A general methodology for teaching norms to social robots

Bertram Malle
Brown University

Eric Rosen
Brown University

A General Methodology for Teaching Norms to Social Robots*

Bertram F. Malle, Eric Rosen, Vivienne B. Chi, Matthew Berg, Peter Haas
Brown University

Abstract— **Human behavior is powerfully guided by social and moral norms. Robots that enter human societies must therefore behave in norm-conforming ways as well. However, there is currently no cognitive, let alone computational model available of how humans represent, activate, and learn norms. We offer first steps toward such a model and apply it to the design of a norm-competent social robot. We propose a general methodology for such a design, from empirical identification of relevant norms to computational implementations of norm learning to thorough and iterative evaluation of the robot's norm compliance by means of community feedback.**

I. Introduction

No human community can exist without norms [1]—the social and moral rules of behavior that community members are expected to follow, for the benefit of the community as a whole. At a steady pace, robots are entering domains such as health care, education, and security, and in these domains robots will likewise be expected to follow the relevant social and moral norms. If robots do not follow such norms, they will be unlikely to benefit human communities, may cause damage, and will not be trusted. If we do require that robots follow social and moral norms, then we must understand and formalize what norms are in the human mind so that we can effectively design robots with the appropriate capacities to represent and obey such norms—in short, so that we can design robots with *norm competence* [2].

However, we currently know very little about how humans represent, learn, activate, and deploy norms to guide their behavior. Moreover, there is scant work on how researchers and designers should identify the relevant norms in the first place. If we decide to introduce a robot as a nurse assistant, for example, what are the norms this robot is expected to follow? And from a technical standpoint, how do we implement these norms in a way that allow the robot to quickly apply the proper norms in the right context? Therefore, the goals for this paper are: (1) propose a framework for how to design robots with norm competence; (2) present new experimental paradigms that help identify the norms of a target domain and community as well as key cognitive properties of norms; (3) outline first steps to implement norm sets in robots and continuously improve these norm sets through human-robot interaction.

II. Background

A. Behavioral Science Work on Norms

In sociology and experimental economics, the importance of norms has been recognized [3], but the focus there is on social forces "against one's rational self-interest," not on how norms operate cognitively and computationally. A sizable set of studies have examined the automatic activation of norms by context-specific cues—for example, garbage on the floor triggering the "don't litter" norm [4] or the sight of a library triggering the "be quiet" norm [5]. But no explicit cognitive model has been offered to account for these and other properties of norms.

People conceptualize norms primarily as *prescriptions* and *prohibitions* [6]. These two types of norms serve different functions. Prescriptions directly guide what action to perform. Prohibitions, by contrast, negate possible courses of action; they define constraints or limits but do not suggest how one *should* act. Moreover, an infinite number of prohibitions exist that prevent action, but even just one prescription suffices to promote action. These asymmetries may have a direct impact on how norms are cognitively represented and can be learned.

B. Formal and Computational Work on Norms

Two prominent research programs have dominated computational investigations of norms. One has long been dedicated to deontic logic and its variants, which formalizes notions of prescription and prohibition in specialized modal operators [7] and has recently been applied to what is called machine ethics [8]. This approach develops formal logics to arrive at morally acceptable conclusions from premises and to offer verification of the validity and reliability of reasoning. It also aims to be expressive enough to represent everyday norm reasoning [9]–[11].

The second research program grew out of work on multi-agent systems in AI [12] and integrated insights from the social sciences and BDI logic into the "BOID" (Belief-Obligation-Intention-Desire) framework [13]. Here, deontic concepts are combined with practical reasoning to allow for the resolution of conflicts between norms and desires, which are common in human agents. Sophisticated extensions of these formalisms aim to capture the interplay between cognitive properties and societal dynamics of norms [14] as well as the internalization

* This project was supported in part by grants from the National Science Foundation (NSF), 1717701, Office of Naval Research (ONR), N00014-14-l-0144, and from the Defense Advanced Research Projects Agency (DARPA), SIMPLEX 14-46-FP-097. The opinions expressed here are our own and do not necessarily reflect the views of NSF, ONR, or DARPA.

Bertram F. Malle is with the Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer St., Providence, RI 02912 USA (phone: 401-863-6820; e-mail: bfmalle@ brown.edu).

Eric Rosen and Matthew Berg are with the Department of Computer Science, Brown University, 115 Waterman St. Providence, RI 02906 USA.

Vivienne B. Chi is with the Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, 190 Thayer Street, Providence, RI 02912 USA

Peter Haas is with the Humanity Centered Robotics Initiative, Brown University, 201 Thayer Street, Providence, RI 02906, USA

of norms to counteract agents' selfishness [15]. These models treat norms as cognitive structures that are highly sensitive to context, to community members' norm-compliant behavior, and their demands on each other [16].

C. Robot Decision Making with Norm Competence

An artificial agent with norm competence must obey its community's norms when making decisions about its goal-directed behavior. Some authors have used deontic logic to combine the agent's goals and normative directives and resolve conflicts between them [17]. Others have criticized this approach because it seems to require all norms to be identified a priori and thus omit learning [18]. The general framework of Reinforcement Learning (RL) has recently gained prominence as it enables agents to learn the reward structure of their environment and to plan optimal action sequences that are sensitive to these rewards, even if not known in advance [19]. An RL framework can incorporate an agent's own goals, other people's interests, and ethical rules, all of which enter the agent's reward function [18].

However, without further differentiation, an RL agent's reward function would not distinguish between others' personal preferences and community-based social norms, nor between reward for intentions and reward for outcomes. These distinctions are critical for norm-guided human actions [20], but even sophisticated inverse reinforcement learners [21] do not make them. Some norm-learning frameworks allow the agent to distinguish widely held norms from merely idiosyncratic preferences by heeding the distributions of other agents' behaviors [22]. One general problem of norm-learning RL agents is that the demonstrations and rewards they observe are limited, often opaque, and the norms are rarely explicitly stated. Arnold et al. [20] suggest that typical RL algorithms are too impoverished to learn moral norms and should be complemented by explicit representations of deontic operators on actions—marrying the strengths of deontic logic with those of machine learning.

In sum, despite steps of progress toward building norm-competent artificial agents, many questions remain: Where should the norms come from? Full specification may be too cumbersome; mere learning by observation is misleading. How should conflicts between goals and norms be resolved? How can agents represent the full complexity of context-specific and community-based norms? And how do robots adapt to the subtle variations, exceptions, and changes of norm systems in social communities?

III. Broader Project: Robots with Norm Competence

Our approach to equipping robots with norms combines social-cognitive science with artificial intelligence and robotics. Figure 1 illustrates the general methodology of developing norm competence, which can be applied to any robot operating in a social domain. It has four stages. (1) *Norm identification.* Once the community of deployment and the robot's role are defined, empirical techniques can identify the relevant norms in that community for a suite of relevant contexts, yielding the *Starting Norm Base (SNB)*. (2) *Norm implementation* initially occurs in a disembodied AI that holds the *SNB* and relies on context-specific and community-sensitive decisions processes to plan its norm-abiding actions for specific tasks. (3) *Evaluation* initially consists of a *Norm Learning Game (NLG)* that the AI plays with human users, where a context is given, the system proposes a sequence of actions, and the user provides feedback on these proposals. This process helps the system improve its *SNB* to a *Refined Norm Base (RNB)*. (4) Iterative rounds of *Implementation* and *Evaluation* follow, and capacities can be expanded. For example, a scene recognition module would allow the system to autonomously detect the context it is in (first from photos, then videos); a communication module would make verbal speech processing and synthesizing more natural and could be extended to processing gesture and eye gaze [23]. Further, NLGs will be enriched by moving to virtual reality environments. Here, full interaction sequences can be tested with a simulated robot controlled with physics engines like Unity 3D, thus providing a safe environment for the robot to explore the boundaries of its norm awareness. When evaluations show the robot to be reliable in obeying norms and trustworthy to human users, deployment of the physical robot in the real world can safely proceed.

IV. Social and Moral Norms: Theory, Definitions

Norms vary by context, community, and era; they vary by degrees of demand (e.g., "suggested" vs. "required"; "discouraged vs. "forbidden"); and they range from highly specific (e.g., "stretching one's hand out for a handshake after the other did") to abstract (e.g., "showing respect"). Despite these variations, there is a fundamental structure of norms that can be captured in the following working definition (integrating multiple proposals [4], [16], [24], [25]):

A norm is an instruction, in a given community, to (not) perform an action in a given context, provided that a sufficient number of individuals in the community (i) demand, to a certain degree, of each other to follow the instruction and (ii) do in fact follow it.

More formally, a norm $\mathcal{N}$ is defined by the tuple:

$$\mathcal{N} := <S, C, \mathcal{A}, \mathcal{D}^f, \mathcal{P}> . \quad (1)$$

$\mathcal{N}$ always exists relative to a social community S and relative to a context C; it operates on an action $\mathcal{A}$ by way of a deontic operator $\mathcal{D}$ with force parameter f (how *strongly* people demand of each other to perform, or not perform, $\mathcal{A}$) and with a prevalence parameter $\mathcal{P}$ (to what extent the community actually obeys the norm). In deontic logic, $\mathcal{D}$ has often been conceptualized as coming in at least two kinds: prescriptions and prohibitions. Some have suggested, instead, that $\mathcal{D}$ be a continuum [26]–[28], for example ranging from -1 for the strongest prohibitions through permissions at 0 to +1 for the strongest prescriptions.

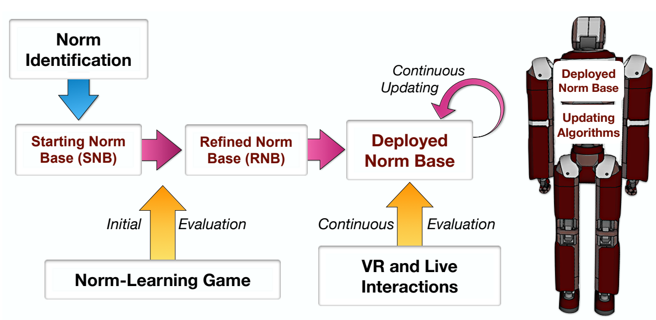

Figure 1. A general methodology for teaching robots norm competence.

One might also allow for uncertainty, such that the agent estimates $\mathcal{D}^f$ and $\mathcal{P}$ parameters within some interval, $[\alpha, \beta]$, potentially treated as Dempster-Shafer uncertainty [29] (see [22], [30] for more detail).

We should note that particular elements of the above definition separate norms from related concepts. The notion property of deontic force separates norms from preferences or goals. Many people put milk in their coffee, but they do not demand of each other to do so, hence this action is not a norm but a wide-spread preference. By contrast, getting in line when ordering coffee in a coffee shop is a norm—that is what people do and demand of each other. Further, context condition *C* distinguishes norms from values. Whereas norms are instructions to act in a particular way, values are ideals (e.g., loyalty, honesty) that may be achieved by a variety of possible actions. Nonetheless, norms are likely to reflect or express the general guidance that values provide.

We now turn to the question of how one might identify a community's norm sets for specific contexts. We present two experimental paradigms and initial empirical results on this question and briefly refer to work that hints at the graded deontic force of norms ($\mathcal{D}^f$). We then sketch how such norm sets can be used to modify robot action planning.

V. Experimental Paradigms and Results

We developed two paradigms to experimentally assess the norms that are activated by a particular context [31], [32]. In the first, people freely generate norms associated with various contexts, akin to the free-listing procedure in semantic network and scene-categorization research [33]. In the second, people are presented with a number of candidate norms and, for a given context, have to decide whether that norm fits or does not fit the context. We operationalized "context" as a pictured scene (e.g., restaurant) in which the person occupies a particular role (e.g., waitress). We collected data on three types of norms—permissions, prescriptions, prohibitions—but will focus here on prescriptions and prohibitions.

A. Norm-Generation Paradigm

In a first study, 80 participants were recruited via Amazon's Mechanical Turk. They were shown four pictures one at a time and asked to type, for up to 60 seconds, as many actions as came to mind that (a) one is "supposed to do here" (eliciting prescribed actions) or (b) one is "not allowed to do here" (eliciting prohibited actions). Each participant answered only one of these questions, for 4 contexts (out of 8 total). We refer to these prescribed or prohibited actions that people generated (e.g., talking, putting on sunscreen) as norms.

Because people express the same norm in linguistic variants (e.g., read, read a book, read books), two research assistants inspected the typed norms for each of the 8 contexts and assigned the same norm identifier to responses with identical or highly similar meaning within context and norm type (interrater agreement > .80%). Overall, participants generated between 3.7 and 7.4 norms per context, with prohibitions eliciting on average 1.5 fewer norms than prescriptions (Figure 2). We computed consensus scores for each generated norm (how many participants wrote down that norm for the given context), and we rank ordered the top 10 most consensual norms for each scene and norm type.

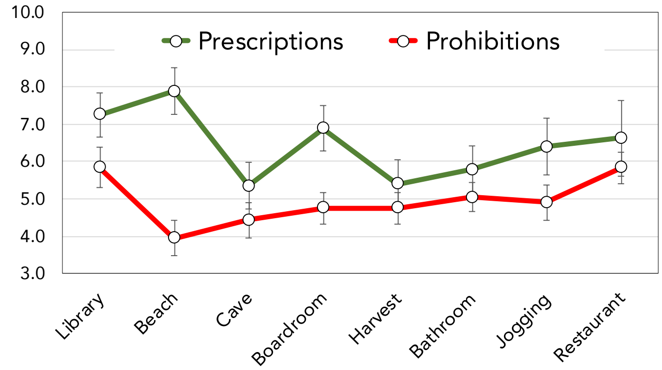

Figure 2. Average number of prescription and prohibition norms that participants generated for eight contexts (time limit: 60 seconds).

As Figure 3 illustrates, prescriptions displayed higher consensus than prohibitions, and overall consensus declined rapidly. This declining consensus is almost certainly a result of the unconstrained nature of the task, in which people are free to report a wide range of applicable norms for each context. In fact, we will see in the second paradigm that these consensus values are likely to be very conservative estimates of true consensus.

Of great interest was the degree of context-specificity of norms across the eight contexts. For each context, we determined how often one of its 10 most consensual norms was also mentioned in one or more other contexts as a top-10 norm. For prohibitions, 37 out of the 80 norms across eight contexts (46%) were unique—they appeared in only a single context. Of the remaining norms, 16 appeared in one other context (8 pairs), 12 in two others (4 triplets), two in three other contexts (2 quadruplets), and one in four other contexts. The average number of other contexts in which any given prohibition reappeared was 1.05 (possible range 0 to 7). For prescriptions, context-specificity was even higher, with 70 out of the 80 generated norms (87.5%) appearing uniquely in only a single context. The remaining prescriptions appeared in one other context (5 pairs), and the average number of contexts in which any given prescription reappeared was 0.125.

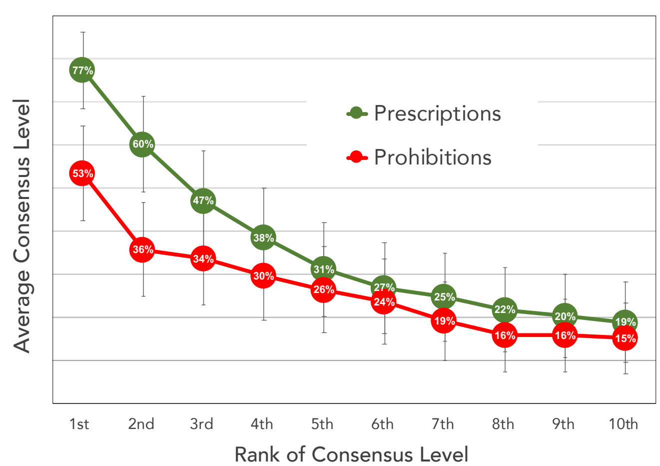

Figure 3. Norm generation consensus (with standard errors) for prescriptions and prohibitions, averaged across eight contexts and ordered by level of consensus. Consensus is the percentage of participants who wrote down the same norm for a given context.

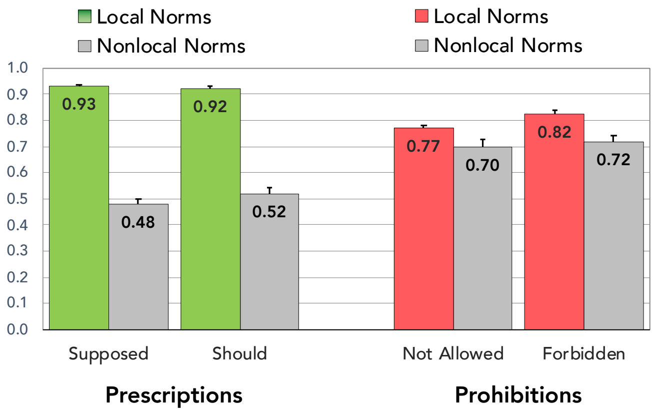

Figure 4. Proportion of participants who detected local (context-specific) and nonlocal norms as applicable to a particular context, averaged across four contexts. Drawn bars are +/- 1 standard error.

B. *Norm-Detection Paradigm*

The second experimental paradigm presents participants with one context at a time (activated by a picture) and introduces a series of actions that, in this context, may or may not be prescribed (in one experimental condition) or prohibited (in another condition). Participants indicate by pressing a Yes or No key whether a given action is indeed prescribed/ prohibited. For each context, participants see 14 actions: half are top-7 norms generated in the previous study for that context ("local" norms) and the other half are top-7 norms from other contexts ("nonlocal"). The proportion of people who designate a given norm as applicable to a specific context indexes prevalence. The difference between these detection rates for local and nonlocal norms indexes context-sensitivity.

In an initial study ($n = 100$), we asked participants to detect prescription or prohibition norms for four contexts (each norm type was probed in multiple ways, such as "forbidden" and "not allowed" for prohibitions). As Figure 4 shows, people nearly unanimously recognized the local prescription norms as applicable to their context, indicating high levels of consensus; people rejected the nonlocal norms half of the time. For prohibitions, this local-nonlocal difference was statistically significant but much weaker. Local prohibitions were detected less reliably, and nonlocal prohibitions differentiated less clearly, than the corresponding prescriptions.

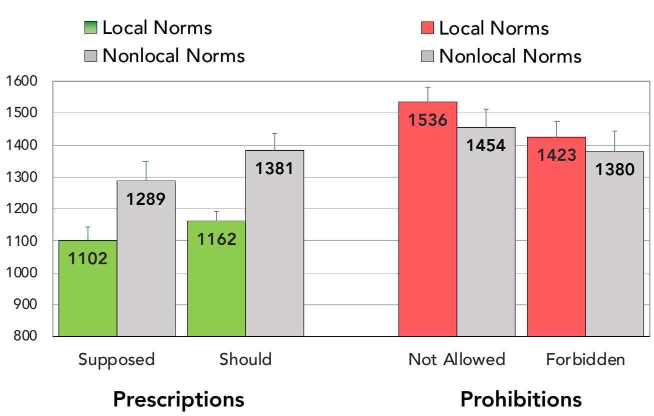

Figure 5. Reaction times in milliseconds for detecting local (context-specific) and nonlocal norms as being applicable to a particular context, averaged across four contexts. Drawn bars are +/- 1 standard error.

We also measured reaction times of recognizing local and nonlocal norms for their context. Figure 5 shows that people were fast in detecting context-specific prescription norms, and substantially faster than for nonlocal norms. This speed advantage was not present for prohibition norms. People were generally much slower at detecting prohibitions and no faster, even slightly slower, for local than nonlocal prohibitions.

To summarize, we presented two experimental paradigms to identify norms that a community perceives to be prevalent and applicable to particular contexts. Though consensus in the generation paradigm was high for only some norms, consensus in the detection paradigm was high throughout. Context specificity was very strong in both paradigms. Significant differences between norm types emerged. Prescriptions were generated in higher numbers, displayed greater consensus, showed higher context sensitivity, and were activated faster than prohibitions. Prescriptions thus appear to be readily on people's mind when they encounter a given context. Applied to robot design, we may conclude that a robot must know the prescriptions for any context it finds itself in, because these norms could efficiently guide the robot's behavior and because people would immediately think of such prescriptions (and impose them on the robot) when entering the context.

C. *Graded Deontic Force*

We also briefly mention our ongoing efforts to develop reliable and intuitive measures of deontic force. Several studies found that people agree substantially on a vocabulary of deontic force—a set of ordered terms that denote weak to strong prescriptions and weak to strong prohibitions [34] (see Figure 6). These linguistic markers could be used in norm generation and detection paradigms to identify norms of particular deontic force, such as all actions in a certain context that are "required" ($\mathcal{D}^f$ high) vs. "suggested" ($\mathcal{D}^f$ low). New work is also needed that relates degrees of deontic force to perceived prevalence, consensus, and speed of activation.

We should close by emphasizing that, for any given robot deployment, the introduced experimental paradigms are not the only sources of the relevant community's norms. For example, to develop a nurse assistant robot, one would also consult nursing training materials, hospital manuals, and interview various stakeholders. The goal here is not a complete and final norm catalogue but a *Starting Norm Base*. Refinements are discussed in the following sections.

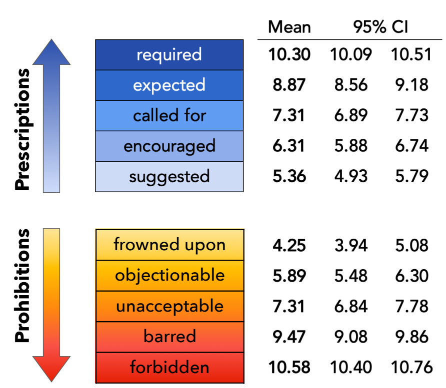

Figure 6. Vocabulary of graded deontic force in prescriptions (top 5) and prohibitions (bottom 5), with means and 95% confidence intervals around those means on scales of rated strength of expression.

VI. Norm Implementation

We now sketch an implementation of a norm-competent robot. We use a simple Markov Decision Process (MDP) as the starting point and then expand it to support norm competence.

A. *Selfish Goal Pursuit*

A Markov Decision Process (MDP) consists of a set of states, S, a set of actions, A, and a transition function, which describes the probabilities of the agent moving from any one state, s, to any other state, s', by taking action a: $P_a = \Pr(s'|s,a)$. These transitions come with rewards specified in a reward function, $R_a(s,s')$. A policy π is then sought that tells the agent what action to take in any given state so that the resulting action path through the state space maximizes the agent's cumulative rewards (sometimes discounted over time). We could use an MDP to have a robot perform a certain task optimally by providing it with a set of actions, which put it in a set of states, and we set up the rewards such that the robot always ends up in state s_g, the state in which the task is accomplished. Optimality might mean that it ends up in s_g in the shortest amount of time, on the shortest path, with the least amount of energy use. A strict MDP robot is the ultimate rationally selfish agent. It moves through state space by choosing actions that maximize its rewards, and the value of each action a is entirely determined by the reward of the state transition $s \rightarrow s'$ that the action achieves. For such an agent, the end entirely justifies the means.

B. *Norm-Competent Goal Pursuit*

A norm-competent robot is not merely selfish [15]. Norms give value to actions above and beyond the rewards those actions confer on the individual agent. For example, norms favor actions that have positive outcomes for the community, even if they come with a cost for the individual—that is, norms are prosocial. Norms also favor actions that other community members have performed in the past—that is, norms maintain customs and traditions. Such a consideration of other people's history directly violates the Markov property of an MDP, in which decisions are solely based on the system's current state. Building the agent's or others' past actions into the state representation would quickly lead to computational explosion.

To accommodate norm competence the robot's decision making system would have to decompose the reward function so that actions would have value irrespective of the states they bring about. In some models, norms are leveraged to prune the state space by ignoring states that violate known norms [35]. However, if potential states S' caused by A are unknown, the system may be paralyzed or make many mistakes. By contrast, because norms specifically instruct the agent to perform *actions,* a norm-competent decision maker can choose a norm-guided action even when the resulting s' is unknown and the agent never received any reward for performing a. That is, the agent trusts the social community that compliance with a norm will lead to an acceptable (though unknown) outcome.

Norms can also add value to specific variants of actions whose outcomes may (at least at first glance) be equally rewarding. For example, delivering information politely is more valued than delivering it bluntly, even if in both cases the resulting outcome of "information delivered" (and its reward) are identical. One might counter by saying that the recipient of politely delivered information will be happier than the recipient of bluntly delivered information, which changes the reward function. However, such an approach not only presupposes the distinction between psychological and world states (a nontrivial addition), it also assumes that the reward function would be tuned to others' psychological states, another nontrivial addition. Following a norm system does not require such tuning because norm-favored actions have value even without knowledge of their valued consequences.

People may expect even more sophistication from a norm-competent robot: Such a robot would not only avoid violating norms but would know when violating a norm is acceptable—e.g., when a moderate prohibition may be violated to comply with a strong prescription (e.g., lying to save someone's life). This demand for norm tradeoffs requires the robot to take into account the graded deontic force of all norms relevant to the context. The robot would choose an action that achieves the set goal s_g but would do so with minimal norm violation costs [28], [36]. The system may still have a hard threshold, such as $D^f < 0.5$, below which the robot needs to consult with a supervisor before making any norm tradeoffs.

C. *Reconciling Goals and Norms*

A norm-competent robot relying on an enriched MDP must also meet another challenge: the tension between being optimal in a selfish goal-pursuing sense (e.g., fast and efficient) and being socially appropriate in a norm-abiding sense (e.g., polite, fair). To resolve such conflicts, the value of goals and norms must be commensurable in some way. We know that people somehow find a common scale on which they perform value trade-offs and therefore resolve goal-norm conflicts; how could a robot acquire such a common scale?

One approach is to build norm considerations into the reward function and thus measure them in units of reward. However, rewards for efficient goal pursuit are often straightforwardly measurable (e.g., time or energy saved) whereas rewards for norm compliance are often merely represented (e.g., an action has symbolic value, even if one has never observed that it actually provided value to someone else). Moreover, rewards for goal pursuit stem from stationary distributions, where, say, the meaning of saving 1 kwh is knowable and does not change. By contrast, rewards for norm compliance are nonstationary: they are often underspecified, may be reevaluated in novel situations, and may change over time in unexpected ways. The agent must be able to adjust to such changes without inflating state space and falling into computational intractability. Even if it can, that still leaves the problem of commensurable reward assignments unsolved.

We suggest the opposite approach: to subsume rational goal efficiency considerations under the normative framework and learn from community members how optimal efficiency and social appropriateness are balanced against each other. Thus, the common currency is deontic force. Efficiency considerations (e.g., of time or energy use) can be framed as normative demands on the robot's behavior, and each demand has a certain degree of deontic force—how important it is for the robot to be fast or to use limited amounts of energy, *compared to* being polite, interrupting the main task to answer a question, etc.

If deontic force D^f [-1; 1] is the degree to which a norm imposes itself on an action, then "Be efficient" could be one of these norms (e.g., with $D^f = 0.3$, a mild prescription). Planning then becomes an attempt to optimally satisfy as many norms as possible (including the standard goals of any MDP), but with keen attention to the norms' deontic force and the goal to minimize norm-violation costs. In this process, however, the agent would not receive rewards for merely complying with norms, because we need to avoid accrual of infinite rewards for the countless prohibitions the agent is not violating (e.g., the countless ways to crash into objects); it would incur costs by violating prohibitions or failing to obey prescriptions.

D. *The Starting Norm Base (SNB)*

Our framework incorporates empirically identified norms in the Starting Norm Base (SNB), but the SNB must be updated in dynamic learning processes throughout evaluation and deployment. We now describe some of the properties of an SNB and likely challenges in its implementation.

The SNB can be thought of as a cube of contexts C × actions $\mathcal{A}$ × deontic operators $\mathcal{D}^f$, where each cell in the cube is a norm $\mathcal{N}_k$ that maps a particular action to a particular context under a particular deontic force (see Fig. 7):

$$\mathcal{N}_k = C_i \longmapsto \mathcal{D}^f(\mathcal{A}_j) \quad (2)$$

We define a *Norm Set* for a given context C_i as one slice of the cube, containing norms that assign deontic force values to a subset of actions in the robot's repertoire. Numerous other actions in the repertoire will remain unmapped and therefore be treated as irrelevant for the given context. This notable parsimony is modeled after the human experimental results (see Section V) in which a limited set of relatively unique action norms was activated by specific contexts. In the robot, each context-specific Norm Set can be implemented as a separate MDP, following a recently proposed modular approach [37]. During action planning, the robot will navigate through a sequence of contexts, so the modular MDPs may be organized hierarchically as Abstract MDPs [38].

Because activating context-specific norm sets is an essential element in norm competence, the robot must identify the context it is in. Thus, one of the major challenges of implementing a norm-competent robot is to equip it with context identification capacities discussed next.

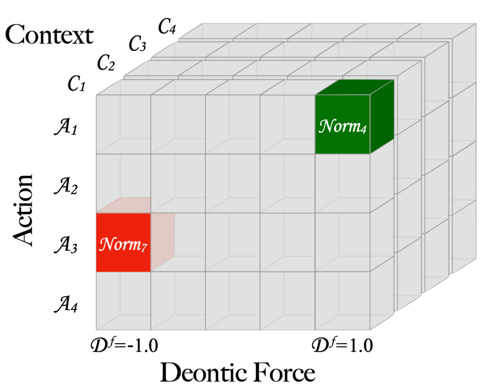

Figure 7. Simplified norm cube: Each cell displays a norm $\mathcal{N}_k$ that maps a particular action $\mathcal{A}_j$ to a particular context C_i under a deontic force $\mathcal{D}^f$. In reality, $\mathcal{D}^f$ values lie on a continuum, and many more actions and contexts exist.

E. *Context Identification*

To start out, we define a dozen or so broader contexts (e.g., nurse break room, hallway, patient room) and enable the robot to infer these contexts from photos or videos, which may be accomplished by the higher layers of a Convolutional Neural Network [39]. Such initial context categories will then activate the context-relevant norm set from the SNB. However, a deeper analysis of specific scene features and objects is necessary to distinguish variants of these broader contexts, such as *entering-patient-room-when-patient-alone* vs. *...-when-doctor-present* vs. *…-when-visitor-present*. These variants will share many elements in their respective norm sets but also activate some distinct norms. A number of the context variants, their distinguishing features, and their specific norm sets can be predefined and added to the Starting Norm Base. However, the robot will have to learn new context variants as well, especially as it acts in those contexts.

For example, for a nurse assistant robot, the following norm will be useful: "When in a patient's room (C_5), be quiet ($\mathcal{A}_{10}$)." But later the robot may learn several new norms, such as "When in a patient's room (C_5) & the patient asks a question (C_{10}), answer the question ($\mathcal{A}_{15}$)." Or, "When in a patient's room (C_5) & delivering medication (C_8), announce yourself ($\mathcal{A}_{20}$)." There are numerous formal ways to integrate such superficially contradictory norms, but they all come down to identifying context variants—here, variants of being in the patient's room, which can be represented as conjunctions of features or as nested conditionals. Over time, the robot will accumulate a context catalogue, greatly expanding and refining its initial context categories.

VII. Learning and Evaluation

The last phase of the proposed methodology of designing norm-competent robots is to have users interact with an SNB-equipped robot and help it advance, through guided feedback, toward a Refined Norm Base (RNB). These interactions will take place in the earlier mentioned *Norm-Learning Games*—human-robot interaction studies that test and improve the system's norm competence. Initially these experiments will be text-based and have constrained communicative channels, but we are also creating a more open-ended virtual-reality format.

In the Norm-Learning Game (NLG), a robot equipped with a predefined SNB proposes an action plan for a task it is assigned, and human teachers (sampled from the community of deployment) give the robot feedback about the normative appropriateness of its action proposals (Fig. 8). Because the interaction takes place in safe virtual space (text-based or VR), the robot can afford to make mistakes, and the human teachers correct them without fear of harm. This safe setting is important for mastering prohibitions, which are difficult to learn from observation (a norm-compliant community will not commit many prohibition violations), and human teachers are more likely to teach prescriptions than prohibitions (as results in Section V. suggest). The safe offline setting also allows the robot to analyze patterns of feedback from the entire set of human teachers, point out inconsistencies it noticed, and ask for clarification. The NLG may thus provide human teachers with opportunities to reflect on limitations, biases, or contradictions in their own norm system and perhaps bring those up for discussion in the community.

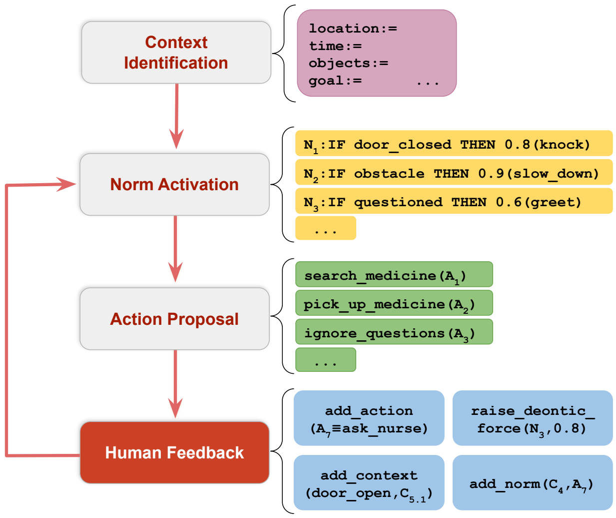

Figure 8. The main components of the Norm Learning Game

There are at least three mistakes a robot can make in the NLG. First, it may *misidentify the context* it is in or fail to distinguish between two variants of a context. To correct this error, the teacher can label the current context and perhaps identify diagnostic features of this context; or the robot can learn to associate the newly learned context label with features present at the time. Second, the robot may activate an *incomplete norm set* for a given context. To correct this error, the teacher can tell the robot about additional norm-governed actions that apply to the context (e.g., "When picking up pain medication you need to ask for the chief physician's permission"). Third, the robot may assign an *incorrect deontic force* to an activated norm. Such an error will be detected when the robot prioritizes one action (e.g., entering a patient's room when the door is open) while the relevant norm set contains a higher-priority action (e.g., self-announcing). The human teacher will provide at least ordinal information (e.g., suggesting that the proposed action is worse than the alternative one) and might even express specific deontic force values within a vocabulary of graded terms (see Fig. 6). In each case, the robot learns to adjust force values while keeping both actions (the proposed and the alternative) in its norm set for the given context. (Note that prioritization is often akin to "do X before Y," which will require an expressive form of action representation such as linear temporal logic [36].)

The initial simplified version of the NLG contains three constraints. First, to simplify context detection (e.g., from visual features), the teacher verbally places the robot in a particular starting context. A nurse assistant robot may be told, "You are at the front desk and the supervising nurse asks you…," followed by a specific task assignment, "…to deliver Ms. Jones' medication to her room." By directly providing the initial context, the teacher and robot learner can focus on refining the robot's norm sets. Second, to simplify learning new norms, the robot's total action repertoire is limited, and the user append actions to a norm set by selecting ones from the repertoire. Third, to limit the demands on the system's natural language processing abilities, the human teacher's feedback will come in the form of specific feedback types—e.g., introducing a new context variant, adding a norm to a context, or changing a norm's deontic force.

In subsequent, more open-ended forms of the NLG, the robot will combine multi-modal input into inferences about the current context it is in; it will learn new action concepts that were not even in its action repertoire (e.g., composed of familiar primitives [39]); and it will interpret the user's feedback as the appropriate kind of correction—e.g., of context identification ("No, we aren't in the patient room"), norm addition ("Here, you also have to…"), or deontic force adjustment ("It's more important to…").

A text-based NLG, though useful, does not fully capture the complex visuospatial information relevant to context and norm identification. To provide the human teacher with environments that are still safe but contain rich features and are more scalable, we are also developing a Virtual Reality (VR) version of the NLG. Game engines like Unity3D [40] allow us to implement high-quality physics simulations and use object meshes for realistic scenes, such as hospital hallways and patient rooms. Initially, people will be in the observer perspective, watching a robot perform various actions in the scenes. Using VR-Head-Mounted Displays (VR-HMDs), people can freely move their head and interact with the 3D environment and recognize subtle context shifts. The robot's action is simulated with VR-robot middleware like ROS Reality [41], and its "cognition" (norm sets, action repertoire) can be ported from the initial text-based learning game to the 3D environments. Expanded NLP components will allow the human teacher to give the robot feedback with naturally expressed language. VR also enables expanded context-detection challenges, such that virtually presented scene features become inputs to the robot's own ability to identify the context it is in.

VIII. Summary

After defining and conceptualizing social-moral norms, we proposed a methodology for designing a norm-competent robot. Using empirical results on central properties of human norm structures, we introduced one possible computational implementation: a hybrid model of predefined and learned norms and goals. Significantly expanding an MDP approach, we sketched a robot with capacities for context detection, norm activation, and action guidance by norm constraints. We suggested a common currency for goal pursuit and norm compliance and outlined a Norm Learning Game that allows the robot to refine its starting norm base through iterative community feedback in a safe VR environment. A full test of this model is yet to come, but we believe to have provided a cognitively realistic, computationally tractable, and practically applicable model of robot norm competence.

References

[1] M. Hechter and K.-D. Opp, Eds., *Social norms*. New York, NY: Russell Sage Foundation, 2001.

[2] B. F. Malle, P. Bello, and M. Scheutz, "Requirements for an artificial agent with norm competence," in *Proceedings of 2nd ACM conference on AI and Ethics (AIES'19)*, New York, NY, 2019, pp. 21–27, doi: 10.1145/3306618.3314252.

[3] E. Krupka and R. A. Weber, "The focusing and informational effects of norms on pro-social behavior," *Journal of Economic Psychology*, vol. 30, no. 3, pp. 307–320, Jun. 2009, doi: 10.1016/j.joep.2008.11.005.

[4] R. B. Cialdini, C. A. Kallgren, and R. R. Reno, "A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior," in *Advances in Experimental*

Social Psychology, vol. 24, M. P. Zanna, Ed. San Diego, CA: Academic Press, 1991, pp. 201–234.

[5] H. Aarts and A. Dijksterhuis, “The silence of the library: Environment, situational norm, and social behavior,” *Journal of Personality and Social Psychology*, vol. 84, no. 1, pp. 18–28, Jan. 2003, doi: 10.1037/0022-3514.84.1.18.

[6] R. J. Janoff-Bulman, S. Sheikh, and S. Hepp, “Proscriptive versus prescriptive morality: Two faces of moral regulation.,” *Journal of Personality and Social Psychology*, vol. 96, no. 3, pp. 521–537, 2009, doi: 10.1037/a0013779.

[7] P. McNamara, “Deontic logic,” in *Handbook of the History of Logic*, vol. 7, D. M. Gabbay and J. Woods, Eds. North-Holland, 2006, pp. 197–288.

[8] M. Anderson and S. L. Anderson, Eds., *Machine ethics*. New York, NY: Cambridge University Press, 2011.

[9] T. Ågotnes, W. V. D. Hoek, J. A. Rodriguez-Aguilar, C. Sierra, and M. Wooldridge, “On the logic of normative systems,” in *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI ’07)*, M. Veloso, Ed. Palo Alto, CA: AAAI Press, 2007, pp. 1181–1186.

[10] G. Governatori and A. Rotolo, “BIO logical agents: Norms, beliefs, intentions in defeasible logic,” in *Normative Multi-agent Systems*, G. Boella, L. van der Torre, and H. Verhagen, Eds. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2007.

[11] L. M. Pereira and A. Saptawijaya, “Modelling morality with prospective logic,” *International Journal of Reasoning-based Intelligent Systems*, vol. 1, no. 3/4, pp. 209–221, 2009, doi: 10.1504/IJRIS.2009.028020.

[12] M. J. Wooldridge, *An introduction to multiagent systems*, 2nd ed. Chichester, U.K: John Wiley & Sons, 2009.

[13] J. Broersen, M. Dastani, J. Hulstijn, Z. Huang, and L. van der Torre, “The BOID Architecture: Conflicts between beliefs, obligations, intentions and desires,” in *Proceedings of the Fifth International Conference on Autonomous Agents*, New York, NY, 2001, pp. 9–16, doi: 10.1145/375735.375766.

[14] R. Conte, G. Andrighetto, and M. Campenni, *Minding norms: Mechanisms and dynamics of social order in agent societies*. New York, NY: Oxford University Press USA, 2013.

[15] G. Andrighetto, D. Villatoro, and R. Conte, “Norm internalization in artificial societies,” *AI Communications*, vol. 23, no. 4, pp. 325–339, Dec. 2010.

[16] G. Andrighetto, C. Castelfranchi, E. Mayor, J. McBreen, M. Lopez-Sanchez, and S. Parsons, “(Social) norm dynamics,” in *Normative multi-agent systems*, vol. 4, G. Andrighetto, G. Governatori, P. Noriega, and L. W. N. van der Torre, Eds. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013, pp. 135–170.

[17] K. Arkoudas, S. Bringsjord, and P. Bello, “Toward ethical robots via mechanized deontic logic,” in *Machine Ethics: Papers from the AAAI Fall Symposium 2005*, vol. FS-05-06, 2005, pp. 17–23.

[18] D. Abel, J. MacGlashan, and M. L. Littman, “Reinforcement learning as a framework for ethical decision making,” 2016.

[19] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, “Planning and acting in partially observable stochastic domains,” *Artificial Intelligence*, vol. 101, no. 1, pp. 99–134, May 1998, doi: 10.1016/S0004-3702(98)00023-X.

[20] T. Arnold, D. Kasenberg, and M. Scheutz, “Value alignment or misalignment – What will keep systems accountable?,” in *The Workshops of the Thirty-First AAAI Conference on Artificial Intelligence: Technical Reports, WS-17-02: AI, Ethics, and Society*, Palo Alto, CA: The AAAI Press, 2017, pp. 81–88.

[21] D. Hadfield-Menell, S. J. Russell, P. Abbeel, and A. Dragan, “Cooperative inverse reinforcement learning,” in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 3909–3917.

[22] V. Sarathy, M. Scheutz, Y. N. Kenett, M. M. Allaham, J. L. Austerweil, and B. F. Malle, “Mental representations and computational modeling of context-specific human norm systems.,” in *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, Austin, TX: Cognitive Science Society, 2017, pp. 1035–1040.

[23] D. Whitney, E. Rosen, J. MacGlashan, L. L. S. Wong, and S. Tellex, “Reducing errors in object-fetching interactions through social feedback,” in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 1006–1013.

[24] C. Bicchieri, *The grammar of society: The nature and dynamics of social norms*. New York, NY: Cambridge University Press, 2006.

[25] B. F. Malle, M. Scheutz, and J. L. Austerweil, “Networks of social and moral norms in human and robot agents,” in *A World with Robots: International Conference on Robot Ethics: ICRE 2015*, M. I. Aldinhas Ferreira, J. Silva Sequeira, M. O. Tokhi, E. E. Kadar, and G. S. Virk, Eds. Cham, Switzerland: Springer International Publishing, 2017, pp. 3–17.

[26] M. Nickles, “Towards a logic of graded normativity and norm adherence,” in *Normative Multi-agent Systems: Dagstuhl Seminar Proceedings*, Dagstuhl, Germany, 2007, [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2007/926.

[28] K. J. Holyoak and D. Powell, “Deontological coherence: A framework for commonsense moral reasoning,” *Psychological Bulletin*, vol. 142, no. 11, pp. 1179–1203, Nov. 2016, doi: 10.1037/bul0000075.

[29] G. Shafer, *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press, 1976.

[30] V. Sarathy, M. Scheutz, and B. F. Malle, “Learning behavioral norms in uncertain and changing contexts,” in *Proceedings of the 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2017, pp. 301–306.

[31] Y. N. Kenett, M. M. Allaham, J. L. Austerweil, and B. F. Malle, “The norm fluency task: Unveiling the properties of norm representation.,” in *Poster presented at the 57th Annual Meeting of the Psychonomic Society, Boston, MA, November 2016*, Boston, MA, 2016.

[32] B. F. Malle and J. L. Austerweil, “From value alignment to norm competence: How robots and AI will conform to human society. Paper presented at,” presented at the American Psychological Association’s conference on Technology, Mind and Society, Washington, DC, Apr. 2018.

[33] B. Tversky and K. Hemenway, “Categories of environmental scenes,” *Cognitive Psychology*, vol. 15, no. 1, pp. 121–149, Jan. 1983, doi: 10.1016/0010-0285(83)90006-3.

[34] B. F. Malle, “Graded representations of norm strength,” in *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society.*, Cognitive Science Society, 2020.

[35] M. S. Fagundes, S. Ossowski, J. Cerquides, and P. Noriega, “Design and evaluation of norm-aware agents based on Normative Markov Decision Processes,” *International Journal of Approximate Reasoning*, vol. 78, pp. 33–61, Nov. 2016, doi: 10.1016/j.ijar.2016.06.005.

[36] D. Kasenberg and M. Scheutz, “Norm conflict resolution in stochastic domains,” in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, 2018, pp. 85–92.

[37] V. Krishnamoorthy, W. Luo, M. Lewis, and K. Sycara, “A computational framework for integrating task planning and norm aware reasoning for social robots,” in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Aug. 2018, pp. 282–287, doi: 10.1109/ROMAN.2018.8525577.

[38] N. Gopalan *et al.*, “Planning with Abstract Markov Decision Processes,” in *Twenty-Seventh International Conference on Automated Planning and Scheduling (IACAPS 2017)*, Jun. 2017, pp. 480–488, Accessed: Apr. 13, 2020. [Online]. Available: https://aaai.org/ocs/index.php/ICAPS/ICAPS17/paper/view/15759.

[39] Z. Kira, W. Li, R. Allen, and A. R. Wagner, “Leveraging deep learning for spatio-temporal understanding of everyday environments,” 2016.

[40] Unity Technologies, *Unity*. 2020.

[41] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, “Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality,” in *Robotics Research: The 18th International Symposium ISRR*, N. M. Amato, G. Hager, S. Thomas, and M. Torres-Torriti, Eds. Cham, Switzerland: Springer Nature, 2019, pp. 335–350.