A Bayesian approach to norm identification

Stephen Cranefield[1] and **Felipe Meneguzzi**[2] and **Nir Oren**[3] and **Bastin Tony Roy Savarimuthu**[4]

Abstract. When entering a system, an agent should be aware of the obligations and prohibitions (collectively *norms*) that affect it. Existing solutions to this *norm identification* problem make use of observations of either norm compliant, or norm violating, behaviour. Thus, they assume an extreme situation where norms are typically violated, or complied with. In this paper we propose a Bayesian approach to norm identification which operates by learning from both norm compliant and norm violating behaviour. We evaluate our approach's effectiveness empirically and compare its accuracy to existing approaches. By utilising both types of behaviour, we not only overcome a major limitation of such approaches, but also obtain improved performance over the state of the art, allowing norms to be learned with fewer observations.

1 Introduction

Norms, as instantiated through obligations, permissions and prohibitions, are a popular approach to declarative behaviour specification within multi-agent systems [25, 21, 7, 1]. Such norms describe the expected behaviour of agents, but can be violated in exceptional circumstances. A large body of work exists on how agents should behave in the presence of norms [10, 3, 14, 15, 13]. Recently, work has emerged addressing *norm identification*—how an agent can identify norms already present in an environment. This problem is important in open, dynamic multi-agent systems, where agents can enter and leave the system at any time, and no assumption regarding norm knowledge can be made. While it is often assumed that norms can be communicated to agents when they enter a system [20], factors such as limited bandwidth, implicit norms (in some systems), lack of a shared ontology, malicious behaviour and changing norms can invalidate this assumption, instead requiring that agents be able to identify norms dynamically.

Previous work on dynamic norms often focuses on the consequences of norm emergence to society [17], that is, evaluating what happens to a society when norms change. However, only recently have researchers started to investigate practical approaches to the problem posed to individual agents of *inferring* new norms as they emerge. Such work often makes a combination of assumptions regarding what available evidence can actually be used to identify new norms. The work of Savarimuthu et al. [18, 19], is a typical example of an existing approach to norm recognition, based on the detection of a *sanctioning signal*—an action responding to an agent's norm violation that may (possibly) be performed by a peer of the agent or an institutional authority. Crucially, it is assumed that these signals may be recognised as conveying some negative emotional or institutional force, even before details of the specific norms in the society have been inferred[5]. By learning the situations in which such sanctioning signals arise, agents are able to infer their triggering norms. However, while such an approach works well when sanctioning signals are common, it is more difficult to apply in systems where agents (largely) comply with norms.

To address this difficulty, Oren and Meneguzzi [16] introduced a plan recognition based mechanism for norm identification. In their approach, by observing the behaviour of agents, and identifying what states these agents avoid or always achieve, prohibitions and obligations can be identified. However, in its simplest form, this approach must assume fully norm-compliant behaviour. An adaptation suggested by Oren and Meneguzzi overcomes this limitation, but is too memory intensive to be practical in any reasonably sized domain.

There have been other works in the realm of norm identification [6, 2, 11]. The work done in the EMIL project by Campenni et al. [6] infers norms using observed behavioural patterns based on a threshold-based approach, where the observations could be from a range of sources: deontic commands, evaluative statements and assertions made by agents that are being observed. Based on aggregating this information, an agent could infer potential norms. For example if behaviour A is more prevalent than behaviour B in a given context, then A is considered to be a norm. This work assumes that an observer already knows how to interpret the normative statements, and hence has an implicit notion of a norm. However, in real life an observer new to a society may not have prior knowledge of what the norms might be and why an agent is being sanctioned. The work of Alrawagfeh et al. [2] aims to extract permission norms similarly to that of Campenni et al. [6], and prohibitions similarly to the work of Savarimuthu et al. [19, 18], thus suffering from the limitations of these two approaches. Additionally, this work does not infer obligations. The work of Mahmoud et al. [11], like the work of Savarimuthu et al. [19, 18], requires a sanctioning signal in order to function.

Existing work on norm identification therefore assumes that norms are almost always either complied with or violated [11, 19, 18, 16, 2], and is not appropriate in less extreme (and more realistic) cases. The core contribution of this paper is *an approach to norm identification that operates well in domains where both norm compliance and violation can regularly occur*. Thus we relax the strong assumptions of all existing work and develop an algorithm that can infer norms using a variety of possible sources of evidence. Our approach, described in Section 2, uses Bayesian inference to compute for each candidate norm the odds that it is an established norm, compared to the null hypothesis that there are no norms, given observations of other agents' behaviour. To act in a norm compliant way, an agent uses these odds to select which norms should be followed. We report on an empirical evaluation of our approach in Section 3, which

[1] University of Otago, New Zealand, email: stephen.cranefield@otago.ac.nz
[2] Pontifical Catholic University of Rio Grande do Sul, Brazil, email: felipe.meneguzzi@pucrs.br
[3] University of Aberdeen, UK, email: n.oren@abdn.ac.uk
[4] University of Otago, New Zealand, email: tony.savarimuthu@otago.ac.nz

[5] For example, in human society, observing someone shouting or gesturing angrily at another can be understood as a message of displeasure, even without overhearing their conversation. Similarly, if we observe a policeman issuing an infringement notice to another, we are aware that a violation has been detected, even if we do not know the details.

shows two key results: first, that norm-compliant behaviour is possible after relatively few observations of the actions of others; and second, that our approach outperforms existing approaches to norm identification. Finally, we contextualise our work and point to future research directions in Section 4.

An extended abstract of this paper has been published previously [8]. That did not provide details of the approach, and reported on some preliminary experimental results. In this paper we provide a full account of the models and algorithms used in the work, and discuss a new experimental evaluation.

2 The Model and Approach

The first question that we must address is how our normative system should be encoded. There is a long history of utilising transition systems to model agents within a multi-agent system. These transition systems model the state space as nodes in a graph, and actions are encoded as edges allowing transitions between nodes (states). Therefore, following Oren and Meneguzzi [16], we consider an abstract normative environment where norms govern motion through such a graph, and seek to identify legal and illegal paths within this graph. As in transition systems, nodes in the graph represent individual states, while edges represent transitions through the space due to agent actions. However, unlike the graphs used in transition systems, we abstract away from the interpretation function used to associate values with variables in each state, and instead consider only motion through the graph itself. Therefore, a path within such a graph represents the actions of an agent following a plan to transition from some initial state (its start node) to a goal state (its destination node). One such graph is shown in Figure 1.

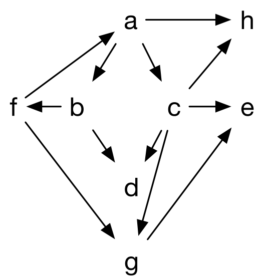


Figure 1: A sample graph

We conceptualise plans as sequences of nodes, and make no assumptions about the source of those plans: they may be generated dynamically given a goal and a set of possible actions, or they may come from a plan library, such as a BDI agent program. Our norm identification mechanism is based on the assumption that the observed agents' plan libraries (or available actions and planning mechanism) are known to the observing agent, at least at some level of abstraction. This would be the case if all agents share the same plan library, if their possible plans can be inferred from public knowledge about the problem domain, e.g. public transport routes and timetables, or (as assumed by Oren and Meneguzzi [16]) if the observing agent has a plan recognition mechanism. Alternatively, in the absence of any other information, an agent may have no other option than to simply assume that other agents are like itself, in order to gain some traction on the norm identification problem. It is important to note that for the purposes of norm identification, agents only need to infer the plans of other agents that govern their *publicly observable* behaviour.

Identifying norms then involves observing the movements of others through the graph to identify their goals and the paths that correspond to plan executions. Given a set of norm hypotheses, we use observations as evidence to compute the odds of each hypothesis being a norm compared to the null hypothesis that there are no norms. In addition, we assume that when violations of norms occur, other agents may choose to sanction the offending agents, that these sanctioning actions can be recognised through sanctioning signals, and that these actions are spontaneously performed by agents rather than being generated by the plans they follow. These signals are another source of evidence that can be used to update the odds of the norm hypotheses.

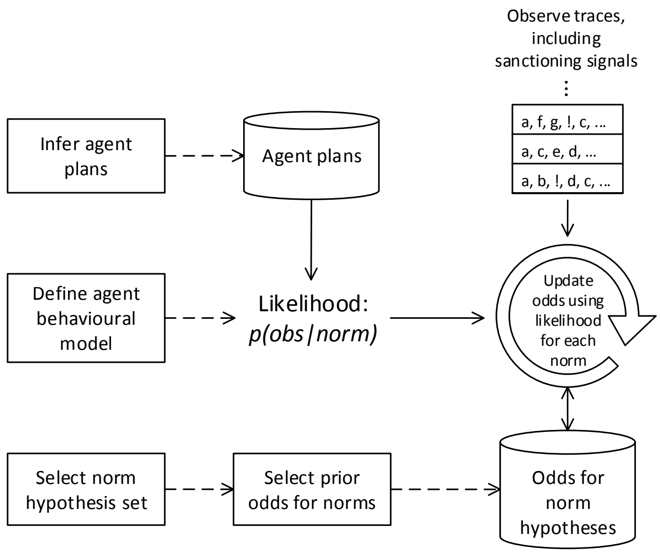


Figure 2: Overview of the approach

Figure 2 illustrates our approach to norm identification. The leftmost boxes in the diagram illustrate the inputs of our approach. The top left shows that, as discussed above, we must obtain an approximation (at least) of the plans that observed agents use to generate their publicly observable behaviour. We must also choose a set of candidate norms, as shown at the bottom left of the diagram. These are the hypotheses for which we iteratively compute their odds of being norms in the agent society (compared to the null hypothesis that there are no norms), as observations of agent behaviour are made. Section 2.1 describes our normative language and the instances of this language that form our hypothesis set in the graph traversal domain.[6] The middle left of Figure 2 shows one other requirement of our approach. Our Bayesian approach to norm identification involves computing the likelihood of observed behaviour, given each candidate norm and the null hypothesis. As we are assuming that agent behaviour is generated by plans, we need a model explaining how agents choose which plans to follow in the presence of norms. This is discussed in Section 2.5. In addition, we assume that agents may choose to sanction others if they have violated a norm, and that this sanctioning behaviour is not part of the plan execution process, but rather a reactive process that runs in parallel with plan execution. We model this by the use of parameters specifying society-wide probabilities of observing and then choosing to sanction norm violations. We also consider the possibility of agents choosing to punish other agents for their own (non-normative) reasons, and model the chance of this occurring using another parameter. These parameters and their use in computing the likelihood of observations are discussed in Section 2.4.

[6] The largest set of norm hypotheses arises if we consider all possible norm formulas generated by the normative language, which is what we use for our experiments.

The right hand side of Figure 2 shows the run-time Bayesian inference that updates the odds of the norm hypotheses as observations are made of agent behaviour. These observations are traces of agent movement on the graph, annotated with any sanctioning signals observed (denoted '!' in the figure). Given prior odds for each norm hypothesis (we currently use a uniform prior distribution), Bayes' Theorem explains how to update the odds for the norm hypotheses after making a new observation, by computing the likelihood of the observation under each hypothesis. This is discussed in Section 2.2.

2.1 Normative language

Our norm hypothesis space is defined by the following subset of linear temporal logic (LTL), where cn and n range over the labels of nodes in the graph[7], $\top$ denotes $true$, and $\Diamond$ and $\bigcirc$ denote "eventually" and "in the next state", respectively.

$$\begin{aligned}\text{NORM} = [\neg]\Diamond n \ &| \ cn \wedge \bigcirc\top \rightarrow [\neg]\bigcirc n \\ &| \ cn \wedge \bigcirc\top \rightarrow [\neg]\bigcirc\Diamond n\end{aligned}$$

These norms are interpreted as obligations or prohibitions constraining the agents' motion through the graph.

These three norm types, with and without the optional negation, are abbreviated and interpreted as follows (in the order shown above):

1. $eventually(n)$ / $never(n)$: These unconditional norms constrain a plan execution to include or exclude node n, and correspond to the obligation that n eventually occurs, or (respectively) that state n is prohibited.
2. $next(cn, n)$ / $not\_next(cn, n)$: These are conditional obligations and prohibitions, triggered whenever the agent reaches node cn (we refer to this as the "condition node" for the norm) and the end state has not been reached[8]. In this case the norm states that node n must be (or, respectively, must not be) the next node reached. We restrict our norm hypotheses to only include $next$ and not_next formulae for which there is an edge from cn to n in the graph.
3. $eventually(cn, n)$ / $never(cn, n)$: These are also conditional norms, expressing that *beginning from the node after the condition node*, node n must be eventually reached or (respectively) never reached.

Since $eventually$ and $next$ norms are obligations and $never$ and not_next are prohibitions, they could alternatively be expressed using explicit deontic modalities, with temporal logic semantics for each modality that specify the traces in which future violations are deemed to occur (e.g. see the approach of Broersen et al. [5]). However, for our purpose in this paper, the syntax above and the semantics of violation given (later) in Table 1 are sufficient.

2.2 Bayesian updating

Bayesian approaches to machine learning make use of Bayes' Theorem, which in its *diachronic interpretation* states how the probability of a hypothesis H should be updated in the light of new data D.

$$p(H|D) = \frac{p(H)p(D|H)}{p(D)}$$

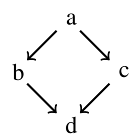
The probability $p(H)$ is known as the *prior* probability of hypothesis H, $p(D|H)$ is the *likelihood* of the observed data D given the hypothesis, and $p(H|D)$ is the *posterior* probability of H given D. The denominator $p(D)$ is the probability of the data being observed under any hypothesis, and is a normalising term for the probabilities $p(H|D)$.

The calculation can be repeated as further data is observed by replacing the prior with the previously calculated posterior and using Bayes' Theorem again to compute an updated posterior. This process is known as *Bayesian updating*.

If $\mathbf{H}$ is a mutually exclusive and collectively exhaustive set of hypotheses, the denominator can be expanded as follows.

$$p(H|D) = \frac{p(H)p(D|H)}{\sum_{H' \in \mathbf{H}} p(H')p(D|H')}$$

However, the hypotheses of interest in a problem domain may not be mutually exclusive (independent of each other), and/or we may not be able to enumerate a finite set of hypotheses. This is the case when the hypotheses are norms that may hold in a society. Norms may not be independent of each other, and this can depend on the environment, for example, given the graph below and the goal of travelling from node a to node d, a norm prohibiting movement to node b after visiting node a has precisely the same effect as a norm obliging travel to c after visiting a.



2.3 Updating the odds of norms

When the normalising term $p(D)$ cannot be easily computed, e.g. because the hypotheses are not mutually exclusive and collectively exhaustive, an alternative approach to using Bayes' Theorem is to work with *odds*. The odds of hypothesis H_1 over hypothesis H_2, given some observed data D, is denoted $O(H_1{:}H_2|D)$ and is defined as follows:

$$\begin{aligned}O(H_1{:}H_2|D) = \frac{p(H_1|D)}{p(H_2|D)} &= \frac{p(H_1)p(D|H_1)/p(D)}{p(H_2)p(D|H_2)/p(D)} \\ &= O(H_1{:}H_2)\frac{p(D|H_1)}{p(D|H_2)}\end{aligned}$$

where $O(H_1{:}H_2) = \frac{p(H1)}{p(H2)}$ denotes the prior odds of H_1 with respect to H_2.

In this formulation the normalising constant $p(D)$ cancels out and the odds of two competing hypotheses given new data can be computed using only the prior odds and the likelihoods of the two hypotheses. The probabilities of all other hypotheses do not need to be considered. In this paper we consider the odds of our hypotheses of interest (norms) compared to a specific null hypothesis: the hypothesis that there are no norms, written $H_\emptyset$. We write $O_\emptyset(H) = O(H{:}H_\emptyset)$ for the prior odds of H and $O_\emptyset(H|D) = O(H{:}H_\emptyset|D)$ for the posterior odds of H given D. By definition, $O_\emptyset(H_\emptyset) = 1$. For other norm hypotheses we set the prior odds uniformly to an arbitrary value less than one. The precise values of prior odds are unimportant for our work as we are interested in finding the norms with the *maximum* odds compared to the null hypothesis.

Whenever new data D is observed, we can then update the posterior odds for each norm hypothesis H by multiplying them by the ratio of the likelihoods of D given H and $H_\emptyset$. We consider two sources of evidence for norms. For each observation, we separately compute its likelihood based on (a) the observed sanctioning signals, and (b)

[7] Technically, cn and n are *nominals* from Hybrid Logic [4, p.435]: propositional symbols that are constrained to be true in exactly one state.

[8] Formally, the end state of a trace can be identified as the one in which $\bigcirc\top$ is false.

procedure update-odds$(p, s, g, P, \mathbf{H})$
begin
likelihood$^{sig}_{H_\emptyset} = p^{sig}(p, s \mid H_\emptyset)$
likelihood$^{plans}_{H_\emptyset} = p^{plans}(p \mid H_\emptyset, g, P)$
for $n \in \mathbf{H}$
$O_\emptyset(n) = O_\emptyset(n) * p^{sig}(p, s \mid n)$ / likelihood$^{sig}_{H_\emptyset}$
$O_\emptyset(n) = O_\emptyset(n) * p^{plans}(p \mid n, g, P)$ / likelihood$^{plans}_{H_\emptyset}$
end
end

Figure 3: The procedure for updating odds

Table 1: Violation indices $v_n(p)$ for the six norm types, given the path $p = \langle p_1, \cdots, p_\ell \rangle$

Norm type	Violation indices
$eventually(n)$	$\{\ell\}$ if $\forall i \in \{1, \cdots, \ell\}\ p_i \neq n$, else $\varnothing$
$never(n)$	$\{i : p_i = n\}$
$next(cn, n)$	$\{i{+}1 : 1 \leq i < \ell \wedge p_i = cn \wedge p_{i+1} \neq n\}$
$not\_next(cn, n)$	$\{i{+}1 : 1 \leq i < \ell \wedge p_i = cn \wedge p_{i+1} = n\}$
$eventually(cn, n)$	$\{\ell\}$ if $\exists i \in \{1, \cdots, \ell\}\ (p_i = cn \wedge \forall j \in \{i{+}1, \cdots, \ell\}\ p_j \neq n)$ else $\varnothing$
$never(cn, n)$	$\varnothing$ if $\forall i \in \{1, \cdots, \ell\}$, $p_i \neq n$, else $\{j : \min(\{i : p_i = cn\}) < j \leq \ell \wedge p_j = n\}$

a plan-based approach, and update the odds based on each of these. Each observation consists of a path p and a set s of path indices at which sanctioning signals were observed. For our norm hypotheses we only consider a single norm at a time[9], i.e., our hypothesis set $\mathbf{H}$ consists of all norms from the language defined in Section 2.1.

The procedure for updating the odds[10] for all norm hypotheses in the hypothesis set $\mathbf{H}$, given a new observation $\langle p, s \rangle$ is shown in Figure 3, where p^{sig} and p^{plans} are as defined in the following sections. The parameters passed to the update-odds function are the observed path and sanctioning signals, a goal g and set P of plans used by the plan-based likelihood computation, and the norm hypothesis set.

2.4 The likelihood of observed sanctions

We assume that agents may (sometimes) observe paths traversed by other agents in the graph. The observed paths represent possibly partial traversals of the graph by the other agents: they may be segments of longer paths traversed, but there are no unobserved nodes internal to the paths. Violations are detected through *signalling actions* (also known as *signals*) that indicate sanctioning of the observed agent [18, 19]. Such a signalling action could occur due to norm violation, or due to the sanctioner sanctioning the observed agent improperly (e.g., due to maliciousness, or a violation of the sanctioner's personal values). We model the latter case by assuming there is a small population-wide probability p_{pun} of a non-normative punishment signal being observed after any step of an observed path. We also assume there are fixed probabilities of norm violations being observed (p_{obs}) and of observed violations being sanctioned (p_{sanc}). We model all signalling actions by a single symbol—we do not assume that sanctions are specific to particular norms, nor that agents can distinguish normative sanctions from non-normative punishments.

[9] Our approach can be extended to consider hypotheses that are non-singleton norm sets, but we leave this for future work. If more than one norm can hold, it is still useful to identify the norms with the highest individual relative odds, before considering which sets of norms to add to the hypothesis space.

[10] In our implementation, we work with log odds.

function choose-plan(goal, plans, norm)
1. poss-plans = plans(goal, plans)
2. Decide whether to be norm-compliant
3a. **if** norm-compliant
nvp = non-viol-plans(poss-plans, norm)
if nvp = $\emptyset$
return null
else
return random-weighted-choice(nvp)
3b. **else**
if poss-plans = $\emptyset$
return null
else
return random-weighted-choice(poss-plans)

Figure 4: Model for an agent's choice of plan

Given an observed trace annotated with sanctioning actions, we can compute the likelihood of this observation given a hypothesized norm as follows.

Let $p = \langle p_1, \cdots, p_\ell \rangle$ be an observed path and the set s be a record of the indices of the path at which signals were observed.[11] The signal is represented by including index i in set s. We define $v_n(p)$ as the set of indices of the path p at which violations of the norm n occurred, defined in Table 1. The occurrence of $i \in v_n(p)$ is interpreted as the violation occurring *after* the action to move to node p_i, and may be a result either of that move or of the path ending if the destination node has been reached and an $eventually$ norm is violated.

Given a norm hypothesis n, the likelihood of observing trace $p = \langle p_1, \cdots, p_\ell \rangle$, where set s contains the indices at which sanction or punishment signals were observed, is then:

$$p^{sig}(p, s \mid n) = \prod_{1 \leq i \leq \ell} p^{sig}_i(i \in v_n(p),\, i \in s \mid n)$$

where p^{sig}_i, which takes two Boolean arguments, denotes the likelihood of the observation at path index i, as defined by the following table.

	$i \in s$	$i \notin s$
$i \in v_n(p)$	$p_{pun} + ((1-p_{pun}) \cdot p_{obs} \cdot p_{sanc})$	$(1-p_{pun}) \cdot (1-p_{obs} . p_{sanc})$
$i \notin v_n(p)$	p_{pun}	$1-p_{pun}$

The first row of the table is for the case when a violation occurs at index i. If a signal is observed at i, then this is either a non-normative punishment or the violation was observed and sanctioned. If no signal is observed, then there is no punishment and the violation has not been both observed and sanctioned. When there is no violation at i (second row), a signal can only be a non-normative punishment, so the likelihood of a signal occurring (or not) is the probability of the punishment occurring (or not).

2.5 Likelihood using knowledge of agent plans

Following the approach of Oren and Meneguzzi [16] we can use knowledge of agent plans (e.g. through plan recognition [22]) to compute the likelihood of an observed path through the graph (ignoring any sanction or punishment signals). We assume that all agents

[11] We currently assume that sanctions are applied (if at all) *immediately* after a movement to a node p_i in the path causes a norm to be violated. Relaxing this assumption would require a more complex likelihood function that considers the possible matches of signals with possible past violations.

$$p^{plans}(o \mid n, g, P) = p_{comp} \left(\sum_{\pi \in \text{non-viol-plans}(\text{plans}(g,P),n)} \frac{\text{weight}(\pi)}{\sum_{\pi' \in \text{non-viol-plans}(\text{plans}(g,P),n)} \text{weight}(\pi')} p(o \mid \pi) \right)$$
$$+ (1 - p_{comp}) \left(\sum_{\pi \in \text{plans}(g,P)} \frac{\text{weight}(\pi)}{\sum_{\pi' \in \text{plans}(g,P)} \text{weight}(\pi')} p(o \mid \pi) \right)$$
$$= p_{comp} \frac{\sum_{\substack{\pi \in \text{non-viol-plans}(\text{plans}(g,P),n) \\ \cap\, \text{plans-containing}(\text{plans}(g,P),o)}} \text{weight}(\pi)}{\sum_{\pi \in \text{non-viol-plans}(\text{plans}(g,P),n)} \text{weight}(\pi)} + (1 - p_{comp}) \frac{\sum_{\substack{\pi \in \text{plans}(g,P) \\ \cap\, \text{plans-containing}(\text{plans}(g,P),o)}} \text{weight}(\pi)}{\sum_{\pi \in \text{plans}(g,P)} \text{weight}(\pi)}$$

Figure 5: Likelihood of an observed path using knowledge of agent plans

share the same set of possible plans (choices of paths in the graph), and that the observing agent can infer the observed agent's goal (comprising starting and destination nodes).

To define the likelihood of an observed path given a norm hypothesis, a goal and a plan library, we require a model for the decision-making process of the observed agents, which must choose and execute plans to achieve their goals in the possible presence of a norm. The analysis in this section is based on the decision-making model shown in Figure 4.

In this model, the agent first generates all plans for the goal. The returned plans may be weighted, (e.g., to indicate agent preferences or execution costs), but our examples use equal weights for simplicity. Next, the agent decides whether it it will act in a norm-compliant manner. If so, it filters the possible plans to keep only those that do not violate the norm, and chooses a plan using a random weighted choice. Otherwise, it makes a random weighted choice from the full set of plans for the goal. Note that this is intended to be a simple abstract model for the purpose of defining a likelihood function in the absence of any information about an observed agent. We do not claim that this is, or should be, the exact control mechanism used in any agent implementation.

We define the likelihood, based on knowledge of agent plans, of an observed path o on the graph, given a norm hypothesis n, an inferred goal g and a set of plans P, as shown in Figure 5. We write p_{comp} for the rate of norm compliance in the society. The first two lines of the figure multiply the probability of choosing a plan π and the probability $p(o \mid \pi)$ that the plan contains the observed path, for the norm-compliance and non-norm-compliance cases. As $p(o \mid \pi)$ is either 1 or 0, the last two lines replace this factor with a union in the limits of the sum. The function non-viol-plans filters out plans that cause violations, using the violation indices function v_n (Table 1).

There are two cases when the formula in Figure 5 cannot be evaluated due to zero values in the denominator of a fraction: when there are no plans for the inferred goal, and when there are no norm-compliant plans for the goal. The former case invalidates our assumption that the observed behaviour is generated using a plan taken from a known set of plans to fulfil the inferred goal, and we abandon the odds update based on plan knowledge for the current observation. In the latter case we replace the first addend in the last line of the figure with 0. This represents the assumption that a norm compliant agent would have abandoned its goal in this case.

3 Experiments

We have performed a set of experiments to validate and evaluate our approach to norm detection. These experiments use a random graph [9] containing 35 nodes, representing states in a state space, and edges connecting these nodes represent actions available to an agent. The algorithm we used to generate observations works exactly like an agent randomly choosing possible plans to reach a goal state, subject to the normative constraints, with a certain probability. Table 2 summarises the parameters common to all experiments.

Table 2: Parameters used across experiments.

p_{pun}	0.01	p_{sanc}	0.99
p_{obs}	0.99	Prior odds $O_\emptyset(n)$	0.5

Here, we differentiate the norms used to generate the observations, which we call N_g, from the norms inferred by our norm detection approach, which we call N_d. Norm likelihood is estimated using log odds against there being no norm (rather than a probability), and we infer the norms from a set of observations by ranking the odds of each norm (against no norm), and consider a given number of the norms with the highest odds to be the norms in the society.

In the experiments, we ran an agent fully aware of the norms to generate a random set of observations following the algorithm of Figure 4 (i.e. random, but norm-compliant behaviour), while allowing for the possibility of occasional non-compliant behaviour with a probability set at 1% unless otherwise noted. Thus $(1 - p_{comp}) = 0.01$. We then applied our norm identification method to these observations in order to compute the odds of all possible norm hypotheses. Our experiments consisted of submitting a sequence of 100 observations to an agent and measuring its ability to produce norm-compliant behaviour after increasing numbers of observations. Each experiment was repeated 50 times and the results were averaged to reduce the impact of any particularly informative or uninformative sequence of observations.

Our aim is to allow an agent to undertake norm-compliant behaviour, even without an exact model of the norms. One approach to doing this is to consider the T most likely norms, even if they are less likely than the null hypothesis (i.e., they may not exist since they are less likely than there being no norm). In such a situation, the agent can be thought of as acting conservatively, as it may avoid potentially permitted courses of action. Another approach is to have the agent consider only norms that are more likely than the null hypothesis[12]. In the remainder of this work, we consider the functioning of an agent utilising norm identification and acting using the conservative approach, and evaluate its effectiveness in generating norm compliant behaviour.

Our experiments measure precision and recall as a function of the number of observations supplied to the detection mechanism.

[12] However, this is dependent on the prior odds chosen.

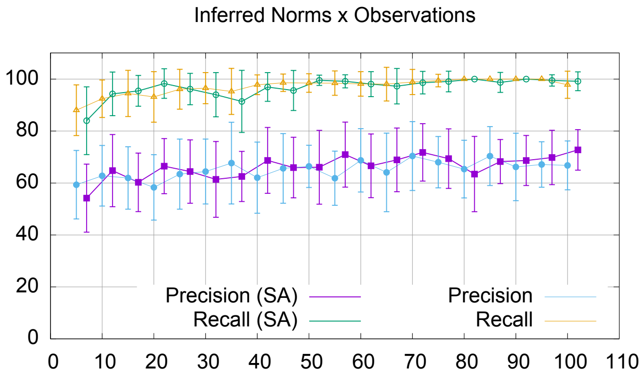


Figure 6: Inferred norms

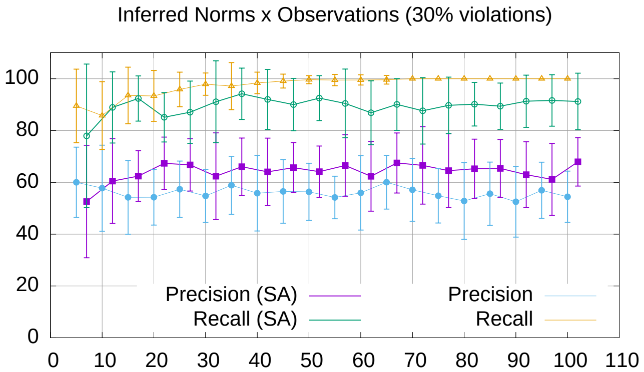


Figure 7: Inferred norms, with 30% violations

As norms may subsume each other, we measured precision and recall in terms of compliant behaviour rather than the exact norms inferred. Thus, in our evaluation, precision means the fraction of norm-compliant plan executions generated by an agent using a sample of the inferred norms N_d to drive its behaviour, and recall is the fraction of plans that are compliant with the true underlying norms N_g that are sampled by the agent. Specifically, we compute precision by updating the odds of each norm hypothesis after an observation and choosing the set N_d with the top T norms.

Our experiments are illustrated in the graphs of Figures 6–8. Our first set of experiments, illustrated in Figure 6, shows precision and recall as a function of the number of observations with error bars denoting standard deviation for these measures. Each experiment was conducted in the presence and absence of sanctioning actions—"SA" in the figure legends indicate the use of sanctioning actions.[13] The results indicate that precision varies from a starting point of around 55 without sanctioning actions and 60 with them. Precision and recall increase rapidly in the first 10 observations and then tend to slowly increase as more observations are taken into consideration, while variance in precision diminishes as more observations are made.

Moreover, as we increase the amount of non-compliant behaviour in the observations, illustrated in Figure 7, the effect of sanctioning actions becomes more pronounced. Precision improves, while recall gets worse, as potential norm compliant plans are filtered out (possibly due to non-normative punishment actions) while fewer non-compliant plans are executed. Thus, as long as the agent has choices

[13] Note that the values for the sanctioning action (SA) case are shifted two points to the right to prevent error bars from overlapping.

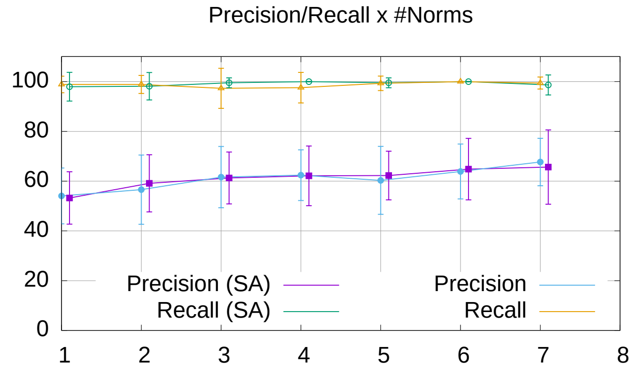


Figure 8: Precision and recall as a function of norms

available to itself with regards to plans which it can execute, sanctioning actions appear to be beneficial. It is also important to note that while recall is lower in the presence of sanctioning actions, it still remains relatively high.

We also computed precision and recall as a function of the number of norms in the system. In this experiment, we randomly selected elements of the power set $\mathcal{P}(N_g)$ of the set of seven norms in the scenario and, for each size of subset of $\mathcal{P}(N_g)$, we determined precision and recall after 100 observations. The results, shown in Figure 8, indicate a slight increase in precision and recall as more norms are introduced into the system. However, the highly overlapping error bars prevent us from asserting that the number of norms present in the system has an impact on the ability of the approach to infer norms.

Finally, we ran experiments comparing the effectiveness of our approach against the previous state-of-the-art approaches, each of which makes different assumptions about the behaviour of the underlying agents. We compared the Bayesian approach to the approaches of Savarimuthu et al. [18, 19], which assume that observed agents generate non-compliant behaviour for norms to be inferred, and that of Oren and Meneguzzi [16], which assumes that observed agents mostly comply with the norms[14]. In all of these comparative experiments we generated 20 observations and included sanctioning actions, since the approaches of Savarimuthu et al. relied on these signals. Like our previous experiment, we performed 50 repetitions to smooth out random variations in the observations. Unlike our previous experiments, which used 100 observations, we stopped at 20 observations, since by the 20th observation precision and recall had dropped to 0 for all comparable approaches. We conducted experiments with compliance ranging from almost full (a 0.01% probability of violations) to none (a 100% probability of violations). The results are presented in Table 3, which shows, for each approach, the mean of the following measures after a given number of observations (indicated by the #Obs column) have been made: the number of norms inferred, precision and recall. Standard deviations for each measure are shown in parentheses. Here, the Bayesian approach clearly outperforms previous approaches after a very small number of observations. For all competing approaches, precision and recall tend to drop towards 0 as the number of observations increase. Since these approaches infer norms more aggressively, they tend to overconstrain the agent's behaviour, inferring many more norms than the Bayesian approach. As a result, the competing approaches tend to generate no "true positive" plan evaluations, that is, for plans that are compliant with the real underlying norms, these approaches tend to erroneously evaluate them as being non-compliant.

[14] In the comparison we describe, a threshold of 0.5 was used in Oren and Meneguzzi's second algorithm.

Table 3: Comparison of the Bayesian approach with the data mining [19, 18] and plan recognition [16] approaches. In each line, the best results are shown in bold.

#Obs	#Norms (σ)	Precision (σ)	Recall (σ)	#Obs	#Norms (σ)	Precision (σ)	Recall (σ)	#Obs	#Norms (σ)	Precision (σ)	Recall (σ)
Bayesian Approach				**Data Mining Approach [19, 18]**				**Plan Recognition Approach [16]**			
Probability of violation=0.01				Probability of violation=0.01				Probability of violation=0.01			
1	18.00 (0.00)	**62.63** (27.50)	**43.16** (28.08)	1	7.90 (2.14)	60.00 (48.98)	9.56 (9.60)	1	30.00 (0.00)	60.00 (48.98)	11.43 (10.46)
5	18.00 (0.00)	**59.59** (15.91)	**86.31** (12.81)	5	10.55 (2.97)	25.00 (43.30)	3.16 (5.88)	5	28.25 (1.84)	0.00 (0.00)	0.00 (0.00)
10	18.00 (0.00)	**63.58** (8.94)	**91.87** (7.22)	10	8.20 (2.42)	50.00 (50.00)	9.40 (11.07)	10	27.95 (1.85)	0.00 (0.00)	0.00 (0.00)
15	18.00 (0.00)	**59.97** (11.64)	**90.17** (12.48)	15	7.35 (1.15)	55.00 (49.74)	6.44 (6.28)	15	27.50 (2.71)	0.00 (0.00)	0.00 (0.00)
20	18.00 (0.00)	**64.76** (12.24)	**95.54** (7.16)	20	7.95 (2.29)	45.00 (49.74)	5.22 (6.47)	20	27.00 (2.38)	0.00 (0.00)	0.00 (0.00)
Probability of violation=0.3				Probability of violation=0.3				Probability of violation=0.3			
1	18.00 (0.00)	**40.43** (32.15)	**31.48** (29.33)	1	7.30 (2.12)	30.00 (45.82)	5.22 (9.02)	1	30.00 (0.00)	40.00 (48.98)	6.51 (10.20)
5	18.00 (0.00)	**52.59** (21.72)	**77.94** (27.72)	5	11.75 (3.25)	15.00 (35.70)	1.37 (3.27)	5	27.25 (2.38)	0.00 (0.00)	0.00 (0.00)
10	18.00 (0.00)	**60.50** (16.34)	**88.92** (13.73)	10	13.45 (3.84)	10.00 (30.00)	1.26 (4.18)	10	26.50 (2.01)	0.00 (0.00)	0.00 (0.00)
15	18.00 (0.00)	**62.42** (9.77)	**92.37** (8.74)	15	15.20 (4.28)	5.00 (21.79)	0.38 (1.67)	15	26.15 (2.72)	0.00 (0.00)	0.00 (0.00)
20	17.95 (0.21)	**67.36** (10.12)	**85.14** (9.54)	20	16.45 (3.33)	5.00 (21.79)	0.90 (3.96)	20	25.15 (2.39)	0.00 (0.00)	0.00 (0.00)
Probability of violation=0.6				Probability of violation=0.6				Probability of violation=0.6			
1	18.00 (0.00)	58.48 (34.09)	**34.77** (30.56)	1	6.40 (2.53)	70.00 (45.82)	10.62 (9.87)	1	30.00 (0.00)	**80.00** (40.00)	14.82 (11.65)
5	18.00 (0.00)	**54.58** (14.59)	**82.34** (15.77)	5	13.50 (3.78)	15.00 (35.70)	1.59 (3.95)	5	27.20 (1.99)	0.00 (0.00)	0.00 (0.00)
10	18.00 (0.00)	**55.63** (11.26)	**82.41** (16.93)	10	13.20 (3.52)	15.00 (35.70)	2.85 (8.91)	10	27.35 (2.57)	0.00 (0.00)	0.00 (0.00)
15	18.00 (0.00)	**58.57** (13.48)	**87.97** (9.58)	15	15.75 (4.49)	0.00 (0.00)	0.00 (0.00)	15	26.45 (2.20)	0.00 (0.00)	0.00 (0.00)
20	18.00 (0.00)	**62.34** (10.90)	**81.03** (12.50)	20	17.05 (4.42)	0.00 (0.00)	0.00 (0.00)	20	25.70 (2.23)	0.00 (0.00)	0.00 (0.00)
Probability of violation=1				Probability of violation=1				Probability of violation=1			
1	18.00 (0.00)	35.13 (35.01)	**22.97** (23.26)	1	6.10 (2.56)	45.00 (49.74)	8.80 (11.05)	1	30.00 (0.00)	**60.00** (48.98)	8.93 (9.38)
5	17.95 (0.21)	**61.71** (26.40)	**50.87** (30.00)	5	11.40 (5.04)	5.00 (21.79)	0.38 (1.67)	5	27.45 (1.59)	0.00 (0.00)	0.00 (0.00)
10	18.00 (0.00)	**52.05** (20.25)	**69.37** (28.50)	10	12.15 (4.87)	0.00 (0.00)	0.00 (0.00)	10	27.15 (1.45)	0.00 (0.00)	0.00 (0.00)
15	18.00 (0.00)	**63.59** (20.83)	**73.97** (24.45)	15	17.25 (6.54)	0.00 (0.00)	0.00 (0.00)	15	26.50 (1.96)	0.00 (0.00)	0.00 (0.00)
20	17.95 (0.21)	**56.83** (19.22)	**74.22** (21.27)	20	16.90 (5.65)	0.00 (0.00)	0.00 (0.00)	20	26.15 (1.98)	0.00 (0.00)	0.00 (0.00)

We note that, given the different expressivity of the competing approaches, and the fact that our experiments used norms expressed in the temporal modalities used in our approach, there may be a mismatch in the detection capabilities within the experiments. Nevertheless, since our experiments measured precision and recall in terms of *compliant behaviours* rather than the specific norms, we believe that our analysis is valid.

4 Discussion and Conclusions

The Bayesian approach presented in this paper combines ideas from the sanctioning action observation [19, 18] and plan recognition [16] norm identification approaches to create a powerful new mechanism. As our experiments indicate, we generate norm-compliant behaviour in a norm-identifying agent approximately 60% of the time across a range of violation likelihoods, and show that the presence of sanctioning actions substantially improves recall.

As mentioned in Section 2, we assume that we can determine an agent's starting point and goal. If we consider AgentSpeak(L) style agents, then the identification of a plan, and from this its context and guard conditions, would allow an agent to determine the observed agent's start point in many situations. Furthermore, once a plan has been identified, its goal can be trivially determined. While we assume that sanctioning actions (signals) are observable, we do not assume that it is possible to associate specific norms with specific signals. We also assume that one cannot differentiate between sanction and punishment signals. Lifting such restrictions would improve the learning rate, but is not realistic.

We intend to pursue several avenues of future work. First, while our model permits it, we have not evaluated the effects of conflicting norms on the norm identification process or on the conservative strategy described in this paper, and we intend to investigate what additional mechanisms must be created to function in such domains. Given that norms can subsume others, we believe the use of a subsumption-based norm conflict resolution mechanism [24] would result in an agent with an enhanced ability to identify norms and act in a norm compliant manner. We also plan to investigate weighting mechanisms and their effects on norm identification. Such mechanisms could, for example, originate from a trust and reputation model [12]. Here, highly trusted agents could be assumed to (normally) act in a norm-compliant manner, while less trustworthy agents would be expected to trigger more sanctioning actions. The addition of trust information could allow us to consider which agents are performing the sanctioning actions. An agent often signalling that a trustworthy agent is acting in a norm-violating manner could have its opinion discounted, while those signalling that untrustworthy agents are violating norms could have their opinion strengthened. The addition of such a mechanism should improve the performance of our model, and given the Bayesian underpinnings of many trust systems [23], should be a relatively straightforward addition.

Another source of weightings we could exploit within the model originates from the plans themselves. In this work, we assumed that all plans to achieve some goal are equally likely to be used by an agent. However, some of these plans could be more expensive (e.g. from a resource utilisation point of view) than others, and a utility maximising agent would be expected to select cheaper plans (subject to normative constraints). We believe that the use of such weights would increase the rate at which norms are identified, and also increase the precision and recall of our approach.

ACKNOWLEDGEMENTS

F. Meneguzzi thanks Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through the Universal Grant (Grant ref. 482156/2013-9) and PQ fellowship (Grant ref. 306864/2013-4).

REFERENCES

[1] Huib Aldewereld, Frank Dignum, Andrés García-Camino, Pablo Noriega, Juan A. Rodríguez-Aguilar, and Carles Sierra, 'Operationalisation of norms for usage in electronic institutions', in *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 223–225. ACM, 2006.

[2] Wagdi Alrawagfeh, Edward Brown, and Manrique Mata-Montero, 'Norms of behaviour and their identification and verification in open multi-agent societies', *International Journal of Agent Technologies and Systems*, **3**(3), 1–16, 2011.

[3] *Normative Multi-Agent Systems*, eds., Giulia Andrighetto, Guido Governatori, Pablo Noriega, and Leendert W. N. van der Torre, volume 4 of *Dagstuhl Follow-Ups*, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2013.

[4] P. Blackburn, M. de Rijke, and Y. Venema, *Modal Logic*, Cambridge University Press, 2001.

[5] Jan Broersen, Frank Dignum, Virginia Dignum, and John-Jules Ch. Meyer, 'Designing a deontic logic of deadlines', in *Deontic Logic in Computer Science*, volume 3065 of *Lecture Notes in Computer Science*, 43–56, Springer, 2004.

[6] Marco Campenní, Giulia Andrighetto, Federico Cecconi, and Rosaria Conte, 'Normal = Normative? The role of intelligent agents in norm innovation', *Mind & Society*, **8**(2), 153–172, 2009.

[7] Cristiano Castelfranchi, Frank Dignum, Catholijn M. Jonker, and Jan Treur, 'Deliberative normative agents: Principles and architecture', in *Intelligent Agents VI. Agent Theories, Architectures, and Languages*, 364–378, Springer, 2000.

[8] Stephen Cranefield, Tony Savarimuthu, Felipe Meneguzzi, and Nir Oren, 'A Bayesian approach to norm identification (extended abstract)', in *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1743–1744. IFAAMAS, 2015.

[9] P. Erdös and A. Rény, 'On random graphs', *Publicationes Mathematicae*, **6**, 290–297, 1959.

[10] Michael Luck, Samhar Mahmoud, Felipe Meneguzzi, Martin Kollingbaum, Timothy J. Norman, Natalia Criado, and Moser Silva Fagundes, 'Normative agents', in *Agreement Technologies*, ed., Sascha Ossowski, volume 8 of *Law, Governance and Technology Series*, 209–220, Springer, 2013.

[11] Moamin A. Mahmoud, Mohd Sharifuddin Ahmad, Azhana Ahmad, Mohd Zaliman Mohd Yusoff, and Aida Mustapha, 'The semantics of norms mining in multi-agent systems', in *Computational Collective Intelligence. Technologies and Applications*, volume 7653 of *Lecture Notes in Computer Science*, 425–435, Springer, 2012.

[12] S. Marsh, *Formalising trust as a computational concept*, Ph.D. dissertation, University of Stirling, 1994.

[13] Felipe Meneguzzi, Odinaldo Rodrigues, Nir Oren, Wamberto W. Vasconcelos, and Michael Luck, 'BDI reasoning with normative considerations', *Engineering Applications of Artificial Intelligence*, **43**, 127–146, 2015.

[14] Felipe Rech Meneguzzi and Michael Luck, 'Norm-based behaviour modification in BDI agents', in *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, pp. 177–184. IFAAMAS, 2009.

[15] Sanjay Modgil, Nir Oren, Noura Faci, Felipe Meneguzzi, Simon Miles, and Michael Luck, 'Monitoring compliance with e-contracts and norms', *Artificial Intelligence and Law*, **23**(2), 161–196, 2015.

[16] Nir Oren and Felipe Meneguzzi. Norm identification through plan recognition. Proceedings of the 15th International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems, http://www.staff.science.uu.nl/~dignu101/coin2013/papers/20130161.pdf, 2013.

[17] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Martin K Purvis, and Maryam A Purvis, 'Norm emergence in agent societies formed by dynamically changing networks', *Web Intelligence and Agent Systems*, **7**(3), 223–232, 2009.

[18] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis, 'Obligation norm identification in agent societies', *Journal of Artificial Societies and Social Simulation*, **13**(4), 3, 2010.

[19] Bastin Tony Roy Savarimuthu, Stephen Cranefield, Maryam A. Purvis, and Martin K. Purvis, 'Identifying prohibition norms in agent societies', *Artificial Intelligence and Law*, **21**(1), 1–46, 2013.

[20] Constantin Serban and Naftaly H. Minsky, 'In vivo evolution of policies that govern a distributed system', in *Proceedings of the 10th IEEE International Conference on Policies for Distributed Systems and Networks*, pp. 134–141, 2009.

[21] Yoav Shoham and Moshe Tennenholtz, 'Emergent conventions in multi-agent systems: Initial experimental results and observations', in *Proceedings of the Third International Conference on the Principles of Knowledge Representation and Reasoning*, pp. 225–231. Morgan Kaufmann, 1992.

[22] Gita Sukthankar, Robert P. Goldman, Christopher Geib, David V. Pynadath, and Hung Hai Bui, *Plan, Activity, and Intent Recognition: Theory and Practice*, Elsevier, 2014.

[23] W.T. Luke Teacy, Michael Luck, Alex Rogers, and Nicholas R. Jennings, 'An efficient and versatile approach to trust and reputation using hierarchical Bayesian modelling', *Artificial Intelligence*, **193**, 149–185, 2012.

[24] Wamberto Vasconcelos, Martin J. Kollingbaum, and Timothy J. Norman, 'Resolving conflict and inconsistency in norm-regulated virtual organizations', in *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 632–639. ACM, 2007.

[25] Georg Henrik von Wright, *Norm and action: a logical enquiry*, Routledge & Kegan Paul, 1963.