A Norm Emergence Framework
for Normative MAS – Position Paper

Andreasa Morris-Martin$^{(\boxtimes)}$, Marina De Vos, and Julian Padget

University of Bath, Bath, UK
{a.l.morris.martin,m.d.vos,j.a.padget}@bath.ac.uk

Abstract. Norm emergence is typically studied in the context of multiagent systems (MAS) where norms are implicit, and participating agents use simplistic decision-making mechanisms. These implicit norms are usually unconsciously shared and adopted through agent interaction. A norm is deemed to have emerged when a threshold or predetermined percentage of agents follow the "norm". Conversely, in normative MAS, norms are typically explicit and agents deliberately share norms through communication or are informed about norms by an authority, following which an agent decides whether to adopt the norm or not. The decision to adopt a norm by the agent can happen immediately after recognition or when an applicable situation arises. In this paper, we make the case that, similarly, a norm has emerged in a normative MAS when a percentage of agents adopt the norm. Furthermore, we posit that agents themselves can and should be involved in norm synthesis, and hence influence the norms governing the MAS, in line with Ostrom's eight principles. Consequently, we put forward a framework for the emergence of norms within a normative MAS, that allows participating agents to propose/request changes to the normative system, while special-purpose synthesizer agents formulate new norms or revisions in response to these requests. Synthesizers must collectively agree that the new norm or norm revision should proceed, and then finally be approved by an "Oracle". The normative system is then modified to incorporate the norm.

Keywords: Norm synthesis · Synthesizer agents · Normative MAS · Normative system

1 Introduction

Multiagent systems (MAS) enable participating agents to interact with each other and their environment to accomplish individual goals and collective goals. MAS utilise norms to encourage coordination and cooperation, and to avoid the occurrence of undesirable states. Hollander and Wu [17] define a normative MAS as a system which combines concepts of norms with explicit representations of normative information in order to provide a solution to problems relating to openness in MAS. In contrast, in some MAS norms may not be considered at

all or the concept of norms is present but only as implicit representations of normative information.

The normative system, the set of norms in the MAS, is normally created as a result of the process of norm synthesis. Norm synthesis is the creation and updating of norms to avoid conflict situations – unwanted states in the MAS. Norm synthesis can be offline, which occurs mostly during design time [22] or as a separate process outside the system governed by the norms [30], where norms are determined by the designer or other stakeholders. A normative system resulting from offline synthesis is typically fixed for the lifetime of the system. But, over time, with changing environments, the norms can become partly or wholly irrelevant. Consequently, it becomes necessary to update the normative system leading to the introduction of online norm synthesis.

Online norm synthesis occurs while the system is live, and (new/revised) norms are typically determined by a centralised mechanism with global knowledge and without any input from the participants [1,25]. There exists a subset of decentralised online norm synthesis mechanisms [14,15], that allow agents to submit their personal strategies, then the most popular strategy is incorporated into the normative system; a similar approach is undertaken in [36]. Additionally runtime norm revision approaches [10,11] and dynamic normative systems [18,20,21] present other avenues where normative systems can be changed while the system is live.

On the other hand, the introduction of norms into a MAS without an explicit normative system, is orchestrated by participating agents in simulation models of norm emergence [32,41,44] , where the norms are defined as the preferred action from a set of actions, and the norm is usually learnt through agent interactions. However these norms do not become part of an explicit normative system as with work by [14,15,36] that also involve individual agent strategies as norms.

In this paper, we introduce a framework that allows participating agents to affect the online norm synthesis process. During runtime, participating agents can identify the norms or situations that require normative regulation, as long as there are the necessary affordances [33] for participating agents to propose or request changes to the normative system. We propose synthesizer agents which, upon request from participating agents, synthesise norms for the MAS to meet the identified need. We believe that after we encode the synthesised norm into the normative system, and it is adopted by a sufficient number of agents, we may ultimately observe the emergence of synthesised norm(s).

The rest of the paper is structured as follows. Section 2 which situates the proposed framework by discussing the gap identified in the literature. Section 3 gives a high level description of the framework, after which we discuss the stages of the framework in more detail: Section 4 highlights the main contribution which defines the norm creation stage, Sect. 5 discusses norm propagation, then norm adoption in Sect. 6, finishing in Sect. 7, with norm emergence. The paper concludes in Sect. 8, where we briefly discuss how the framework can facilitate the emergence of norms in open normative MAS and allow for the normative

component to be changed in response to issues affecting agents participating in the MAS.

2   Research Context

Norms in the (MAS) literature are looked at from two distinct perspectives: norms as deontic concepts and norms as a preference behaviour. Each perspective implements the life cycle of norms differently, but can broadly be seen to follow similar stages. The norm life cycle defines the stages a norm goes through in its lifetime. The widely accepted norm life cycle [40], based on simulation studies has a three stage process: norm formation/creation, norm propagation and norm emergence. The literature also provides variations to the norm life cycle for example: Hollander and Wu [17] present a ten stage process which is encapsulated into three super-processes, and Savarimuthu and Cranefield [40] and Mahmoud et al. [24] present separate five stage processes. Frantz and Pigozzi [12], after an analysis of existing life cycle models present a refined evolutionary, rather than end to end model, with five stages namely: Creation, Transmission, Identification, Internalization and Forgetting. They consider Evolution and Emergence as phenomena rather than explicit stages in the model [12].

A closer look at the norm life cycle enables us to better understand the different perspectives on norms. We consider the widely accepted three stage process in our analysis.

2.1   The Prescriptive Perspective and the Norm Life Cycle

The first perspective, norms as deontic concepts, views norms as permissions or prohibitions and obligations, or commitments and is referred to as *norms as prescriptions* [7], or the *prescriptive approach* [39]. Here we refer to it as the *prescriptive perspective*. Its literature uses an explicit representation of norms, typically internal as agent beliefs, but can also be external and referenceable.

The norm life cycle in a society with explicit norm representation exhibits the following stages: (i) creation: a norm is introduced into the society, source often unknown, (ii) identification: agents become aware of the norm and update their beliefs, (iii) adoption: agents reason about adopting the norm, (iv) propagation: agents deliberately, or possibly unintentionally, inform others of the norm, (v) emergence: not usually considered, but inferrable when a predetermined threshold of agents have adopted the norm.

The prescriptive perspective branch of the literature presents the norm life cycle with a focus on norm identification and adoption. It also considers the synthesis of norms as an alternative to the norm creation stage, as it defines the norms that will regulate the behaviour of the agents within the MAS. Norm synthesis is geared towards creating norms to avoid conflicts and can occur both online and offline. Recently in normative MAS, online norm synthesis is becoming a popular topic, as it allows for the introduction of norms into the MAS at runtime to meet changing circumstances. Another benefit of online

norm synthesis is that it can cater for observed states, or agent capabilities that had not been conceived at the time the MAS was designed [16]. The online norm synthesis mechanisms cited [1,6,25–28] here operate by monitoring the state of the environment and proposing norms to resolve conflicts, that can or have occurred, to prevent their occurrence in the future. For example, [26–28] do so by prohibiting the action of an agent in the timestep before the conflict. The majority of online methods cited above utilise a centralised mechanism with global knowledge, and to the best of our knowledge only AOCMAS [6] employs a decentralised mechanism.

AOCMAS is a two-level distributed MAS architecture that equips an organisation with adaptive capabilities. At one level there are domain agents who are concerned with their individual goals, and above them are assistant agents who are concerned with the organisation's goals and help to facilitate the adaption. Assistants are responsible for or oversee a cluster of domain agents. Assistant agents partially observe the state of the organisation at runtime and propose regulations, rules or legal norms, for problems identified. Before proposing new solutions, assistants check if an existing stored solution is applicable using case-based reasoning (CBR). If none is found, they propose a set of regulations, and each regulation in the set is voted on by all assistants. The regulations with the majority vote become the new set of regulations, that over time are evaluated for effectiveness and may be updated or removed.

Another branch of online norm synthesis mechanisms [14,15,36] incorporates the strategies of individual agents into the normative system through voting, utilising a decentralised mechanism. In [14,15] the most popular strategy becomes part of the normative system, whereas in [36] the most popular strategy and associated sanction is selected, then put to a vote, before it can become part of the normative system. The norms derived here are geared at providing agents with a consistent strategy from among alternatives, presumably operationalised as obligations rather than prohibitions.

We note that the literature on norm synthesis, which is applicable to the norm creation stage of the norm life cycle, fails to investigate the impact of the remainder of the norm life cycle. Instead the focus is on determining an appropriate set of norms, after which it is assumed that either these norms will be adopted by agents or that the norms will be regimented. It is surprising that a disconnect exists within a single perspective of the norm literature, but research in different sections of the literature investigate various processes of the norm life cycle independently, with little attempt to combine or sequence the activities together. This gap, we believe, can readily be filled, by using norm synthesis in lieu of norm creation, while the other stages continue as normal in the prescriptive approach.

2.2 The Emergence Perspective and the Norm Life Cycle

The second perspective of norms is as a preference behaviour. This views norms as a predetermined or computed preference behaviour to execute from among a set of behaviours in a given situation. The preference behaviour perspective is

referred to as *norms as conventions* [7] or the *emergence approach* [39], here we refer to it as the *emergence perspective*. Its literature typically uses an implicit norm representation. However it is imperative to note that the deliberate communication of an agent's strategy to another agent would necessitate an explicit representation of the norm. Additionally [8,42] for example, enable agents to make explicit representations of social norms that have been identified via observation.

The norm life cycle in a society with implicit norms usually has these stages: (i) creation: the initial or predetermined strategy for an agent from a set of available actions, (ii) propagation: the sharing of an agent's strategy either unintentionally or deliberately, (iii) emergence: a percentage of agents follow this strategy for every instance of the triggering situation. The norm life cycle with implicit norms is usually seen in simulation models of norm emergence.

The emergence perspective literature is focused primarily on the study of norm emergence. Norm emergence is normally defined as the point in a MAS when a threshold or predetermined percentage of agents adopt a norm. There is generally no discussion of the activities that precede or follow this point in the state of the MAS. In [31] we present the notion that a refinement of norm emergence is needed, as it is difficult to fully understand the emergence of a norm without taking into account the preceding activities of norm creation and propagation. Therefore norm emergence is henceforward what is normally the norm life cycle, but including norm creation and propagation.

Utilising the interactions of agents at runtime as a basis for the creation of norms, as in norm emergence, is another potential solution for the source of norms in the prescriptive perspective, this is similar to the discussion from [12] of a second entry point to the life cycle model. Using norm emergence as a basis for norm creation is not an automatic fix however, as there are disadvantages to using norm emergence alone. Morales et al. [30] suggest that norm emergence is inappropriate to synthesise norms for MAS where there are numerous inter-dependent conflict situations. Additionally, norm emergence does not usually allow for the explicit representations of norms. This can be problematic, because when norms are implicit, there can be confusion among the agents about the prevailing norm, because different subsets of agents may adopt different norms. Additionally agents can have different interpretations about the prevailing norm(s) based on their beliefs [17], whereas [12] consider these different interpretations of norms as a form of norm change.

2.3   Research Opportunity

As we have alluded to earlier, norm emergence in its traditional sense is comparable to online norm synthesis, since both involve the creation of norms at runtime. In norm synthesis the creation of the norms is motivated by a single external entity: external in the sense that it does not involve participating agents, while in norm emergence it is achieved internally by the participating agents' behaviour. The creation of norms in norm emergence is naturally distributed and online, while norm synthesis is usually centralised and often times offline.

We submit that it can be beneficial to use participating agents within the MAS in the norm synthesis process, thereby taking the benefits of the distributed and online approach of norm emergence and adopting it into the norm synthesis process, to produce explicitly represented norms. The final product is an approach that facilitates the emergence of norms in a normative MAS system. Building on the concept of distributed agents having the ability to create norms, we put forward a framework that allows participating agents to request the inclusion/modification of norms in the normative system, based on their experiences within the MAS.

We situate this notion of participating agents contributing to the norms on two of Ostrom's [34] eight principles for governing the commons. Foremost, the need to allow participatory decision making, especially from those who are likely to be affected by the rules/decisions, and secondly the need to ensure the rules in place meet the needs of the local context, that is the participant's needs. We believe these two principles, which are intended to aid proper governance of the commons, are applicable to defining rules for normative MAS with different contexts.

Haynes *et al.* [16] examine how emergent behaviours in a system can be beneficial to the system and should be encouraged and spread, while non-beneficial emergent behaviour should be discouraged. We suggest that beneficial emergent behaviour can further be encoded explicitly as legal norms within the normative MAS. Therefore beneficial emergent behaviours can give rise to the emergence of obligation norms, permission norms and prohibition norms for alternative behaviours. Likewise, non-beneficial emergent behaviours could give rise to the emergence of prohibition norms, obligation norms to avoid certain states and revocation of permission norms.

Frantz et al. [13] examine how agents can derive normative information from their observations and present them as explicit prescriptions with a focus on social norms. In addition, [8,42] show how agents can identify the prevailing norms based on their observations and experiences. The preceding allow us to see the usefulness in equipping agents with the capacity to identify and formulate norms based on their interactions with the system. Therefore the inclusion of agents capable of identifying and formulating norms in a MAS supports the notion of agents that can synthesise different types of norms (for the MAS), a concept we believe presents an opportunity for future research. This is different from the approaches presented in [14,15,36], which allow the synthesis of norms for an action selection strategy from among existing options.

Note however that we do not expect participating agents to be able to properly synthesise norms for the MAS in our approach, because they would need access to domain knowledge and history of the MAS, which a participating agent would not normally have. Additionally, there is a need for higher minimum requirements on the cognitive abilities of participating agents that may make the MAS inaccessible to some agent types. Therefore, we propose agents communicate requests to synthesizer agents, who are responsible for a subset of agents within the MAS. Each synthesizer agent is capable of synthesising norms,

based on the request from the participating agents, and their (partial) knowledge of the domain and the environment. We contend that the request for norm change from a participating agent aligns with triggering the initiation of norm synthesis from the bottom up.
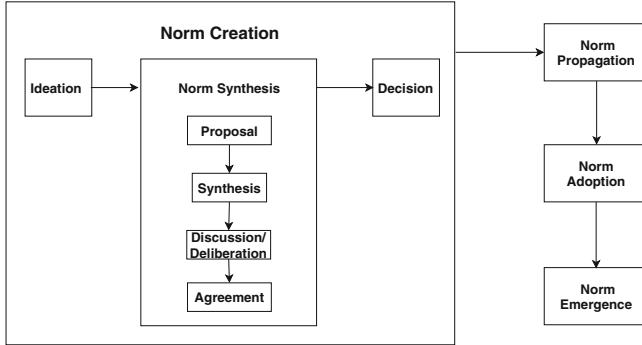

Fig. 1. A conceptualisation of norm emergence for normative MAS

3   Norm Emergence Framework

Agents operating in a MAS, like in human societies, could potentially determine the norms that would better regulate the society. In the emergence perspective literature, it is accepted that agents learn their behaviour from interacting with other agents. This means that the strategy of one agent or a set of agents can become the norm in the society. The usefulness of employing the concept of the agents determining the norm, as a potential answer to the source of the norm in the prescriptive perspective, is worth investigating, thereby allowing the norms of a system to be determined by the agents that participate in the system. [13] show how agents can prescribe social norms based on their observations, but this focuses on an agent's individual norms, rather than incorporating these into the normative system, on the other hand [14,15,36] demonstrate mechanisms to embed the social norm/strategies of agents into the normative system for governing the entire MAS.

The goal of our approach is to explore techniques for developing self-governing systems, in which agents participate in the revision of the norms that affect them. This would preclude the need for direct human intervention to synthesise norms for changing environments, or at the least, require minimal human intervention. The removal of human involvement though useful in this context unearths risks that must be considered in developing MAS. We provide a brief explanation in Sect. 4.3. Decision Stage.

We conceptualise a model that pulls together three research strands, namely: (i) the life cycle of norms in the prescriptive perspective (ii) the life cycle of norms in the emergence perspective, and (iii) the norm synthesis process, in order to construct a framework that facilitates the emergence of norms in a normative MAS. Thereby, we aim to identify the complementary activities – norm creation in emergence and norm identification and adoption in normative MAS and explicitly represented norms from norm synthesis, remove duplicate and unnecessary activities, and sequence them in a way that a new model is revealed. The resulting model enables us to define a norm emergence framework for normative MAS which we sketch in Fig. 2.

The model details the process of norms being synthesised, which begins with the initial recognition of the need for a new norm in the MAS by a participating agent. We conceive of two types of agents operating within the MAS: participant agents and synthesizer agents. The topology of the agents within the MAS will not be considered as we do not consider interactions within the MAS to be influenced by social distance or connections between participating agents. We will however aim for there to be a uniform distribution of agents to synthesizers based on the number of synthesizer agents in use.

Participating Agents. We put forward some assumptions about the agents that may participate in such a MAS, specifically: (i) the agent has an explicit internal representation of norms and some non-trivial cognitive or reasoning abilities, (ii) the agent considers norms in their action-selection and planning processes, and (iii) the agent is capable of perceiving a need for a norm change, either as a new or revised norm.

The agent's perception of the need for a norm change is predicated on agents considering norms when acting, and being able to observe the effect of their actions on the environment via action feedback. The participating agents in our model are inherently *normative agents* and rather *naive* because they assume that the norm(s) they enact are solely responsible for the feedback they receive.

Synthesizer Agents. The framework calls for a distinguished set of agents with partial perception of the MAS, that we refer to as synthesizers. Participating agents are assigned a synthesizer agent upon entry to the MAS. Synthesizer agents are agents designed with knowledge of the domain context: goals, actions, conflicting states, norms. They are capable of perceiving all the actions of the agents for which they are responsible and the environment state at any given time. Synthesizer agents in this model are inspired by the assistant agents in [6], which compute regulations utilising a partial perspective of the MAS, but do so after observing a problem, whereas synthesizer agents here await requests from participating agents. Additionally, assistants [6] vote on the regulations determining which shall be included, and similarly synthesizer agents must vote to decide whether or not to include all the norms proposed after discussions.

We note that synthesizer agents in our model can potentially be proactive as well. They can be proactive by examining traces and identifying when conflicting states occur, then proposing norms to avoid them in the future, without waiting on instruction from the agent to do so. Such an approach echoes elements of [26,

28–30]. Synthesizer agents could also have other functions within a MAS, such as being responsible for the enforcement of sanctions if violation occurs. We have however decided to limit the functions of the synthesizers in this framework to participating in the process of synthesising norms for the normative system only upon receipt of requests from agents. This we believe allows us to demonstrate norm synthesis initiated from the bottom-up which is different from existing research on norm synthesis.

Components of the Framework. The stages of the conceptual model of our framework are: (i) norm creation, (ii) norm propagation, (iii) norm adoption, and (iv) norm emergence as shown in Fig. 1 and Fig. 2.

Of the existing life cycle models, the stages of our conceptual framework are most closely aligned with the life cycle model of norm emergence in [12], with respect to the sequencing of stages but with some differences. Our norm creation stage encapsulates the norm creation activities of [12] that are shown as two entry points to their model at norm creation and identification. Our decision for one entry point is premised on utilising special-purpose synthesizer agents which encode the norms based on requests from individual agents. Therefore, though agent's experiences form the basis for changes to the normative system, these changes are managed by a group of domain agents, synthesizers, and changes must be approved before the rules/norms of the system are changed.

The activities of norm propagation are similar to that of transmission and norm adoption with identification, however norm emergence is represented as a specific stage in our model, while it is viewed as a phenomenon in [12]. Emergence is a specific stage in our framework because we consider it to be an activity that will be realised when a number of agents acting in a particular way exceeds a threshold.

We do not consider the forgetting of norms in our model. Our assumption is that modifications to the normative system could result in additions, modifications or deletions to the norms, which can be viewed as similar to the concept of forgetting being a byproduct of evolving norms in [12], though their lifecycle provides a specific stage to represent it. Additionally [12] discuss evolution as another phenomenon and though we do not consider it in our framework, it can be considered to exist in a similar way in our model, through the actions of the synthesizer agents on the set of norms.

The key difference we claim between our framework and the existing life cycle models, is that it builds a bridge from a theoretical understanding of the norm life cycle to a framework that operationalises the stages to bring about the emergence of norms in a normative MAS, utilising synthesizer agents. We discuss each of the stages of our conceptual model: norm creation, norm propagation, norm adoption, and norm emergence, in the remaining sections of the paper.

4   Norm Creation Stage

The creation stage comprises three sub-processes (i) ideation (ii) norm synthesis, and (iii) decision, depicted in Fig. 2, component 1. In the ideal scenario, at the
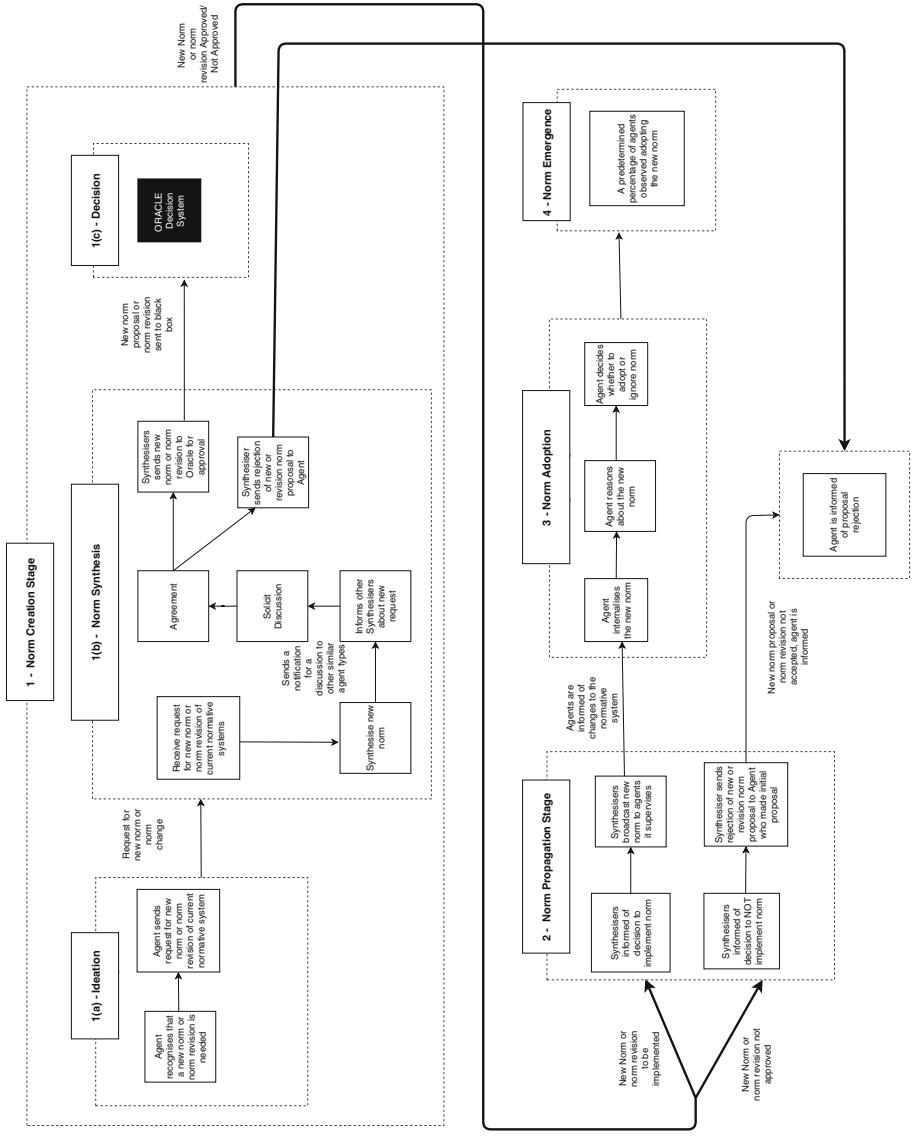
Fig. 2. Conceptual Model for norm emergence in normative MAS

end of creation, there is a norm that must be incorporated into the normative system. If not, the request would not have been approved, and thereafter the synthesizers and the initial requesting agent must be informed.

4.1   Ideation

The ideation process models the norm creation stage of agents in the emergence perspective, Fig. 2, component 1(a), by enabling the agent to request changes based on feedback from its interaction with the environment and other agents, similar to how an agent would learn a norm in the emergence perspective. The difference here is that the agent itself is not capable of making changes to the normative system, but can communicate with a synthesizer agent to initiate that change on their behalf.

A participating agent in a society may determine there is a need for a norm change: a new norm or the revision of an existing norm governing the MAS. Norm change requests are typically predicated on the following circumstances: (i) conflict situation/state arising from compliance with prevailing norms, (ii) reasoning that repeatedly determines that violation of current norms is a rational choice, (ii) a new norm or norm revision that can potentially bring about a better outcome for the agent or a better state in the MAS; we refer to this as an innovation norm, (iv) dissatisfaction with the current norms, e.g. prohibition of actions that can help agents achieve goals more efficiently.

An agent operating within the society may likely recognise one of the above situations developing long before an external observer can do so. The agent then informs their assigned synthesizer agent of their perceived need for a norm change. Each request must specify the context of the request, the reason for the request and the actual norm proposed, if the agent is capable of synthesising the norm. It may be useful to provide a template to the agents which specifies what needs to be included in the request.

Agents similar to those described in [13] are capable of encoding their observations as norm prescriptions, which would make them suitable to suggest actual norm revisions in a request to a synthesizer. This could be specifically useful for agents suggesting an innovation norm based on observations or experiences with the system.

Individual agents may or may not be able to determine what the new norm can be, since they have limited knowledge of the system. We could put more responsibility on the participating agents to be able to perform norm synthesis themselves, but this would require us to set minimum capabilities, such as their having high cognitive ability and providing them access to knowledge of activities within the MAS. The framework is intended to be accessible for most normative agents so we prefer an approach which does not impose significant requirements.

4.2   Norm Synthesis Stage

The norm synthesis process, as depicted in Fig. 2, component 1(b), enables the encoding of the agents' request, which would have been influenced by their

behaviour, into an explicit legal norm that can address the issue within the request. Synthesizer agents, upon receipt of a proposal, begin the process of synthesising a new norm to address the proposal, but additionally, the synthesizer agents engage in discussion and the establishment of a consensus about the request. The sub-processes of the norm synthesis stage are: proposal, synthesis, discussion/deliberation, and agreement, each of which we now discuss in more detail.

Proposal. After receiving a proposal, synthesiser agents must parse the request in preparation for synthesising one or more norms. Agents utilise a provided template to make proposals and the synthesizer will be equipped to interpret it. Synthesizers will normally address requests chronologically and will only handle one request at a time; if multiple requests occur over a short space of time, they are queued. It is at this point that synthesisers have to decide which requests warrant action, by applying a filtering mechanism. Synthesizers could employ an automated rejection process, which returns a message to participating agents when their norm requests have been deemed to not warrant action. We note however that a filtering service could potentially be made available to participating agents, whereby an intermediate mechanism processes the complaint, provides feedback and then the agent decides whether to make a request of the synthesizer agent.

Synthesis. The act of synthesising or encoding the norm occurs here. This process utilises the originating agent's request, information about the environment that it can perceive and the domain context. Alternatively, if the participating agent is able to synthesise a norm that will meet the identified need and include it in the request, then this process entails the synthesizer agent determining the validity of the norm.

During synthesis, the synthesizer agent must reason whether the synthesised norm(s) is capable of addressing the need identified by the agent. For example, for a conflict situation, the norm should ensure that once followed, that conflict situation is no longer observed, or for an innovation norm, the norm once adopted ensures that the agents are more efficient in accomplishing their goals, and those of the MAS, or the observation of more acceptable states in the MAS.

Existing literature employs different mechanisms for synthesising norms. SENSE [30] considers the context of the interacting agents in the time-step before the conflict situation and the actions taken by the agents in that time-step. It then synthesises a norm by prohibiting the action of any one of the participating agents. This is similar to IRON [26,28]. Another technique, utilised in [2,23], is built on inductive logic programming and uses the following inputs: a normative system (which could be an empty set), the observation traces and the normative conditions that must hold in the final states. It utilises the proceeding to revise the normative system producing norms (rules) that are compatible with the supplied trace and the condition. Both [2,23] and [30] are offline mechanisms, but can potentially be replicated for online use.

Finally, the proposed new norm is put forward to the other synthesizer agents for discussion and agreement.

Discussion/Deliberation. The synthesising agent informs other synthesizer agents about the new norm and solicits a discussion via a discussion mechanism, which we will call a discussion board, that is visible and accessible only to synthesizer agents. There is potential to have the agents in the system be able to view this discussion board. We considered the possibility of providing *read-only* access to participating agents, but this would require them to have the ability to parse and understand the discussions on the board, erecting further technical barriers to access.

Synthesizers are alerted when there is a new activity on the discussion board and they can engage in the discussion when ready. The presenting synthesizer agent may be required to defend their proposed norm. An argumentation or negotiation scheme may be appropriate here and can be developed based on existing argumentation frameworks such as [3–5,37]. Synthesiser agents will be required to defend the validity of the proposed norm under scrutiny to other synthesizers, not unlike the agents in [3], who have to defend their preferred action or abort.

The resolution of internal norm conflicts and modifications, based on new perspectives highlighted by the other synthesizers, can result in the revision of the proposed norm. It is the responsibility of the synthesizer that proposes the norm to ensure that after any changes made during this process, the modified norm can still meet the needs of the initial request. Therefore, it is possible that the initial proposed norm may be revised, and the actual outcome is a consequence of modifications made as part of this stage.

Agreement. To proceed to the next stage, there needs to be agreement, where, say, a predetermined percentage of synthesisers must agree that the new norm should be introduced into the normative system. Synthesizers can come to agreement using several methods: (i) a clear direction forward can be established based on the outcome(s) of the previous discussion phase, (ii) synthesizers can attempt to reach a consensus on the norm proposal for or against it proceeding, or (iii) synthesizers can utilise a voting mechanism (agreed in advance).

If a decision to allow the norm proposal to proceed cannot be reached, the presenting synthesizer must inform the proposing agent that the norm change was rejected. The synthesizer can provide reasons for the rejection so that the proposing agent can possibly refine and submit a new request/proposal. Once a decision is reached, the agreed norm change is passed on to the decision stage for a final decision on its inclusion in the normative system of the MAS.

4.3   Decision Stage

The decision about whether a norm change will be made is done by a decision making mechanism, which we will refer to as the "Oracle" as depicted in Fig. 2, component 1(c). The Oracle is assumed to have access to the entire state of the MAS. The context and domain of the MAS may require that this decision incorporates human input [35,46]. Human input might be necessary to preclude the risk of destabilisation of the MAS due to repeated norm synthesis or the

synthesis of norms that are potentially detrimental to the purpose of the MAS. The domain of the MAS being modelled will determine whether the Oracle's final decision to modify the normative system requires human input, or if an automated decision only can be allowed. For example, in a MAS interacting directly with humans or whose impact can be catastrophic, it may be preferred if a human(s) authorises the changes to the normative system.

A possible solution would be for the mechanism to make a recommendation that needs to be accepted by a human oversight process before implementation. We note that though including a human at this stage can provide human accountability, it raises the challenge of ensuring that the human and the system have the same understanding of the goals of the system. Otherwise we risk a situation where both human and system may need to prove which party is correct. This could necessitate further arbitration mechanisms to avoid stagnation within the MAS.

Though norm conflicts would have been considered in discussions with other synthesizers, there may still be norm conflicts arising from external interaction. This can occur when the MAS may be itself governed by a higher level MAS [19] and the norms of this MAS must not conflict with the norms of the governing MAS. These conflicts, if they exist must be resolved, and this can be achieved utilising techniques similar to [19]. The stability of the normative system within the MAS must also be considered and should be incorporated into the final decision on the modification of the normative system. At the end of this process, the decision to include or reject the norm proposed is communicated to the synthesizer agents.

5   Norm Propagation

At the end of the norm creation stage, a norm has either been accepted for inclusion into the normative system or been rejected. The preceding decision must be communicated to the synthesizer agents by the Oracle. The norm propagation activities are depicted in Fig. 2, component 2. In the case of a modification to the normative system, synthesiser agents are tasked with spreading this information to the agents they are responsible for as it is necessary that all participating agents in the MAS be aware of any changes to the normative system. Synthesizers can choose an appropriate norm propagation mechanism to communicate with their assigned participating agents. Broadcasting is one solution, a common knowledge source, as used in [43] is another, which advises agents when new information is available, or synthesizers can use a distributed sharing mechanism, where they inform some agents, who then inform other agents they are connected to. This latter approach could possibly affect whether all agents become aware of the modifications to the normative system in a timely manner. Alternatively, if the norm change proposal was rejected, the synthesizer needs to communicate this to the agent that made the initial proposal.

6   Norm Adoption

The adoption of the norm requires the agent to internalise and reason about adopting the norm. The norm adoption activities are depicted in Fig. 2, component 3. Internalising the norm means making an internal representation of the norm, while the adoption of the norm is the decision whether to adopt the norm after reasoning about it. The internalisation of the norm is initiated when agents are made aware of modifications to the normative system. As agents are assumed to have an explicit internal representation of norms, see Sect. 3, this internalisation of the norm means agents will have to incorporate the norm as part of their beliefs. We note the potential for misinterpretations and changes to the norms as a result of this internalisation process [12,17], however we consider the inclusion of a centralised, externally referenceable normative system as helping to minimise the effects of this.

Once internalised, agents can reason about the validity and applicability of the norm and decide whether it is useful to adopt it. Some agents will do this once and adopt the norm in every situation that it is applicable going forward, as with the normative agents in [9]. Other agents will reason about it every time a decision to act needs to be made, similar to agents in [4,38] and graded normative agents in [9].

7   Norm Emergence

The norm emergence activities are depicted in Fig. 2, component 4. Once an agent has adopted the norm and complies with it, the percentage of agents doing so may reach the threshold for emergence. At this point it can be said that the norm has emerged within the MAS. If the percentage of agents adopting the norm never reaches the threshold for emergence, it may be useful to see if the opposite prevails. That is, where a large percentage of agents violate the norm. If this is the case then it may be necessary to reevaluate the norm's place within the normative system, which could ultimately lead to its removal.

8   Conclusions

In this paper we present a framework for norm emergence in normative MAS premised on utilising the experiences of participating agents to trigger changes in the normative system. We describe a distinguished population of synthesizer agents, who accept requests from participating agents and synthesise norms in response to these requests. We believe that this is a useful approach for open MAS as there is no need to impose any requirements on the participating agents but can instead provide a set of synthesizer agents that have the requisite capabilities to perform the role. The introduction of synthesisers, that can propose changes to the normative system, provides participating agents with recourse when they are not satisfied with their experience in the MAS. Instead of leaving the MAS, they can potentially initiate changes within the MAS that will not

only improve the experience for themselves, but perhaps for other agents. Participating agents will only need to know who their synthesizer agent is and what information should be provided.

We posit that the use of special-purpose synthesizer agents that address the needs of participating agents, equips the MAS to modify its normative system at runtime, thereby facilitating decentralised runtime norm synthesis in the MAS. Furthermore we posit that the inclusion of discussion and agreement phases for synthesizer agents to agree on the inclusion of norms is an important addition to the framework. A synthesizer will synthesise norms based on a partial context, and as a result it could synthesise norms that, though capable of resolving the proposed issue, are not useful for the collective MAS.

The remaining synthesiser agents will be able to assess the proposed norm and determine how it will affect their own view of the MAS. The intention is that only norms that are beneficial to a majority or all agents of the MAS shall achieve consensus to proceed to the final stage of verification in the MAS: the *decision stage*. The goal is that this additional layer will aid the MAS by helping to maintain the stability of the normative system and preventing agents from introducing norms that are detrimental to it. It is also the stage where norms can be checked for compliance with governing external MASs, if they exist and/or non-negotiable MAS norms/rules. This stage can also allow for human input depending on the domain of the MAS.

The normative system is modified if the norm is approved by the Oracle mechanism in the decision stage, and agents within the MAS need to be informed of changes to the normative MAS. Participating agents will then reason about adopting the norm and over time a threshold of agents may choose to adopt the norm. Ultimately we may observe the emergence of the norm in the MAS that was synthesised based on a request by a participating agent.

References

1. Alechina, N., Bulling, N., Dastani, M., Logan, B.: Practical run-time norm enforcement with bounded lookahead. In: Weiss, et al. [45], pp. 443–451
2. Athakravi, D., Corapi, D., Russo, A., De Vos, M., Padget, J., Satoh, K.: Handling change in normative specifications. In: Baldoni, M., Dennis, L., Mascardi, V., Vasconcelos, W. (eds.) DALT 2012. LNCS (LNAI), vol. 7784, pp. 1–19. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37890-4_1
3. Atkinson, K., Bench-Capon, T.: Practical reasoning as presumptive argumentation using action based alternating transition systems. Artif. Intell. **171**(10–15), 855–874 (2007). https://doi.org/10.1016/j.artint.2007.04.009
4. Bench-Capon, T., Modgil, S.: Norms and value based reasoning: justifying compliance and violation. Artif. Intell. Law **25**(1), 29–64 (2017). https://doi.org/10.1007/s10506-017-9194-9
5. Boella, G., van der Torre, L.: BDI and BOID argumentation. In: Proceedings of IJCAI Workshop on Computational Models of Natural Argument, vol. 3 (2003)
6. Campos, J., Lopez-Sanchez, M., Salamó, M., Avila, P., Rodríguez-Aguilar, J.A.: Robust regulation adaptation in multi-agent systems. ACM Trans. Auton. Adapt. Syst. **8**(3), 1–27 (2013). https://doi.org/10.1145/2517328

7. Conte, R., Castelfranchi, C.: From conventions to prescriptions. Towards an integrated view of norms. Artif. Intell. Law **7**(4), 323–340 (1999)

8. Cranefield, S., Savarimuthu, T., Meneguzzi, F., Oren, N.: A Bayesian approach to norm identification. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, pp. 1743–1744. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2015)

9. Criado, N., Argente, E., Botti, V.: Normative deliberation in graded BDI agents. In: Dix, J., Witteveen, C. (eds.) MATES 2010. LNCS (LNAI), vol. 6251, pp. 52–63. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16178-0_7

10. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime norm revision using Bayesian networks. In: Miller, T., Oren, N., Sakurai, Y., Noda, I., Savarimuthu, B.T.R., Cao Son, T. (eds.) PRIMA 2018. LNCS (LNAI), vol. 11224, pp. 279–295. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03098-8_17

11. Dell'Anna, D., Dastani, M., Dalpiaz, F.: Runtime revision of norms and sanctions based on agent preferences. In: Elkind, E., Veloso, M., Agmon, N., Taylor, M.E. (eds.) Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2019, Montreal, QC, Canada, May 13–17, pp. 1609–1617 (2017). International Foundation for Autonomous Agents and Multiagent Systems (2019). http://dl.acm.org/citation.cfm?id=3331881

12. Frantz, C., Pigozzi, G.: Modeling norm dynamics in multiagent systems. FLAP **5**(2), 491–564 (2018). https://www.collegepublications.co.uk/downloads/ifcolog00022.pdf

13. Frantz, C.K., Purvis, M.K., Savarimuthu, B.T.R., Nowostawski, M.: Modelling dynamic normative understanding in agent societies. Scalable Comput. Pract. Exp. **16**(4), 355–380 (2015). http://www.scpe.org/index.php/scpe/article/view/1128

14. Ghorbani, A., Bravo, G.: Managing the commons: a simple model of the emergence of institutions through collective action. Int. J. Commons **10**(1), 200–219 (2016). https://doi.org/10.18352/ijc.606

15. Ghorbani, A., Bravo, G., Frey, U., Theesfeld, I.: Self-organization in the commons: an empirically-tested model. Environ. Modell. Softw. **96**, 30–45 (2017). https://doi.org/10.1016/j.envsoft.2017.06.039

16. Haynes, C., Luck, M., McBurney, P., Mahmoud, S., Ví tek, T., Miles, S.: Engineering the emergence of norms: a review. Knowl. Eng. Rev. **32**, e18 (2017). https://doi.org/10.1017/S0269888917000169

17. Hollander, C.D., Wu, A.S.: The current state of normative agent-based systems. J. Artif. Soc. Soc. Simul. **14**(2) (2011). https://doi.org/10.18564/jasss.1750

18. Huang, X., Ruan, J., Chen, Q., Su, K.: Normative multiagent systems: The dynamic generalization. In: Kambhampati, S. (ed.) Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9–15 July 2016, pp. 1123–1129. IJCAI/AAAI Press (2016). http://www.ijcai.org/Abstract/16/163

19. King, T.C., De Vos, M., Dignum, V., Jonker, C.M., Li, T., Padget, J., van Riemsdijk, M.B.: Automated multi-level governance compliance checking. Auton. Agent. Multi-Agent Syst. **31**(6), 1283–1343 (2017). https://doi.org/10.1007/s10458-017-9363-y

20. Knobbout, M., Dastani, M., Meyer, J.-J.C.: Reasoning about dynamic normative systems. In: Fermé, E., Leite, J. (eds.) JELIA 2014. LNCS (LNAI), vol. 8761, pp. 628–636. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11558-0_46

21. Knobbout, M., Dastani, M., Meyer, J.C.: A dynamic logic of norm change. In: Kaminka, G.A., et al. (eds.) ECAI 2016–22nd European Conference on Artificial Intelligence, 29 August–2 September 2016, The Hague, The Netherlands - Including Prestigious Applications of Artificial Intelligence (PAIS 2016). Frontiers in Artificial Intelligence and Applications, vol. 285, pp. 886–894. IOS Press (2016). https://doi.org/10.3233/978-1-61499-672-9-886

22. Lee, J.H., Padget, J., Logan, B., Dybalova, D., Alechina, N.: *N-Jason*: run-time norm compliance in AgentSpeak(L). In: Dalpiaz, F., Dix, J., van Riemsdijk, M.B. (eds.) EMAS 2014. LNCS (LNAI), vol. 8758, pp. 367–387. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-14484-9_19

23. Li, T.: Normative Conflict Detection and Resolution in Cooperating Institutions. Ph.D. thesis, University of Bath, July 2014. https://researchportal.bath.ac.uk/files/123040954/Tingting_thesis_final.pdf

24. Mahmoud, M.A., Ahmad, M.S., Mohd Yusoff, M.Z., Mustapha, A.: A review of norms and normative multiagent systems. Sci. World J. (2014). https://doi.org/10.1155/2014/684587. Hindawi Publishing Corporation

25. Mashayekhi, M., Du, H., List, G.F., Singh, M.P.: Silk: a simulation study of regulating open normative multiagent systems. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, pp. 373–379. AAAI Press (2016)

26. Morales, J., Lopez-Sanchez, M., Rodriguez-Aguilar, J.A., Wooldridge, M., Vasconcelos, W.: Automated synthesis of normative systems. In: Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, pp. 483–490. International Foundation for Autonomous Agents and Multiagent Systems (2013)

27. Morales, J., López-Sánchez, M., Rodríguez-Aguilar, J.A., Wooldridge, M.J., Vasconcelos, W.W.: Synthesising liberal normative systems. In: Weiss et al. [45], pp. 433–441

28. Morales, J., López-sánchez, M., Rodriguez-Aguilar, J.A., Vasconcelos, W., Wooldridge, M.: Online automated synthesis of compact normative systems. ACM Trans. Auton. Adapt. Syst. **10**(1), 1–33 (2015). https://doi.org/10.1145/2720024

29. Morales, J., Wooldridge, M., Rodríguez-Aguilar, J.A., López-Sánchez, M.: Evolutionary synthesis of stable normative systems. In: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, pp. 1646–1648. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2017)

30. Morales, J., Wooldridge, M., Rodríguez-Aguilar, J.A., López-Sánchez, M.: Off-line synthesis of evolutionarily stable normative systems. Auton. Agents Multi-Agent Syst. (2018). https://doi.org/10.1007/s10458-018-9390-3

31. Morris-Martin, A., De Vos, M., Padget, J.: Norm emergence in multiagent systems: a viewpoint paper. Auton. Agent. Multi-Agent Syst. **33**(6), 706–749 (2019). https://doi.org/10.1007/s10458-019-09422-0

32. Mukherjee, P., Sen, S., Airiau, S.: Emergence of norms with biased interactions in heterogeneous agent societies. In: 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology Workshops, pp. 512–515. IEEE (2007). https://doi.org/10.1109/WI-IATW.2007.115

33. Noriega, P., Sabater-Mir, J., Verhagen, H., Padget, J., d'Inverno, M.: fIdentifying affordances for modelling second-order emergent phenomena with the $\mathcal{WIT}$ framework. In: Sukthankar, G., Rodriguez-Aguilar, J.A. (eds.) AAMAS 2017. LNCS (LNAI), vol. 10643, pp. 208–227. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71679-4_14

34. Ostrom, E.: Governing the Commons. The Evolution of Institutions for Collective Action, CUP (1990)

35. Rahwan, I.: Society-in-the-loop: programming the algorithmic social contract. Ethics Inf. Technol. **20**(1), 5–14 (2018). https://doi.org/10.1007/s10676-017-9430-8

36. Riveret, R., Artikis, A., Busquets, D., Pitt, J.: Self-governance by transfiguration: from learning to prescriptions. In: Cariani, F., Grossi, D., Meheus, J., Parent, X. (eds.) DEON 2014. LNCS (LNAI), vol. 8554, pp. 177–191. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08615-6_14

37. Riveret, R., Rotolo, A., Sartor, G.: Probabilistic rule-based argumentation for norm-governed learning agents. Artif. Intell. Law **20**(4), 383–420 (2012). https://doi.org/10.1007/s10506-012-9134-7

38. dos Santos Neto, B.F., da Silva, V.T., de Lucena, C.J.P.: Developing goal-oriented normative agents: the NBDI architecture. In: Filipe, J., Fred, A. (eds.) ICAART 2011. CCIS, vol. 271, pp. 176–191. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-29966-7_12

39. Savarimuthu, B.T.R., Arulanandam, R., Savarimuthu, S.: Emergence of a sharing norm in a simulated hunter-gatherer society. In: Proceedings of the 2011 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2011, Campus Scientifique de la Doua, Lyon, France, 22–27 August 2011, pp. 34–37 (2011). https://doi.org/10.1109/WI-IAT.2011.141

40. Savarimuthu, B.T.R., Cranefield, S.: Norm creation, spreading and emergence: a survey of simulation models of norms in multi-agent systems. Multiagent Grid Syst. **7**(1), 21–54 (2011). https://doi.org/10.3233/MGS-2011-0167

41. Savarimuthu, B.T.R., Cranefield, S., Purvis, M.K., Purvis, M.A.: Norm emergence in agent societies formed by dynamically changing networks. Web Intell. Agent Syst. Int. J. **7**(3), 223–232 (2009). https://doi.org/10.3233/WIA-2009-0164

42. Savarimuthu, B.T.R., Cranefield, S., Purvis, M., Purvis, M.K.: Identifying prohibition norms in agent societies. Artif. Intell. Law **21**(1), 1–46 (2013). https://doi.org/10.1007/s10506-012-9126-7

43. Savarimuthu, B.T.R., Purvis, M., Purvis, M., Cranefield, S.: Social norm emergence in virtual agent societies. In: Baldoni, M., Son, T.C., van Riemsdijk, M.B., Winikoff, M. (eds.) DALT 2008. LNCS (LNAI), vol. 5397, pp. 18–28. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-540-93920-7_2

44. Villatoro, D., Sen, S., Sabater-Mir, J.: Topology and memory effect on convention emergence, pp. 233–240. IEEE (2009). https://doi.org/10.1109/WI-IAT.2009.155

45. Weiss, G., Yolum, P., Bordini, R.H., Elkind, E. (eds.) Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, 4–8 May 2015. ACM (2015)

46. Zanzotto, F.M.: Viewpoint: human-in-the-loop artificial intelligence. J. Artif. Int. Res. **64**(1), 243–252 (2019). https://doi.org/10.1613/jair.1.11345