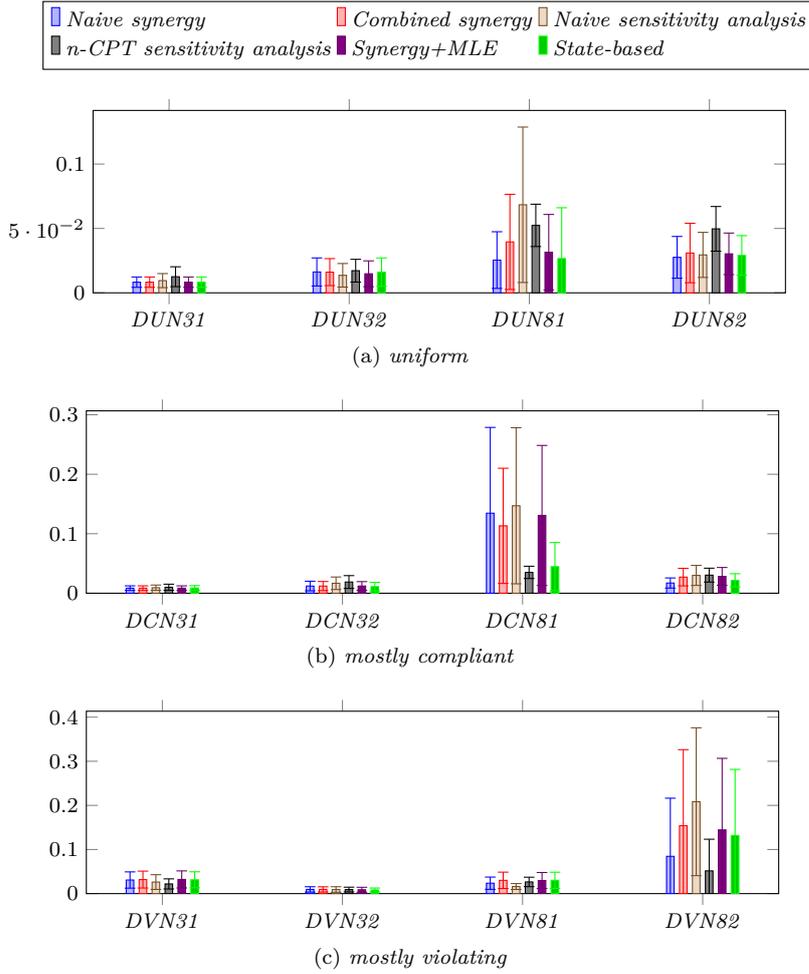
Fig. 8: Average percentage of explored configurations before finding an optimal one.

shows the robustness of *n-CPT sensitivity analysis*: despite performing slightly worse than other strategies in some experiments, *n-CPT sensitivity analysis* never required to explore more than 5% of all configurations, even in problems with about 1% of optimal configurations.

7 Discussion

The results reported in Sec. 6 show that our proposed strategies can be employed to effectively revise at runtime the sanctions of the enforced norms to quickly improve the performance of the system (in terms of achievement of
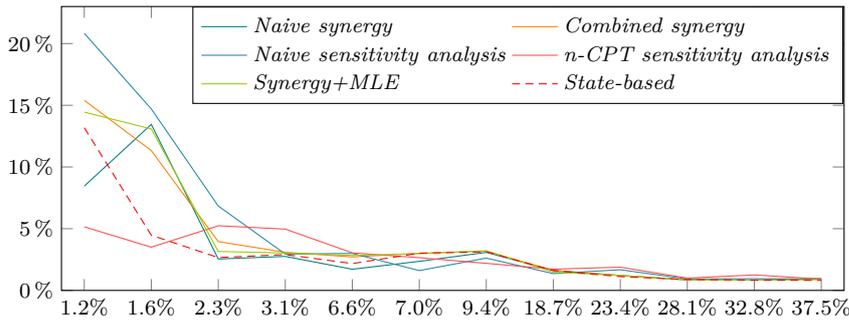
Fig. 9: Average percentage of explored configurations (y-axis) compared to the percentage of optimal configurations in the 12 different experiments (x-axis).

the system-level objectives). In particular, on 12 problems of different difficulty, our strategies reached optimal system's configurations after very few norm revisions. Starting with no initial knowledge about the effectiveness of the possible configurations, all the strategies explored on average less than 10% of all possible configurations before finding an optimal one. In the simplest experiment (*DVN32*), all strategies required to explore on average less than 1% of all possible configurations. In the same experiment, an uninformed strategy that does not consider runtime information and *randomly* tries a new configuration when the current one is not optimal would explore, on average, 62.5% of the configurations. In the most difficult experiment (*DVN82*), while a *random* strategy would explore on average 98.8% of the configurations to find one of the 1.2% optimal ones, our best performing strategy *n-CPT sensitivity analysis* explored, on average, only 5% of all possible configurations.

Our experiments identified three best-performing strategies: *Naive synergy*, *State-based* and *n-CPT sensitivity analysis*. We discuss each of these strategies and interpret the results and the conditions for their applicability.

Naive synergy determines, for each of the enforced norms, what type of synergy exists between the norm and the system-level objectives. Based on the identified synergy, *Naive synergy* increases or decreases the sanction for violating the norm. This strategy suits well cases where the observed data from MAS execution clearly highlights that a norm is better when either obeyed or violated. In experiment *DUN81*, for instance, in both contexts **c2** and **c3** the speed limit norm is effective only when fully obeyed by all agents (i.e., system configurations where some agents violate the speed limit are not optimal). In such experiment, and also in similar experiments such as *DUN82* and *DCN82*, the results confirmed that *Naive synergy* outperforms the other strategies.

The *State-based* strategy extends *Combined synergy*. Just like the latter, it considers the synergy between norms and objectives. Unlike *Combined synergy*, it also considers the most likely explanation for the objectives being not achieved. Furthermore, *State-based* takes also into account the global state of the system (the average norm obedience and objectives achievement) and sug-

gests to revise only a certain type of norms at every iteration. This strategy is suitable for cases where many norms are enforced and where the obedience of agents to a norm is likely to affect also the obedience to other norms. In our experiments, *State-based* performed well in most of the cases, with the exception of the most difficult ones *DCN81* and *DVN82*, where, similarly to *Naive synergy* it required a higher number of revisions.

Note that, in experiments *DCN81* and *DVN82*, the optimal configurations are only 4 and 3, respectively, out of 256. To find the few optimal configurations quickly, it is necessary to have a strategy that precisely directs the norm revision. For this reason, synergy-based or category-based strategies, which only provide a direction for the revision (i.e., they simply suggest to either increase or decrease violations), were not the best in these experiments.

n-CPT sensitivity analysis, instead, provides a quantitative measure of how much change in the violations of each norm is required. This strategy is more precise, and, although it performed slightly worse than other strategies in a few cases, it showed a consistent convergence speed in all the experiments, including complex ones such as *DCN81* and *DVN82*. Thus, this strategy proved to be the most robust in terms of convergence speed. It is worth noting, however, that in cases where the desired achievement of the system-level objectives is not particularly restrictive and where many norms are enforced, *n-CPT sensitivity analysis* may be less adequate due to the higher computational effort it requires, especially if compared to simpler strategies like *Naive synergy*.

The worst-performing strategy, on average, is *Naive sensitivity analysis*. This strategy performed particularly bad (compared to the others) especially in the most difficult experiments, where, as explained above, very few configurations were needed to be found. This result, which may seem surprising since sensitivity-based strategies are generally more precise than the others, can be explained by the *naive* approach of the strategy in determining the amount of change in the violations of norms that is required to achieve the system-level objectives. In doing so, unlike *n-CPT sensitivity analysis*, this strategy considers the changes for only one norm at a time, assuming that the other parameters of the Bayesian Network (i.e., the amount of violations of other norms) would not change. After providing a suggestion, however, the strategy applies a sanctions revision to all norms together (i.e., it changes all the parameters of the network together), creating a discrepancy between the way the suggestions are provided and the implementation of such suggestions. This discrepancy appears evident in cases where the precision of the suggestions is essential to identify one of the few optimal solutions (e.g., *DCN81* and *DVN82*). Note, however, that all the proposed strategies are heuristics. Therefore, there is no guarantee that one strategy will always perform better or worse than the others. This is visible in the results: every strategy that we proposed, including *Naive sensitivity analysis*, performed better than the others in at least one experiment.

7.1 Limitations and Possible Extensions

In the following, we provide a discussion of some of the limitations and assumptions related to our framework and to the revision strategies that we proposed, outlining some possible future directions.

7.1.1 Preferences Changing over Time and Context

We considered agents with same preferences in all operating contexts. This simplification does not affect the generality of our approach. Our framework supports agents with different preferences in multiple operating contexts. In Sec. 4.2, we have shown how to use the estimation of the preferences of agents to determine an upper bound of the probability of violating a norm. In Sec. 5, we used such upper bound to guide the revision of the sanctions of the enforced norms in a certain operating context c. In order to use different preferences in varying operating contexts, it is possible to explicitly model the different contexts (as proposed, for example, in context-aware systems such as Ambient Intelligence systems [46]), and use an adequate upper bound in each of them.

This is made possible by the assumption that the preferences of agents (and therefore our estimation) do not change over time, i.e., we assumed that the behavior of the agent is consistent over time. We did not study the case of preferences changing over time. Preferences may change over time due to external factors inducing changes in the end-user's preferences, the introduction of new norms in the MAS, or changes in agents' own evaluation of states of affairs due to the acquisition of new experience [41, 58].

To support preferences that change over time, our framework needs to be adapted in a number of ways, briefly listed below. First, depending on the type of system, mechanisms for the dynamic elicitation of preferences should be employed and the estimation of the preferences should be dynamically replaced or updated (see, for example, mechanisms to learn and update dynamic preferences [9, 20]). Given the new preferences, the upper bound of the probability of violating a norm should be recomputed. System configurations that are ineffective when certain behaviours are exhibited by the agents, may be instead effective when different behaviors are exhibited, and vice-versa. When the preferences of the agents are changed, therefore, the knowledge acquired during the norm revision process about the effectiveness of the norms and about the relationship between norm violation and system-level objectives should be reconsidered and opportunely weighted. If the preferences of the agents change very quickly and repeatedly over time, the use of a static *Norm Bayesian Network* as the one described in Sec. 4 may be unfavourable and the use of different more dynamic learning techniques, e.g., Dynamic Bayesian Networks [44], may be necessary. Supporting partial and inaccurate preferences of agents, as briefly discussed in Sec. 7.1.2, could also help to cope with preferences changing over time.

7.1.2 Partial or Inaccurate Information

When looking for a new sanction set, we assumed not to have any knowledge about the norm violations that will be actually exhibited when a never-tried-before sanction set is used to enforce norms. To guide the norm revision, we used the upper bound of a norm violation, a "safe" estimation of the actual norm violation that will be exhibited by agents. To calculate such upper bound we assumed an accurate (i.e., perfect) estimation of the preferences of the agents concerning the aspects of the system we aim to regulate.

The advantage of having an accurate estimation of the preferences of the agents is that we can define an upper bound for the probability of violating a (well defined) norm that is not too coarse-grained (e.g., a trivial upper bound is obviously a probability of 1, but this provides little information). As shown in Sec. 6, such an estimation, combined with our revision strategies, allows us to efficiently revise ineffective norms.

In some MASs, however, it is not possible to ensure a correct estimation of the agents' preferences [27]. Extending our work to support partial and/or inaccurate information about the agents' preferences requires an in-depth investigation. Based on the amount and type of information available, the accuracy and usefulness of the upper bound could significantly change. For partial information (e.g., we know that an agent type prefers a state of affairs over another, but we do not have information about all possible comparisons of alternative states of affairs), it is still possible to estimate a possibly more coarse-grained upper bound. For example, a trivial estimation could be obtained by assuming that agents always prefer to violate the norms related to aspects for which we do not have information. Less trivial estimations could be obtained for example by approximating the complete preferences by expressing the uncertain information as a belief function and leveraging the rationality principles of the preferences [18]. The estimated upper bound could be then refined over time by monitoring the behavior (i.e., the number of violations) of the agents. In case of inaccurate information (e.g., some of the available information about the preferences of agents is wrong, or the information available is only obtained from statistical data about the behavior of typical agents, or by learning the preferences from observed agents' choices [27]), the estimation of the probability of violating a norm should be treated more as a prediction, rather than an upper bound. In this case, techniques such as Bayesian Optimization [47], which attempts to find the minimum value of an unknown function, could be used for selecting new sanction sets and to refine over time the current estimation.

Nevertheless, a correct estimation of the preferences of the agents, as used in this paper, does not imply perfect revision strategies. This is because the trend of the upper bound may be different from the trend of the actual norm violation, which is unknown a priori. The consequence of this can be illustrated on the example of Fig. 10, which reports a comparison between an upper bound (red dashed line) of the probability of violating a norm N, and N's exhibited violation (blue solid line), w.r.t. the sanction associated to N.
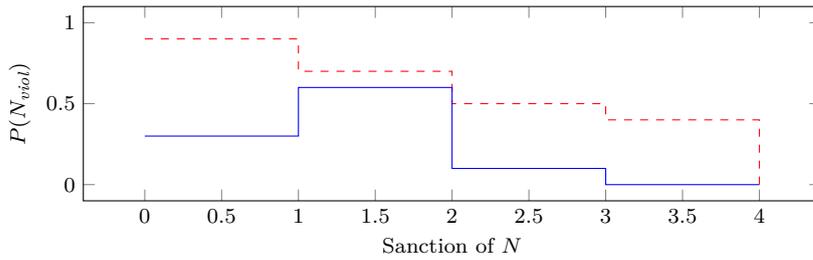
Fig. 10: Comparison between the upper bound (red dashed line) of the probability of violating a norm N, and N's exhibited violation (blue solid line), w.r.t. the sanction associated to N.

Suppose the current sanction for a norm N is 0, with an exhibited norm violation $P(N_{viol}) = 0.3$, and the employed revision algorithm (e.g., *Naive synergy*) suggests to *reduce* violations of N. Here, the only possible choice for Algorithm 1, which relies on the estimation of the upper bound of violating a norm, is to select sanction 4 as new sanction, since for all other sanctions the upper bound is higher than the currently exhibited norm violation. Although sanction 2 would also satisfy the suggestion, this will remain unknown until such sanction is tried. If the optimal value of $P(N_{viol})$ for the achievement of the system-level objectives is, for instance, around 0.1, our supervision framework will need to perform additional revision steps to select sanction 2.

7.1.3 Complexity of Preferences Representation

In this paper, we introduced several types of preferences of rational agents as lists of tuples ordered according to different rational criteria. In our discussion and experiments, we considered complete preferences, i.e., we explicitly represented all possible alternative states of affairs. Such representation, however, grows exponentially with the number of norms and budgets. In real world scenarios, doing so may be possible only in restricted domains where the number of norms and the possible budgets of the agents is limited. In the general case, however, representing the complete preferences of agents may be infeasible. In this work we attempted to lay down well founded principles for understanding the interplay between norms and the preferences of rational agents. For this reason, we provided a formal definition of different types of rational agents and we studied the properties of their preferences in relation with the chances to violate the enforced norms. We consider this as a necessary starting point for approaches to the runtime supervision of normative multiagent systems involving rational agents. In Sec. 7.1.2, we outlined some guidelines for our framework to support also partial (and inaccurate) preferences, which is one obvious way to reduce the complexity of explicitly representing the complete preferences. We leave this as future work, together with the integration of automated preferences elicitation techniques within our framework.

7.1.4 Norms Importance

Our strategies do not make any distinction between norms: revisions are applied to all the norms. This approach can be extended to support a selective revision that takes into account of the importance of a certain norm for the achievement of the objectives. Consider the derivative in Eq. 7, which describes the impact of changes in $P(N_{viol})$ on $P(\mathbf{O}_{true})$ in a context $\mathbf{c}$. High values of such derivatives imply that changes in the violations of norm N have bigger impact on $P(\mathbf{O}_{true})$. We call such derivative for a norm N the *importance* [54] *of norm N* in context $\mathbf{c}$. By computing the importance of all norms, we obtain an ordering between norms w.r.t. the system-level objectives. The strategies of Sec. 5.1 could be then applied to the k most important norms. Although there is no guarantee that this approach will be more effective, it applies to cases in which revising norms comes at a cost, and therefore minimizing the number of revisions is important.

In addition to the importance of a norm, the observed data from MAS execution allows to analyze the relationship between pairs of norms and to detect weather some of the following properties hold.

Additive synergy between two norms. This property, based on the concept of additive synergies in qualitative probabilistic networks [55], describes a situation where it is more likely to achieve the objectives when two norms are either both obeyed or both violated. Formally, two norms $N1$ and $N2$ exhibit an additive synergy when $P(\mathbf{O}_{true}|N1_{ob}N2_{ob}) + P(\mathbf{O}_{true}|N1_{viol}N2_{viol}) \geqslant P(\mathbf{O}_{true}|N1_{ob}N2_{viol}) + P(\mathbf{O}_{true}|N1_{viol}N2_{ob})$. The norms that exhibit an additive synergy with some of the k most important ones, could also be considered among the norms to be revised.

Product synergy between two norms. This property, based on the concept of product synergies in qualitative probabilistic networks [56], expresses how the value of one norm (e.g., $N1$ obeyed) influences the probability of the values of another norm (e.g., $N2$ obeyed), upon knowing the value for a common child (e.g., $\mathbf{O}$ true). For instance a negative product synergy says that observing $N1$ obeyed makes less likely to observe $N2$ being obeyed. Formally, two norms $N1$ and $N2$ exhibit a negative product synergy when $P(\mathbf{O}_{true}|N1_{ob}N2_{ob}) \cdot P(\mathbf{O}_{true}|N1_{viol}N2_{viol}) \geqslant (\leqslant)P(\mathbf{O}_{true}|N1_{ob}N2_{viol}) \cdot P(\mathbf{O}_{true}|N1_{viol}N2_{ob})$. This property can be used to choose between two norms to revise: it is enough to revise one of them to obtain an effect on the other.

7.1.5 Conflicting Norms

In this paper we assumed that the norms that are enforced are not conflicting, i.e., obeying a norm does not prevent *a priori* agents to obey other norms. This work focuses on regulative norms: norms enforced by an institution in order to regulate the behaviour of the agents so to achieve desired system-level properties. In this context, we believe that an institution should not

enforce conflicting norms, and we rely on normative conflict resolution mechanisms [51]. Despite this, our framework currently supports conflicting norms as long as the agents are aware of such conflicts, i.e., as long as the preferences of agents already take into account the conflicts. If two norms N_1 and N_2 are conflicting, obeying N_1 prevents the agents to obey N_2 and vice-versa. The preference of an agent that is aware of the conflict, determines whether the agent prefers to obey N_1 and pay a sanction for N_2, or vice-versa. This information is sufficient in our framework to estimate the upper bound for the violation of the norms and revise the sanctions of the norms when needed. Additionally, the information of the conflict could also be explicitly used to improve the performance of our revision strategies, similarly to the use of the product synergies described in Sec. 7.1.4: if obeying a norm agents cannot obey another norm, then it is sufficient to revise one sanction to obtain an effect also on the violation of the other norm.

7.1.6 Neighborhood Expansion

When a norm revision is triggered, our supervision mechanism searches for a new sanction set that satisfies the suggestions provided by one of the heuristic strategies. The neighborhood of a configuration, in the current hill climbing implementation of the supervisor, is composed by exactly one sanction set (configuration): the one that best satisfies the suggestions. An immediate extension of this approach is to expand the neighborhood definition, by including not only the best satisfying configuration, but also sub-optimal ones: those configurations that "almost" satisfy the suggestions provided. This extension is easily supported by our supervisor, and it better fits the typical usage of the hill climbing optimization technique. By expanding the neighborhood, the number of revision steps required by the supervision mechanism to find an optimal configuration could possibly further decrease. The challenge in expanding the neighborhood is in appropriately defining almost-satisfying suggestions. Different distance metrics and criteria could be considered in order to do so. Adopting a neighborhood composed only by the best satisfying configuration allowed us, however, to analyze the quality of the suggestions provided by our algorithms without further overloading the experimentation with additional parameters. Experiments with different neighborhood definitions will be carried on in future work, considering also a bigger case study.

8 Conclusions

In a MAS, the complexity and unpredictability of the agent interactions and of the environment must be taken into account to maximize the achievement of the system-level objectives. When engineering such systems, the available knowledge of these dynamics is only partial and incomplete. As a consequence, MASs need to be supervised and regulated at runtime.

In this paper, we proposed a supervision mechanism that relies on *norms with sanction* to influence agent behavior and regulate a MAS [12]. We considered MASs where agents are rational, i.e., they always choose to achieve their most preferred state of affairs. We characterized rational agents through their preferences and we made use of an estimation of the agents' preferences to guide the supervision of the MAS. Our mechanism automatically revises the sanctions that are employed to enforce the norms. To do so, it first interprets—through a Bayesian Network—observed data from MAS execution in terms of how well certain norms contribute to the achievement of the system-level objectives in different operating contexts. Then, it suggests how to revise the sanctions based on the knowledge learned at runtime and on the agents' preferences. We proposed six heuristics for the suggestion of sanction revisions.

An evaluation of the strategies through a traffic regulation simulation shows that our heuristics quickly identify optimal norm sets. We performed 12 different experiments on a ring-road traffic simulation, differing for the difficulty of the problem: the number of optimal norm sets to be found among all the possible ones ranged from 1.2% to 37.5%. All the proposed strategies explored a small number of norm sets before finding an optimal one. In particular, the strategy *n-CPT sensitivity analysis*, based on the sensitivity analysis technique from probabilistic reasoning [15], on average never required to explore more than 5% of all possible norm sets in order to find one of the optimal ones.

This work paves the way for numerous future directions, some of which are sketched in Sec. 7.1. An in-depth evaluation of the scalability and computational complexity of the presented approach is necessary to assess its suitability for MASs with many norms and sanctions. Our simple language for representing norms and agents' preferences can be extended to consider complex norm types beyond atomic propositions. Our agent population was defined according to specific types. Future work should study the effect of agents that deviate from the prototypical agent types. Finally, we are planning to extend our strategies to support, in addition to the revision of the sanctions, also the revision of the norm proposition, and to synthesize new norms.

Acknowledgement

References

1. Alechina, N., Bulling, N., Dastani, M., Logan, B.: Practical run-time norm enforcement with bounded lookahead. In: Proc. of AAMAS, pp. 443–451 (2015). URL http://dl.acm.org/citation.cfm?id=2772937
2. Alechina, N., Dastani, M., Logan, B.: Reasoning about normative update. In: Proc. of IJCAI, pp. 20–26 (2013). URL http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6884

3. Alechina, N., Dastani, M., Logan, B.: Norm approximation for imperfect monitors. In: Proc. of AAMAS, pp. 117–124 (2014)

4. Aucher, G., Grossi, D., Herzig, A., Lorini, E.: Dynamic context logic. In: Proc. of LORI, pp. 15–26 (2009)

5. Baarslag, T., Kaisers, M., Gerding, E.H., Jonker, C.M., Gratch, J.: When will negotiation agents be able to represent us? the challenges and opportunities for autonomous negotiators. In: C. Sierra (ed.) Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017, pp. 4684–4690. ijcai.org (2017). DOI 10.24963/ijcai.2017/653. URL `https://doi.org/10.24963/ijcai.2017/653`

6. Baier, J.A., McIlraith, S.A.: On domain-independent heuristics for planning with qualitative preferences. In: Logical Formalizations of Commonsense Reasoning, Papers from the 2007 AAAI Spring Symposium, Technical Report SS-07-05, Stanford, California, USA, March 26-28, 2007, pp. 7–12 (2007). URL `http://www.aaai.org/Library/Symposia/Spring/2007/ss07-05-003.php`

7. Balke, T., da Costa Pereira, C., Dignum, F., Lorini, E., Rotolo, A., Vasconcelos, W.W., Villata, S.: Norms in MAS: definitions and related concepts. In: G. Andrighetto, G. Governatori, P. Noriega, L.W.N. van der Torre (eds.) Normative Multi-Agent Systems, *Dagstuhl Follow-Ups*, vol. 4, pp. 1–31. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2013). DOI 10.4230/DFU.Vol4.12111.1. URL `https://doi.org/10.4230/DFU.Vol4.12111.1`

8. Benabbou, N., Perny, P.: Adaptive elicitation of preferences under uncertainty in sequential decision making problems. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017, pp. 4566–4572 (2017). DOI 10.24963/ijcai.2017/637. URL `https://doi.org/10.24963/ijcai.2017/637`

9. van Benthem, J., Liu, F.: Dynamic logic of preference upgrade. Journal of Applied Non-Classical Logics **17**(2), 157–182 (2007). DOI 10.3166/jancl.17.157-182. URL `https://doi.org/10.3166/jancl.17.157-182`

10. Bienvenu, M., Fritz, C., McIlraith, S.A.: Planning with qualitative temporal preferences. In: Proceedings, Tenth International Conference on Principles of Knowledge Representation and Reasoning, Lake District of the United Kingdom, June 2-5, 2006, pp. 134–144 (2006). URL `http://www.aaai.org/Library/KR/2006/kr06-016.php`

11. Boutilier, C., Patrascu, R., Poupart, P., Schuurmans, D.: Constraint-based optimization and utility elicitation using the minimax decision criterion. Artif. Intell. **170**(8-9), 686–713 (2006). DOI 10.1016/j.artint.2006.02.003. URL `https://doi.org/10.1016/j.artint.2006.02.003`

12. Bulling, N., Dastani, M.: Norm-based mechanism design. Artificial Intelligence **239**, 97–142 (2016). DOI 10.1016/j.artint.2016.07.001

13. Cardoso, H.L., Oliveira, E.C.: Adaptive deterrence sanctions in a normative framework. In: Proceedings of the 2009 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2009, Milan, Italy, 15-18 September 2009, pp. 36–43 (2009). DOI 10.1109/WI-IAT.2009.123. URL `https://doi.org/10.1109/WI-IAT.2009.123`

14. Castelfranchi, C., Dignum, F., Jonker, C.M., Treur, J.: Deliberative normative agents: Principles and architecture. In: Intelligent Agents VI, Agent Theories, Architectures, and Languages (ATAL), 6th International Workshop, ATAL '99, Orlando, Florida, USA, July 15-17, 1999, Proceedings, pp. 364–378 (1999). DOI 10.1007/10719619_27. URL `https://doi.org/10.1007/10719619_27`

15. Chan, H., Darwiche, A.: Sensitivity analysis in bayesian networks: From single to multiple parameters. In: UAI '04, Proceedings of the 20th Conference in Uncertainty in Artificial Intelligence, Banff, Canada, July 7-11, 2004, pp. 67–75 (2004)

16. Chen, L., Pu, P.: Survey of preference elicitation methods. Tech. rep. (2004)

17. Chopra, A.K., Dalpiaz, F., Giorgini, P., Mylopoulos, J.: Reasoning about agents and protocols via goals and commitments. In: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems, pp. 457–464 (2010)

18. Coletti, G., Petturiti, D., Vantaggi, B.: Rationality principles for preferences on belief functions. Kybernetika **51**(3), 486–507 (2015). DOI 10.14736/kyb-2015-3-0486. URL `https://doi.org/10.14736/kyb-2015-3-0486`

19. Cranefield, S., Meneguzzi, F., Oren, N., Savarimuthu, B.T.R.: A bayesian approach to norm identification. In: Proc. of ECAI, pp. 622–629 (2016). DOI 10.3233/978-1-61499-672-9-622

20. Crawford, E., Veloso, M.M.: Learning dynamic preferences in multi-agent meeting scheduling. In: A. Skowron, J.A. Barthès, L.C. Jain, R. Sun, P. Morizet-Mahoudeaux, J. Liu, N. Zhong (eds.) Proceedings of the 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Compiegne, France, September 19-22, 2005, pp. 487–490. IEEE Computer Society (2005). DOI 10.1109/IAT.2005.94. URL https://doi.org/10.1109/IAT.2005.94

21. Dasgupta, A., Ghose, A.K.: BDI agents with objectives and preferences. In: Declarative Agent Languages and Technologies VIII - 8th International Workshop, DALT 2010, Toronto, Canada, May 10, 2010, Revised, Selected and Invited Papers, pp. 22–39 (2010). DOI 10.1007/978-3-642-20715-0_2. URL https://doi.org/10.1007/978-3-642-20715-0_2

22. Dastani, M., Dignum, F., Meyer, J.C.: Autonomy and agent deliberation. In: M. Nickles, M. Rovatsos, G. Weiß (eds.) Agents and Computational Autonomy - Potential, Risks, and Solutions - Postproceedings of the 1st International Workshop on Computational Autonomy - Potential, Risks, Solutions (AUTONOMY 2003), held at the 2nd International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2003), July 14, 2003, Melbourne, Australia, *Lecture Notes in Computer Science*, vol. 2969, pp. 114–127. Springer (2003). DOI 10.1007/978-3-540-25928-2_10. URL https://doi.org/10.1007/978-3-540-25928-2_10

23. Dastani, M., Grossi, D., Meyer, J.C., Tinnemeier, N.A.M.: Normative multi-agent programs and their logics. In: Normative Multi-Agent Systems (2009). URL http://drops.dagstuhl.de/opus/volltexte/2009/1907/

24. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime Norm Revision using Bayesian Networks. In: Proceedings of the 21st International Conference on Principles and Practice of Multi-Agent Systems (PRIMA2018) (2018)

25. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime revision of norms and sanctions based on agent preferences. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '19, Montreal, QC, Canada, May 13-17, 2019, pp. 1609–1617 (2019). URL http://dl.acm.org/citation.cfm?id=3331881

26. Dell'Anna, D., Dastani, M., Dalpiaz, F.: "Runtime Revision of Sanctions in Normative Multi-Agent Systems" - Supplementary Material - SASS (2020). DOI 10.5281/zenodo.3712045. URL http://doi.org/10.5281/zenodo.3712045

27. Evans, O., Stuhlmüller, A., Goodman, N.D.: Learning the preferences of ignorant, inconsistent agents. In: D. Schuurmans, M.P. Wellman (eds.) Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA, pp. 323–329. AAAI Press (2016). URL http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12476

28. Fishburn, P.C.: Axioms for lexicographic preferences. The Review of Economic Studies **42**(3), 415–419 (1975)

29. Governatori, G., Rotolo, A.: Changing legal systems: legal abrogations and annulments in defeasible logic. Logic Journal of the IGPL **18**(1), 157–194 (2010). DOI 10.1093/jigpal/jzp075

30. Hausman, D.M.: Preference, value, choice, and welfare. Cambridge University Press (2011)

31. Jiang, J., Aldewereld, H., Dignum, V., Tan, Y.: Norm contextualization. In: Coordination, Organizations, Institutions, and Norms in Agent Systems VIII - 14th International Workshop, COIN 2012, Held Co-located with AAMAS 2012, Valencia, Spain, June 5, 2012, Revised Selected Papers, pp. 141–157 (2012). DOI 10.1007/978-3-642-37756-3_9. URL https://doi.org/10.1007/978-3-642-37756-3_9

32. Jiang, J., Thangarajah, J., Aldewereld, H., Dignum, V.: Reasoning with agent preferences in normative multi-agent systems. In: International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14, Paris, France, May 5-9, 2014, pp. 1373–1374 (2014). URL http://dl.acm.org/citation.cfm?id=2617479

33. Knobbout, M., Dastani, M.: Reasoning under compliance assumptions in normative multiagent systems. In: Proc. of AAMAS, pp. 331–340 (2012)

34. Knobbout, M., Dastani, M., Meyer, J.C.: Reasoning about dynamic normative systems. In: Proc. of JELIA, pp. 628–636 (2014). DOI 10.1007/978-3-319-11558-0_46

35. Knobbout, M., Dastani, M., Meyer, J.C.: A dynamic logic of norm change. In: Proc. of ECAI, pp. 886–894 (2016). DOI 10.3233/978-1-61499-672-9-886. URL https://doi.org/10.3233/978-1-61499-672-9-886

36. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.: Recent development and applications of SUMO - Simulation of Urban MObility. International Journal On Advances in Systems and Measurements **5**(3&4), 128–138 (2012)

37. Kwisthout, J.: Most probable explanations in bayesian networks: Complexity and tractability. Int. J. Approx. Reasoning **52**(9), 1452–1469 (2011). DOI 10.1016/j.ijar.2011.08.003. URL https://doi.org/10.1016/j.ijar.2011.08.003

38. Mas-Colell, A., Whinston, M.D., Green, J.R., et al.: Microeconomic theory, vol. 1. Oxford university press New York (1995)

39. Miralles, J.C., López-Sánchez, M., Salamó, M., Avila, P., Rodríguez-Aguilar, J.A.: Robust regulation adaptation in multi-agent systems. TAAS **8**(3), 13:1–13:27 (2013). DOI 10.1145/2517328. URL https://doi.org/10.1145/2517328

40. Mostafa, S.A., Ahmad, M.S., Mustapha, A.: Adjustable autonomy: a systematic literature review. Artif. Intell. Rev. **51**(2), 149–186 (2019). DOI 10.1007/s10462-017-9560-8. URL https://doi.org/10.1007/s10462-017-9560-8

41. Natarajan, S., Tadepalli, P.: Dynamic preferences in multi-criteria reinforcement learning. In: L.D. Raedt, S. Wrobel (eds.) Machine Learning, Proceedings of the Twenty-Second International Conference (ICML 2005), Bonn, Germany, August 7-11, 2005, *ACM International Conference Proceeding Series*, vol. 119, pp. 601–608. ACM (2005). DOI 10.1145/1102351.1102427. URL https://doi.org/10.1145/1102351.1102427

42. Padgham, L., Singh, D.: Situational preferences for BDI plans. In: International conference on Autonomous Agents and Multi-Agent Systems, AAMAS '13, Saint Paul, MN, USA, May 6-10, 2013, pp. 1013–1020 (2013). URL http://dl.acm.org/citation.cfm?id=2485080

43. van Riemsdijk, M.B., Hindriks, K.V., Jonker, C.M.: Programming organization-aware agents. In: Engineering Societies in the Agents World X, 10th International Workshop, ESAW 2009, Utrecht, The Netherlands, November 18-20, 2009. Proceedings, pp. 98–112 (2009). DOI 10.1007/978-3-642-10203-5_9. URL https://doi.org/10.1007/978-3-642-10203-5_9

44. Russell, S.J., Norvig, P.: Artificial Intelligence - A Modern Approach (3. internat. ed.). Pearson Education (2010)

45. Schmid, S., Gerostathopoulos, I., Prehofer, C., Bures, T.: Self-adaptation based on big data analytics: A model problem and tool. In: Proc. of SEAMS, pp. 102–108 (2017). DOI 10.1109/SEAMS.2017.20

46. Serral, E., Sernani, P., Dragoni, A.F., Dalpiaz, F.: Contextual requirements prioritization and its application to smart homes. In: A. Braun, R. Wichert, A. Maña (eds.) Ambient Intelligence - 13th European Conference, AmI 2017, Malaga, Spain, April 26-28, 2017, Proceedings, *Lecture Notes in Computer Science*, vol. 10217, pp. 94–109 (2017). DOI 10.1007/978-3-319-56997-0_7. URL https://doi.org/10.1007/978-3-319-56997-0_7

47. Snoek, J., Larochelle, H., Adams, R.P.: Practical bayesian optimization of machine learning algorithms. In: Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States., pp. 2960–2968 (2012). URL http://papers.nips.cc/paper/4522-practical-bayesian-optimization-of-machine-learning-algorithms

48. Sugiyama, Y., Fukui, M., Kikuchi, M., Hasebe, K., Nakayama, A., Nishinari, K., Tadaki, S.i., Yukawa, S.: Traffic jams without bottlenecksâĂŤexperimental evidence for the physical mechanism of the formation of a jam. New journal of physics **10**(3), 033001 (2008)

49. Testerink, B., Dastani, M., Bulling, N.: Distributed controllers for norm enforcement. In: Proc. of ECAI, pp. 751–759 (2016). DOI 10.3233/978-1-61499-672-9-751

50. Tumer, K., Welch, Z.T., Agogino, A.: Aligning social welfare and agent preferences to alleviate traffic congestion. In: Proceedings of the 7th international joint conference

on Autonomous agents and multiagent systems-Volume 2, pp. 655–662. International
Foundation for Autonomous Agents and Multiagent Systems (2008)

51. Vasconcelos, W.W., Kollingbaum, M.J., Norman, T.J.: Normative conflict resolution in multi-agent systems. Autonomous Agents and Multi-Agent Systems **19**(2), 124–152 (2009). DOI 10.1007/s10458-008-9070-9. URL https://doi.org/10.1007/s10458-008-9070-9

52. Vázquez-Salceda, J., Aldewereld, H., Grossi, D., Dignum, F.: From human regulations to regulated software agents' behavior. Artif. Intell. Law **16**(1), 73–87 (2008). DOI 10.1007/s10506-007-9057-x. URL https://doi.org/10.1007/s10506-007-9057-x

53. Visser, S., Thangarajah, J., Harland, J., Dignum, F.: Preference-based reasoning in BDI agent systems. Autonomous Agents and Multi-Agent Systems **30**(2), 291–330 (2016). DOI 10.1007/s10458-015-9288-2. URL https://doi.org/10.1007/s10458-015-9288-2

54. Wang, H., Rish, I., Ma, S.: Using sensitivity analysis for selective parameter update in bayesian network learning. Association for the Advancement of Artificial Intelligence (AAAI) (2002)

55. Wellman, M.P.: Fundamental concepts of qualitative probabilistic networks. Artificial Intelligence **44**(3), 257–303 (1990). DOI 10.1016/0004-3702(90)90026-V

56. Wellman, M.P., Henrion, M.: Explaining 'explaining away'. IEEE Trans. Pattern Anal. Mach. Intell. **15**(3), 287–292 (1993). DOI 10.1109/34.204911. URL https://doi.org/10.1109/34.204911

57. Wooldridge, M.J.: An Introduction to MultiAgent Systems (2. ed.). Wiley (2009)

58. Ziebart, B.D., Dey, A.K., Bagnell, J.A.: Fast planning for dynamic preferences. In: J. Rintanen, B. Nebel, J.C. Beck, E.A. Hansen (eds.) Proceedings of the Eighteenth International Conference on Automated Planning and Scheduling, ICAPS 2008, Sydney, Australia, September 14-18, 2008, pp. 412–419. AAAI (2008). URL http://www.aaai.org/Library/ICAPS/2008/icaps08-051.php

Appendix A   Properties of Rational Agents' Preferences

We report here a formal definition of the properties of the rational agents' preferences described in Sec. 3.3 and Sec. 4.2.

Proposition 1. *A basic preference $Pref(a) = (A, \geq)$ for an agent $a \in Ag$ is*

- *transitive: $\forall x, y, z \in A$ if $x \geq y$ and $y \geq z$ then $x \geq z$; and*
- *complete: $\forall x, y \in A$ either $x \geq y$ or $y \geq x$ or $x \sim y$.*

Proof. Consider a list $AL = (L_1, ..., L_n)$, a set $\mathcal{B} \subset \mathbb{N}$, a set $BL \subseteq \mathcal{B}^n$ and an ordered set $\mathcal{N}$ of n norms. Let $Pref(a) = (A, \geq)$ be a basic preference and x, y, z be alternatives in A.

(*Transitivity*) Assume that $x \geq y$ and $y \geq z$. We prove the transitivity for two cases: either the preference adheres to Def. 2a (case 1) or the preference adheres to Def. 2b (case 2). For both cases, we show that $x \geq z$.

Case 1: we have that $req\_bud(x) \leqslant req\_bud(y)$ and $req\_bud(y) \leqslant req\_bud(z)$. By transitivity of $\leqslant$, we have $req\_bud(x) \leqslant req\_bud(z)$. Moreover, we have $\forall k, l \in A, \forall B, B' \in BL : k[B] > l[B] \Rightarrow k[B'] > l[B']$ for both $x \geq y$ and $y \geq z$, such that we have it also for $x \geq z$. Therefore, we conclude that $x \geq z$.

Case 2: we have either (i) $prop(x) = prop(y) = prop(z)$ or (ii) $prop(x) \neq prop(y)$ and either $prop(y) = prop(z)$ or $prop(y) \neq prop(z)$. In case (i), we have that $req\_bud(x) \leqslant req\_bud(y)$ and $req\_bud(y) \leqslant req\_bud(z)$, and, by transitivity of $\leqslant$, we have that the first condition of Def. 2b is also satisfied by x and z. In case (ii), instead, we have that $\forall B, B' \in BL : x[B] \geq y[B']$.

If $prop(y) = prop(z)$, we have that $prop(x) \neq prop(z)$ and $\forall B, B' \in BL :$ $x[B] \geq z[B']$, so the second condition of Def. 2b is satisfied also by x and z. If instead also $prop(y) \neq prop(z)$, then we have that $\forall B, B' \in BL : x[B] \geq y[B']$ and $\forall B, B' \in BL : y[B] \geq z[B']$. Take three arbitrary alternatives $x[B], y[B'], z[B'']$. We have that $x[B] \geq y[B']$ and $y[B'] \geq z[B'']$. Since the choice of B, B' and B'' was arbitrary, this holds for any possible budget list. Therefore there is no $B, B'' \in BL$ such that alternative $z[B'']$ is strictly preferred to an alternative $x[B]$, i.e., $\forall B, B'' \in BL : x[B] \geq z[B'']$. Again the second condition of Def. 2b is satisfied by x and z, such that $x \geq z$.

(*Completeness*) By definition of basic preference, every pair of alternatives $x, y \in A$ has to satisfy either Def. 2a or Def. 2b. Notice that, given x and y, if it is not the case that $x \geq y$, then we have that $y > x$, therefore for every pair of alternatives $x, y \in A$ either $x \geq y$ or $y > x$. $\qquad\square$

Proposition 2. *A preference $Pref(a) = (A, \geq)$ for an agent $a \in Ag$ is*

- *transitive: $\forall x, y, z \in A$ if $x \geq y$ and $y \geq z$ then $x \geq z$; and*
- *complete: $\forall x, y \in A$ either $x \geq y$ or $y \geq x$ or $x \sim y$.*

Proof. Consider a list $AL = (L_1, ..., L_n)$, a set $\mathcal{B} \subset \mathbb{N}$, a set $BL \subseteq \mathcal{B}^n$ and an ordered set $\mathcal{N}$ of n norms. Let $A = \{(\langle p_1, b_1 \rangle, ..., \langle p_n, b_n \rangle) \mid p_i \in L_i \ \& \ (b_1, \ldots, b_n) \in BL\}$ be the set of alternatives over which agents have preferences. Let $A_1, ..., A_k$ be k disjoint subsets of A as per Def. 4, and x, y, z be alternatives in A.

(*Transitivity*) Assume that $x \geq y$ and $y \geq z$. If both x, y and z belong to the same A_i for $1 \leq i \leq k$ then, by Prop.1, $x \geq z$. Otherwise, if $x \in A_i$, $y \in A_j$ and $z \in A_l$ with $i < j < l$, then, by Def. 4, given $i < l$, $\forall v \in A_i \forall w \in A_l : v \geq w$, therefore $x \geq z$.

(*Completeness*) By Prop.1, for every pair of alternatives $x, y \in A_i$ for $1 \leq i \leq k$, either $x \geq_i y$ or $y \geq_i x$. Furthermore, by definition of preference, for all $x \in A_j$ and $y \in A_i$, we have $x \geq y$, for $1 \leq j < i \leq k$. We have therefore that for every pair of alternatives $x, y \in A$ either $x \geq y$ or $y \geq x$. $\qquad\square$

Proposition 3. *Given an ordered set of norms $\mathcal{N} = \langle N_1, ..., N_n \rangle$, and a set of t agent types $\mathcal{T}$, each type corresponding to a consistent preference (as per Def. 5), increasing the sanction of a norm N_j in $\mathcal{N}$ without changing the sanctions of other norms, does not increase the upper bound of the probability $P(N_{viol})$, i.e., $UB(N_{viol}, \mathcal{N})$, for all N in $\mathcal{N}$.*

Proof. In this paper the agent's preferences are not affected by the preferences of other agents. Since the upper bound $UB(N_{viol}, \mathcal{N})$ of the probability of violating a norm N in the context of a norm set $\mathcal{N}$ is determined by the number of agents with reason to violate N, as per Sec.3.2, if Prop.3 holds for one agent type, then Prop.3 must hold also for all agent types. In the following we consider, therefore, one agent type T. Furthermore we assume $\mathcal{N}$ composed by at least two different norms (if only one norm is enforced, Prop.3 is trivially satisfied).

We prove Prop. 3 by contradiction.

Let M be the set of most preferred alternatives to act upon for agent type T in the context of $\mathcal{N} = \langle N_1, \ldots, N_j, \ldots, N_n \rangle$ (as per Def. 6). Suppose we increase the sanction of norm $N_j = (p_j, s_j)$, obtaining $N_j' = (p_j, s_j' > s_j)$. Let now M' be the set of most preferred alternatives to act upon for agent type T in the context of $\mathcal{N}' = \langle N_1, \ldots, N_j', \ldots, N_n \rangle$.

Suppose, by contradiction, that $UB(N_{viol}, \mathcal{N}') > UB(N_{viol}, \mathcal{N})$ for $N = (p, s) \neq N_j$, with $p \in L_i$ and L_i in AL. This means that, in the context of $\mathcal{N}$, T has no reason to violate N, while in the context of $\mathcal{N}'$, T has reason to violate N (i.e., there exists no alternative $c \in M$ with $viol(c, N)$, while there exists an alternative $c' \in M'$ such that $viol(c', N)$).

In order to modify the set of most preferred alternatives M when increasing the sanction from s_j to s_j', it must be the case that there exists at least one alternative $c \in M$ s.t. $viol(c, N_j)$ and $s_j \leqslant b_j < s_j'$ (with b_j budget of the j-th pair in c). If it is not the case increasing s_j to s_j' does not affect T's most preferred alternatives and thus the proposition holds.

Consider the alternative from M with highest budget b_j in the j-th pair. Consider also an $i \neq j$. Let $\bar{a}wx[B] = (\langle p_1, b_1 \rangle, \ldots, \langle w, b_w \rangle, \ldots, \langle x, b_x \rangle, \ldots, \langle p_n, b_n \rangle)$ be such alternative, with $w \in L_j, x \in L_i$ and $s_j \leqslant b_w < s_j'$. Let $\bar{b}zy[B'] = (\langle p_1', b_1' \rangle, \ldots, \langle z, b_z \rangle, \ldots, \langle y, b_y \rangle, \ldots, \langle p_n', b_n' \rangle)$ be an alternative $c' \in M'$ such that $c' \notin M$ and $viol(c', N)$ with $z \in L_j, y \in L_i$ and $b_z \geqslant s_j'$. Notice that $\bar{a}wx[B]$ is compliant w.r.t. N and $\bar{b}zy[B']$ is not compliant w.r.t. N, hence $y \neq x$. Notice also that $\bar{a}wx[B] > \bar{b}zy[B']$. This is because by Def. 6 we have that, since $\bar{b}zy[B'] \notin M$, $\bar{b}zy[B'] \geq \bar{a}wx[B]$ iff $\bar{b}zy[B']$ violates a norm N_k but the budget is not enough to pay the sanction. Such N_k cannot be N_j, since $b_z \geqslant s_j'$, and if it's another N_k then $\bar{b}zy[B']$ cannot be also in M' because we only increased the sanction of norm N_j. Therefore $\bar{b}zy[B']$ must be strictly less preferred than $\bar{a}wx[B]$. Furthermore, let $\bar{c}$ be a fully compliant alternative.[10]

We first consider the case of $T = (A, \geq)$ basic preference as per Def. 2, which can adhere to either Def. 2a or Def. 2b, then we uplift the proof to the preference as per Def. 4.

Basic Preference.

(Case Def. 2a) Since $\bar{a}wx[B] > \bar{b}zy[B']$, it holds that, due to Def. 2a, $c[B] > \bar{b}zy[B']$ for all alternatives c in A. This means that also a fully compliant alternative $\bar{c}[B]$ is such that $\bar{c}[B] > \bar{b}zy[B']$. However, since $s_j' > b_w$, after revising s_j into s_j', there is at least one alternative in M' (i.e., $\bar{c}[B]$) that is strictly preferred to $\bar{b}zy[B']$, because already present in M, and that is compliant to N. Therefore, $\bar{b}zy[B']$ cannot be among the most preferred alternatives to act upon, i.e., $\bar{b}zy[B'] \notin M'$ (contradiction).

(Case Def. 2b) Since $\bar{a}wx[B] \in M > \bar{b}zy[B'] \in M'$, by Def. 2b it holds that $\bar{a}wx[B'] \geq \bar{a}wx[B] > \bar{b}zy[B'] \geq \bar{b}zy[B]$ for all $B' \in BL$, i.e., $\bar{a}wx > \bar{b}zy$ regardless of the required budget, including the maximum possible budget in

---

[10] We call $\bar{a}$ and $\bar{b}$ the list of propositional atoms that are different from w, x, y, z respectively in $\bar{a}wx$ and $\bar{b}zy$. Also, we use notation $c[B]$ to indicate an alternative c with list of budgets B, as described in Sec.3.2.

$\mathcal{B}$, $max(\mathcal{B})$. Therefore, since $\bar{a}wx$ is the alternative in M with highest budget in the j-th pair, we have that $s'_j > max(\mathcal{B})$, and no alternative that violates N'_j can be chosen in the context of $\mathcal{N}'$. $\bar{b}zy[B']$ is then compliant w.r.t. N'_j, hence $z \neq w$. By Def. 2, T contains at least another alternative $\bar{a}wy[B']$. If $\bar{a}wy[B'] \succeq \bar{a}wx[B']$, then, according to Def. 2b, $\bar{a}wy[B'] \succeq \bar{a}wx[B']$ for all $B' \in BL$. But if this is the case, we have that $s_j > max(\mathcal{B})$ (otherwise at least one alternative $\bar{a}wy$ with $b_w = max(\mathcal{B})$ is in M and, in contradiction with out hypothesis, T has reason to violate N since $viol(y, N)$). But if $s_j > max(\mathcal{B})$ we have $\bar{a}wx[B] \notin M$ (contradiction). If, instead, $\bar{a}wx[B'] > \bar{a}wy[B']$ then $\bar{a}wx > \bar{a}wy$ regardless of the budget. By consistency (Def 5), then, we also have $\bar{b}zx > \bar{b}zy$. We distinguish 2 cases: (a) $\bar{a}wx > \bar{b}zx$, this implies $\bar{a}wx > \bar{b}zx > \bar{b}zy$, which contradicts $\bar{b}zy[B'] \in M'$, since alternatives $\bar{b}zx$ (compliant w.r.t N'_j) are strictly preferred to $\bar{b}zy[B']$; (b) $\bar{b}zx > \bar{a}wx$, this implies that, since $\bar{a}wx[B] \in M$ and $\bar{b}zx$ is compliant w.r.t. both L_j and L_i, then for every other norm violated by $\bar{b}zx$, the sanction associated to such norm is bigger than $max(\mathcal{B})$ (otherwise $\bar{a}wx \notin M$ and at least one alternative $\bar{b}zx \in M$). But if this is the case, also $\bar{b}zy[B'] \notin M'$, since the only sanction that we change is s_j, and again we have a contradiction.

Preference.

In the case of preference $T = (A, \succeq)$, the k basic preferences composing A adhere to either Def. 2a or Def. 2b. The only case non considered above is when $\bar{a}wx[B] \in A_i, \bar{b}zy[B'] \in A_p$ and it does not exists an alternative $c \in A_p$ s.t. $\bar{a}wx[B] > c$, and it does not exists an alternative $c' \in A_q$ s.t. $c' > \bar{b}zy[B']$, for two basic preferences A_p, A_q composing A, with $1 \leqslant p < q \leqslant k$. Since all alternatives in A_i have required budget lower or equal than $max(\mathcal{B}_i)$, then, due to Def. 2, among A_p there is at least one fully compliant alternative with required budget $\leqslant max(\mathcal{B}_i)$. Therefore, even if sanction for N_j is increased to a value $s > max(\mathcal{B})$, M' does not contain any alternative from A_q that was not already in M, therefore $\bar{b}zy[B'] \notin M'$, and again we have a contradiction. $\quad\square$

Appendix B   Experiments Agent Types

We report here the full preferences of the four types of agents considered in our experimentation, described in Sec. 6.

- *BraveRich*:

$(\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 0\rangle) > (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 1\rangle) > (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 2\rangle) >$
$(\langle sp_8, 0\rangle, \langle dist_{0.5}, 0\rangle) > (\langle sp_8, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 1\rangle) > (\langle sp_8, 2\rangle, \langle dist_{0.5}, 2\rangle) >$
$(\langle sp_3, 0\rangle, \langle dist_{0.5}, 0\rangle) > (\langle sp_3, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_3, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 0\rangle) >$
$(\langle sp_3, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 1\rangle) > (\langle sp_3, 2\rangle, \langle dist_{0.5}, 2\rangle) >$

$$(\langle sp_{15},0\rangle,\langle dist_1,0\rangle) \succ (\langle sp_{15},0\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_{15},1\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_{15},0\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_1,1\rangle) \succ (\langle sp_{15},2\rangle,\langle dist_1,2\rangle) \succ$$
$$(\langle sp_8,0\rangle,\langle dist_1,0\rangle) \succ (\langle sp_8,0\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_8,1\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_8,1\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_8,0\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_8,2\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_8,1\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_8,2\rangle,\langle dist_1,1\rangle) \succ (\langle sp_8,2\rangle,\langle dist_1,2\rangle) \succ$$
$$(\langle sp_3,0\rangle,\langle dist_1,0\rangle) \succ (\langle sp_3,0\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_3,1\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_3,0\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_3,1\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_1,1\rangle) \succ (\langle sp_3,2\rangle,\langle dist_1,2\rangle) \succ$$
$$(\langle sp_{15},0\rangle,\langle dist_2,0\rangle) \succ (\langle sp_{15},0\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_{15},1\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_{15},0\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_2,1\rangle) \succ (\langle sp_{15},2\rangle,\langle dist_2,2\rangle) \succ$$
$$(\langle sp_8,0\rangle,\langle dist_2,0\rangle) \succ (\langle sp_8,0\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_8,1\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_8,1\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_8,0\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_8,2\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_8,1\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_8,2\rangle,\langle dist_2,1\rangle) \succ (\langle sp_8,2\rangle,\langle dist_2,2\rangle) \succ$$
$$(\langle sp_3,0\rangle,\langle dist_2,0\rangle) \succ (\langle sp_3,0\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_3,1\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_3,0\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_3,1\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_2,1\rangle) \succ (\langle sp_3,2\rangle,\langle dist_2,2\rangle)$$

– *BraveMiddleClass*:

$$(\langle sp_{15},0\rangle,\langle dist_{0.5},0\rangle) \succ (\langle sp_{15},0\rangle,\langle dist_{0.5},1\rangle) \succeq (\langle sp_{15},1\rangle,\langle dist_{0.5},0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_{0.5},1\rangle) \succ (\langle sp_8,0\rangle,\langle dist_{0.5},0\rangle) \succ (\langle sp_8,0\rangle,\langle dist_{0.5},1\rangle) \succeq$$
$$(\langle sp_8,1\rangle,\langle dist_{0.5},0\rangle) \succ (\langle sp_8,1\rangle,\langle dist_{0.5},1\rangle) \succ (\langle sp_3,0\rangle,\langle dist_{0.5},0\rangle) \succ$$
$$(\langle sp_3,0\rangle,\langle dist_{0.5},1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_{0.5},0\rangle) \succ (\langle sp_3,1\rangle,\langle dist_{0.5},1\rangle) \succ$$
$$(\langle sp_{15},0\rangle,\langle dist_1,0\rangle) \succ (\langle sp_{15},0\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_{15},1\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_1,1\rangle) \succ (\langle sp_8,0\rangle,\langle dist_1,0\rangle) \succ (\langle sp_8,0\rangle,\langle dist_1,1\rangle) \succeq$$
$$(\langle sp_8,1\rangle,\langle dist_1,0\rangle) \succ (\langle sp_8,1\rangle,\langle dist_1,1\rangle) \succ (\langle sp_3,0\rangle,\langle dist_1,0\rangle) \succ$$
$$(\langle sp_3,0\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_1,0\rangle) \succ (\langle sp_3,1\rangle,\langle dist_1,1\rangle) \succ$$
$$(\langle sp_{15},0\rangle,\langle dist_2,0\rangle) \succ (\langle sp_{15},0\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_{15},1\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_2,1\rangle) \succ (\langle sp_8,0\rangle,\langle dist_2,0\rangle) \succ (\langle sp_8,0\rangle,\langle dist_2,1\rangle) \succeq$$
$$(\langle sp_8,1\rangle,\langle dist_2,0\rangle) \succ (\langle sp_8,1\rangle,\langle dist_2,1\rangle) \succ (\langle sp_3,0\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_3,0\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_2,0\rangle) \succ (\langle sp_3,1\rangle,\langle dist_2,1\rangle) \succ$$
$$\text{\# from here ordered by budget}$$
$$(\langle sp_{15},0\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_{0.5},0\rangle) \succeq (\langle sp_8,0\rangle,\langle dist_{0.5},2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_{0.5},0\rangle) \succeq (\langle sp_3,0\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_{0.5},0\rangle) \succeq$$
$$(\langle sp_{15},0\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_1,0\rangle) \succeq (\langle sp_8,0\rangle,\langle dist_1,2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_1,0\rangle) \succeq (\langle sp_3,0\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_1,0\rangle) \succeq$$
$$(\langle sp_{15},0\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_2,0\rangle) \succeq (\langle sp_8,0\rangle,\langle dist_2,2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_2,0\rangle) \succeq (\langle sp_3,0\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_2,0\rangle) \succ$$
$$(\langle sp_{15},1\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_{0.5},1\rangle) \succeq (\langle sp_8,1\rangle,\langle dist_{0.5},2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_{0.5},1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_{0.5},1\rangle) \succeq$$
$$(\langle sp_{15},1\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_8,1\rangle,\langle dist_1,2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_1,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_1,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_1,1\rangle) \succeq$$
$$(\langle sp_{15},1\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_{15},2\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_8,1\rangle,\langle dist_2,2\rangle) \succeq$$
$$(\langle sp_8,2\rangle,\langle dist_2,1\rangle) \succeq (\langle sp_3,1\rangle,\langle dist_2,2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_2,1\rangle) \succ$$
$$(\langle sp_{15},2\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_8,2\rangle,\langle dist_{0.5},2\rangle) \succeq (\langle sp_3,2\rangle,\langle dist_{0.5},2\rangle) \succeq$$

$$(\langle sp_{15}, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 2\rangle) \succeq$$
$$(\langle sp_{15}, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 2\rangle)$$

– *BravePoor*:

$$(\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_{15}, 0\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_{15}, 0\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_2, 0\rangle) \succ$$
$$(\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_{15}, 0\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_{15}, 0\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_2, 0\rangle) \succ$$
$$(\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 0\rangle) \succ$$
$$(\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 1\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_1, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 1\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_2, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 1\rangle) \succ$$
$$(\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 2\rangle) \succeq$$
$$(\langle sp_{15}, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 2\rangle) \succeq$$
$$(\langle sp_{15}, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 2\rangle)$$

– *Cautious*:

$$(\langle sp_3, 0\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_3, 0\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_3, 0\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 0\rangle) \succ$$
$$(\langle sp_3, 0\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_2, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_3, 0\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_1, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_3, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 1\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 0\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 0\rangle) \succ$$
$$(\langle sp_3, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_2, 0\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 0\rangle) \succeq$$

$$(\langle sp_8, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_1, 0\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_3, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_8, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 0\rangle) \succeq$$
$$(\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 0\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 0\rangle) >$$
$$(\langle sp_3, 1\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_2, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_2, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_2, 1\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_1, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_1, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_1, 1\rangle) \succeq$$
$$(\langle sp_3, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_3, 2\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_8, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq$$
$$(\langle sp_8, 2\rangle, \langle dist_{0.5}, 1\rangle) \succeq (\langle sp_{15}, 1\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 1\rangle) >$$
$$(\langle sp_3, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_2, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_2, 2\rangle) \succeq$$
$$(\langle sp_3, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_1, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_1, 2\rangle) \succeq$$
$$(\langle sp_3, 2\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_8, 2\rangle, \langle dist_{0.5}, 2\rangle) \succeq (\langle sp_{15}, 2\rangle, \langle dist_{0.5}, 2\rangle)$$