Formal Verification of Opinion Formation in Swarms

Panagiotis Kouvaros
Department of Computing
Imperial College London
p.Kouvaros@imperial.ac.uk

Alessio Lomuscio
Department of Computing
Imperial College London
a.lomuscio@imperial.ac.uk

ABSTRACT

We investigate the formal verification of consensus protocols in swarm systems composed of arbitrary many agents. We put forward templates to define the behaviour of the agents in an opinion dynamics setting and formulate their verification in terms of the associated parameterised model checking problem. We define a finite abstract model that we show to formally simulate any swarm of any size, thereby precisely encoding any concrete instantiation of the swarm. We give an automatic procedure for verifying temporal-epistemic properties of consensus protocols by model checking the associated finite abstract model. We present a toolkit that can be used to generate the abstract model automatically and verify the protocol by symbolic model checking. We use the toolkit to verify the correctness of majority rule protocols in arbitrary large swarms.

1. INTRODUCTION

Robotic swarm systems have been put forward as a robust alternative to single-robots in a variety of domains, e.g., remote exploration, maintenance of industrial plants, etc. In its simplest form a swarm is a collection of agents all running a simple program. The physical agents in a swarm are often relatively low-powered devices with limited sensing and communication capabilities. Even if their capabilities are limited by physical and computational constraints, their collective capabilities can be significant. For example, simple protocols can ensure robotic swarms can ensure flocking behaviour, or other emerging properties [2, 26, 27].

Consensus, or opinion formation, protocols, [6, 12, 17, 18, 21, 28, 29, 30] are of particular significance in the context of robot swarms as they can be used as the basis for coordination. The aim of a consensus protocol is for the agents in the system to agree on a particular outcome, e.g., which area to move to as a swarm, electing a leader. Before being applied and developed for swarms, they were initially introduced in distributed computing [8, 11] and also used for reasoning about social, economic, and natural sciences problems and scenarios.

In a consensus formation protocol agents maintain a state encoding their present opinion on the issue they need to converge upon. The opinion is associated with an action that an agent may perform. For example, if the agents need to decide where to travel to, this may simply be the direction of travel. The agent's opinion changes at each time step following observations and communication with its peers.

A key issue with opinion formation protocols is to investigate their convergence. In this paper we put forward a formal methodology that can be applied to analyse consensus protocols that follow the majority rule. In these protocols the agents in the system update their opinions simply by considering the opinions of their neighbours and adopting the one that is favoured by the simple majority of its neighbours. While this mechanism appears simple, a large number of applications including collective transport [6], task sequencing [22], and the best-of-n decision problem [30] rely on the majority rule or simple variations of it. For example, variations of the rule have been put forward to account for *latency* or *nesting*. In latency models agents do not change their opinion for a time that is proportional to the quality of their current opinions [6]. In nesting models, the process of opinion formation only takes place in a nest, where the agents with opinions of better quality spend proportionally more time [30].

The analysis of these systems is normally conducted by means of two techniques. Optimisation techniques can provide assurances of the behaviour of the swarm; these use differential equations on continuous domains and assume an infinite number of agents in the system [28, 29]. In contrast, simulation techniques compute the actual evolution but only for a swarm of a given size [6, 30]. However, an ideal analysis of a swarm should give guarantees of a behaviour of the system irrespective of the number of agents in the system when it is deployed. A key essence of protocol verification is, indeed, that conclusions ought to be drawn independently from or with minimal assumptions on the number of agents in the system.

Parameterised model checking has been applied for the analysis of generic swarm systems of an arbitrary number of components [15]. However, the models there introduced are generic and their concrete applicability to classes of protocols is largely open. In this paper we extend the methodology put forward in [15] to model consensus protocols following the majority rule. The results we report indicate that under limited assumptions the methodology and the toolkit we present can be used to analyse any consensus protocol in this class. We are not aware of other automated techniques that can provide formal guarantees on the outcome of consensus protocols irrespective of the number of agents in the system.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), John Thangarajah, Karl Tuyls, Stacy Marsella, Catholijn Jonker (eds.), May 9–13, 2016, Singapore.*

The rest of the paper is organised as follows. In Section 2 we introduce a template-based semantics for swarm systems specialised to encode consensus protocols, and we give the syntax of the temporal-epistemic language that we will use to analyse the protocols. In Section 3 we introduce a class of abstract models that can be used to reason about the infinitely many instantiations of a given consensus protocol. In Section 4 we present an implementation of the theoretical results in the form of a novel toolkit that can be used to analyse consensus protocols. We use the toolkit to validate a majority rule protocol. We conclude in Section 5 by discussing related work.

2. OPINION FORMATION SEMANTICS

In this section we introduce an opinion formation semantics for robotic swarms. In line with the standard treatment of robot swarms, the semantics here introduced accounts for an unbounded collection of behaviourally identical agents [2, 26]. Each agent interacts with its peers through local exchanges, realised by the repeated application of the majority-rule, that enable the swarm to reach consensus on a certain opinion.

In its simplest form, a majority-rule protocol is described as follows [17]. At each time step a team of three randomly picked agents is formed. Each agent can be in one of two states; each state is associated with one of two opinions. The members of a team adopt the opinion held by the majority of the members in team. This process is repeated until consensus is reached on one of the two opinions. Studies have shown that a collective agreement is eventually established on the opinion with the highest initial density [17].

Several extensions to the simple protocol above have been proposed. For example, [6, 18] introduce the concept of *latency*. Latency refers to a period of time in which the agents do not engage in the opinion formation procedure. Specifically, following the adoption of an opinion, an agent enters a latent state, from which the agent does not interact with other agents. The time in which the agent remains in the latent state depends on the quality of the recently adopted opinion. As a result, simulation studies indicate that the swarm collectively adopts the opinion characterised by the shortest latency period [6]. For instance, if opinions represent actions, the swarm converges on the action requiring less time to perform. For example, a swarm may decide to take the shortest path to a destination [6, 28], the best site to explore [12, 30], and so forth. In these cases the majority-rule protocol is often used as a decision making mechanism to solve the best-of-n decision problem [21, 30], i.e., the problem of establishing consensus on the opinion with the highest quality among a set of n opinions. In the rest of this section we introduce a formal semantics for reasoning about the temporal-epistemic properties of opinion formation protocols based on the majority-rule.

2.1 Model

We begin by specifying a generic *agent template* modelling the agents in a swarm. The concrete system of n agents can be constructed from the template by providing the number $n \geq 1$ of actual agents in the swarm.

The models we define are loosely based on interpreted systems [10] and parameterised interpreted systems [15, 14]. They are, however, specialised and extended to the modelling of opinion formation protocols. The agent template is defined as follows.

Definition 2.1 (Agent Template). *An agent template $\mathcal{A}$ is a tuple $\mathcal{A} = (O, h, \alpha, t)$, where:*

- *O is a nonempty and finite set of opinions;*
- *$h : O \to \mathbb{N}$ is a mapping from the set of opinions into the set natural numbers, where $h(o)$ represents the quality of opinion o. O and h define a set*

$$L = \{(o, v, l) \colon o \in O, 0 \leq v \leq \max(h(o) \colon o \in O), l \in \{\mathit{false}, \mathit{true}\}\}$$

  of local states, where each triple (o, v, l) encodes an opinion o, a latent value v, and whether or not the template is into latent state ($l = true$, and $l = false$, respectively).
- *$\alpha \in \mathbb{N}$ is the size of the neighbourhood for the template at any given time step.*
- *$t : L \times O \to L$ is a transition function that returns the next local state given the current local state and the majority opinion held by neighbouring agents.*

Note that a local state is built from an observable component representing an opinion, and a non-observable component associated with a latent value and the latent status. The domain of the latent value depends on the opinions' maximum quality. Intuitively, different behaviour may be associated with different opinions, thereby allowing for the modelling of quality-depended protocols. For instance, as we exemplify in Section 4, the latent value can be used to keep track of the time an agent is engaged in the protocol before it goes into latent state; this period of time is proportional to its currently held opinion. For a local state l, we write $\mathit{opinion}(l)$, $\mathit{value}(l)$, and $\mathit{latent}(l)$ to denote the opinion, the latent value, and the latent status, respectively, encoded in l. We assume that whenever $\mathit{latent}(l) = \mathit{false}$ and $t(l, o) = l'$, then $\mathit{opinion}(l') = o$; i.e, at each time step, an agent switches to the majority opinion in its neighbourhood if not in latent state.

Note the special case of the majority rule in which an agent's neighbourhood is equally split among opinions is typically resolved by either considering neighbourhoods of an odd number of agents, or withholding the currently held opinion, or randomly adopting an opinion [17, 30]. This can be easily added to the framework without altering any of the technical details presented in this paper.

An agent template describes an unbounded family of *concrete opinion formation systems*; each system is obtained by instantiating the template with the actual number of agents in the system. In other words, given $n \geq 1$, the concrete system $\mathcal{S}_{\mathcal{A}}(n)$ is the result of composing precisely n concrete agents participating in the opinion formation system. Each concrete agent is represented in the interpreted systems formalism [10], a standard semantics for multi-agent systems. That is, a concrete agent i is associated with a set of local states L_i, a set of actions Act_i, a protocol P_i that governs which actions may be performed in a given local state, and a transition function t_i that determines the temporal evolution of the agent.

Definition 2.2 (Concrete Agent). *Given an agent template $\mathcal{A} = (O, h, \alpha, t)$ and $n \geq 1$, the concrete agent A_i, with $1 \leq i \leq n$, is a tuple $A_i = (L_i, I_i, Act_i, P_i, t_i)$, where:*

- $L_i = L$ *is the set of local states for agent* i*;*
- $I_i = \{(o, v, l) : o \in O, v = h(o), l = false\}$ *is the set of initial states for agent* i*;*
- $Act_i = \{act_o : o \in O\}$ *is the set of actions for agent* i*;*
- *the local protocol for agent* i *is* $P_i : L_i \to \mathcal{P}(Act_i)$ *is such that for each* $l \in L_i$*,*
  - $P_i(l) = \{act_{opinion(l)}\}$ *if* $latent(l) = true$*;*
  - $P_i(l) = Act_i$ *if* $latent(l) = false$*.*
- *the local transition function for agent* i $t_i : L_i \times Act_i \to L_i$ *is such that* $t_i(l, act_o) = l'$ *iff* $t(l, o) = l'$*;*

So, a concrete agent inherits from its template the set of local states and its local transition function. The agent is initially active and it holds an arbitrary opinion. For each opinion o, the agent admits a corresponding action act_o that is enabled by the protocol whenever the agent holds the opinion o; intuitively, act_o represents the majority opinion in its neighbourhood upon which the agent acts. As such, whenever the agent is in latent state, its protocol only enables the action associated with its currently held opinion; i.e, an agent does not engage in the opinion formation protocol when in latent state in that it can only update its latent value and status independently of the other agents' opinions.

We now describe the overall system. A global state $g = (l_1, \ldots, l_n)$ is a tuple of local states for all the agents in the system; g describes the configuration of the system at a particular instant of time. Given a global state g, we write $g.i$ for the local state l_i of agent A_i in g. Given an opinion o, we write $\#(g, o)$ to denote the number of agents with opinion o in g. By $\#^{false}(g)$, we mean the number of agents that are not in latent state in g. We use $\#^{false}(g, o)$ to express the number of agents with opinion o that are not in latent state in g, and $\#^{true}(g, o)$ for the number of agents with opinion o that are in latent state in g.

Following the application of the majority rule at a global state g, an agent updates its current opinion to opinion o if there are at least $mj(g)$ agents with opinion o in its neighbourhood, where $mj(g)$ is equal to the following:

$$mj(g) = \left\lceil \frac{\alpha}{|\{o : opinion(g.i) = o \text{ for some } 1 \leq i \leq n\}|} \right\rceil$$

The probability $P(g, o)$ that an agent will adopt opinion o when applying the majority rule at state g is calculated as follows:

$$P(g, o) = \sum_{r=mj(g)}^{\min(\alpha, \#^{false}(g,o))} \frac{\binom{\alpha}{r}\binom{\#^{false}(g)-\alpha}{\#^{false}(g,o)-r}}{\binom{\#^{false}(g)}{\#^{false}(g,o)}}$$

We now define the concrete semantics, i.e, the notion of a concrete opinion formation system.

Definition 2.3 (Concrete System). *Given an agent template* $\mathcal{A} = (O, h, \alpha, t)$ *and* $n \geq 1$*, the* concrete opinion formation system *(OFS) with* n *agents is a tuple* $\mathcal{S}_{\mathcal{A}}(n) = (G(n), I(n), R(n), V(n))$*, where:*

- $G(n) \subseteq L_1 \times \ldots \times L_n$ *is the set of global states reachable via* $R(n)$ *from the set of initial global states* $I(n) = I_1 \times \ldots \times I_n$*;*
- $R(n) \subseteq G(n) \times G(n)$ *is the global transition relation that is defined as* $(g, g') \in R(n)$ *iff the following hold:*
  - *for all* $o \in O$*, we have that*
    $$\#(g', o) = [\#^{false}(g).\, P(g, o)] + \#^{true}(g, o)$$
    where $[x]$ *denotes the nearest integer to* x *plus-minus 1 such that* $\sum_{o \in O} \#(g', o) = n$*;*
  - *there is a joint action* $(a.1, \ldots, a.n) \in Act_1 \times \ldots \times Act_n$ *such that for all* $1 \leq i \leq n$*, we have that* $t_i(g.i, a.i) = g'.i$*;*
- $V(n) : G(n) \to \mathcal{P}(AP)$ *is a labelling function for a set* $AP = \{p(o) : o \in O\}$ *of atomic propositions that is defined as follows:* $p(o) \in V(n)(g)$ *iff* $opinion(g.i) = o$ *for every* $1 \leq i \leq n$*.*

A path is a sequence $\pi = g^0 g^1 g^2 \ldots$ with $(g^i, g^{i+1}) \in R(n)$, for every $i \geq 0$. Given a path π we write $\pi(i)$ for the i-th state in π. The set of all paths originating from a state g is denoted by $\Pi(g)$.

Following the above definition, an agent template generates a family of systems; each system is composed of a different number of agents. The concrete transition relation $R(n)$ is such that whenever $(g, g') \in R(n)$, the density of each opinion o in g' corresponds to the probability that an agent will have said opinion in the next time step; this is equal to the probability $P(g, o)$ that an agent will adopt opinion o with the application of the majority rule, plus the ratio $\frac{\#^{true}(g,o)}{n}$ of agents with opinion o that are in latent state in g. As such, our analysis is not probabilistic, since we do not consider transitions that reflect every possible outcome of a given probability distribution, but it is qualitative in the sense that it aims to establish the correctness of a given protocol w.r.t its average behaviour on an infinite number of rounds. Further, note that $R(n)$ does not explicitly depend on the neighbourhood of each agent. Indeed, we abstract away the spacial position for a robot and, in line with existing literature [6, 17], we assume a random neighbourhood for each agent at any instant of time. The labelling function assigns an atomic proposition $p(o)$ on a state iff all the agents agree on opinion o in the state. As we explain below, this will enable us to define consensus specifications which can be interpreted on a concrete system.

2.2 Specifications

We express OFSs specifications in ACTLK, the universal fragment of the temporal-epistemic logic CTLK [23]. CTLK has long been used to express temporal-epistemic properties of the agents in a multiagent system. We fix the notation below but refer to [24] for more details. Given a set *Agents* of agents and a set AP of atomic propositions, ACTLK formulae are defined by the following BNF grammar:

$$\phi ::= p(o) \mid \neg p(o) \mid \phi \wedge \phi \mid \phi \vee \phi \mid AX\phi \mid A(\phi U \phi) \mid A(\phi R \phi) \mid K_i\phi \mid E_\Gamma\phi \mid C_\Gamma\phi$$

where $p(o) \in AP$, $i \in Agents$, and $\Gamma \subseteq Agents$. The epistemic modality $K_i\phi$ is read as "agent i knows that ϕ"; $E_\Gamma\phi$ encodes "every agent in group Γ knows that ϕ"; and $C_\Gamma\phi$ expresses "it is common knowledge in Γ that ϕ" [10]. The temporal modality $AX\phi$ represents "for all paths, ϕ holds at the next step"; $A(\phi U \psi)$ stands for "for all paths, at some point ψ holds and before then ϕ is true along the path"; and

$A(\phi R \psi)$ denotes "for all paths, ψ holds along the path up to and including the point when ϕ becomes true in the path". The interpretation of ACTLK formulae on an OFS $\mathcal{S}_{\mathcal{A}}(n)$ is given as usual: the temporal modalities are interpreted by means of the global transition relation [5], and the epistemic modalities are interpreted by using the epistemic possibility relations [10]. The epistemic possibility relation for an agent $i \in \mathit{Agents}$ is defined as follows:

$$\sim_i = \left\{(g, g') \in G(n) \times G(n) : g.i = g'.i\right\}.$$

We write $(\mathcal{S}(n), g) \models \phi$ to mean that a formula ϕ is true at state g in $\mathcal{S}_{\mathcal{A}}(n)$. If $\mathcal{S}_{\mathcal{A}}(n)$ is clear, then we simplify the notation to $g \models \phi$.

Definition 2.4 (Satisfaction of ACTLK). *Given an OFS $\mathcal{S}_{\mathcal{A}}(n)$, the satisfaction relation $\models$ is inductively defined as follows:*

$g \models p(o)$	*iff*	$p(o) \in V(n)(g)$ *for any* $p(o) \in AP$;
$g \models \neg p(o)$	*iff*	$g \not\models p(o)$;
$g \models \phi \wedge \psi$	*iff*	$g \models \phi$ *and* $g \models \psi$;
$g \models \phi \vee \psi$	*iff*	$g \models \phi$ *or* $g \models \psi$;
$g \models AX\phi$	*iff*	*for every* $\pi \in \Pi(g)$, *we have that* $\pi(1) \models \phi$;
$g \models A(\phi U \psi)$	*iff*	*for every* $\pi \in \Pi(g)$, *there is* $i \geq 0$ *such that* $\pi(i) \models \psi$ *and* $\pi(j) \models \phi$ *for all* $0 \leq j < i$;
$g \models A(\phi R \psi)$	*iff*	*for every* $\pi \in \Pi(g)$ *and for all* $i \geq 0$, *if* $\pi(j) \not\models \phi$, *for all* $0 \leq j < i$, *then* $\pi(i) \models \psi$;
$g \models K_i\phi$	*iff*	*for all* $g' \in G(n)$, $g \sim_i g'$ *implies* $g' \models \phi$;
$g \models E_\Gamma \phi$	*iff*	*for all* $g' \in G(n)$, $g \sim_{E,\Gamma} g'$ *implies* $g' \models \phi$;
$g \models C_\Gamma \phi$	*iff*	*for all* $g' \in G(n)$, $g \sim_{C,\Gamma} g'$ *implies* $g' \models \phi$.

In the definition above, the relation $\sim_{E,\Gamma}$ is defined as the union of the epistemic relations for all the agents in Γ: $E\phi \triangleq \bigcup_{i \in \Gamma} \sim_i$, and the relation $\sim_{C,\Gamma}$ is defined as the transitive closure of $\sim_{E,\Gamma}$. An ACTLK formula ϕ is said to be true in $\mathcal{S}_{\mathcal{A}}(n)$, denoted as $\mathcal{S}_{\mathcal{A}}(n) \models \phi$, if $(\mathcal{S}_{\mathcal{A}}(n), g) \models \phi$ for every $g \in I(n)$. The customary abbreviations of *truth* and *falsity* are assumed: $\top \triangleq p(o) \vee \neg p(o)$, $\bot \triangleq p(o) \wedge \neg p(o)$. Further we define $AF\phi \triangleq A(\top U \phi)$ representing "for all paths, ϕ eventually becomes true", and $AG\phi \triangleq A(\bot R \phi)$ standing for "for all paths, ϕ is globally true".

We now express some specifications of interest. We are interested in verifying whether an OFS will eventually reach consensus on a certain opinion. Observe consensus ought to be stable, i.e., it should not be violated at future points. This is expressed by the following formula:

$$\phi_1 = AF \bigvee_{o \in O} AGp(o)$$

Further, a typical requirement of opinion formation protocols is that the swarm will eventually agree on the opinion of the highest quality:

$$\phi_2 = AFAGp(\boldsymbol{o})$$

where $\boldsymbol{o}$ denotes an opinion with $h(\boldsymbol{o}) \geq h(o)$ for all $o \in O$.

Additionally, we would like to check whether every agent knows the above properties and whether the swarm has common knowledge of the above properties:

$$\phi_3 = E_\Gamma \phi_1 \quad \phi_4 = E_\Gamma \phi_2 \quad \phi_5 = C_\Gamma \phi_1 \quad \phi_6 = C_\Gamma \phi_2$$

where $\Gamma = \mathit{Agents}$. Finally, we are motivated in assessing whether it is always the case that individual knowledge of consensus implies group and common knowledge of consensus, as expressed by the following formulae:

$$\phi_7 = AG(K_i\phi_1 \rightarrow E_\Gamma \phi_1) \qquad \phi_8 = AG(K_i\phi_2 \rightarrow E_\Gamma \phi_2)$$
$$\phi_9 = AG(K_i\phi_1 \rightarrow C_\Gamma \phi_1) \qquad \phi_{10} = AG(K_i\phi_2 \rightarrow C_\Gamma \phi_2)$$

where $i \in \mathit{Agents}$ and $\Gamma = \mathit{Agents}$. Following the unbounded nature of OFSs, for the rest of the paper we restrict ACTLK to specifications ϕ in which:

1. for each K_{A_i} appearing within, $i = 1$. Thus the epistemic modalities appearing in a formula ϕ can only refer to agent A_1. Note, however, any verification result on ϕ can be read as ϕ is referring to any agent of any concrete system. Indeed, studies on the inherent symmetry present in systems of homogeneous agents have shown the following [16]: the interpretation of ϕ on a concrete system $\mathcal{S}_{\mathcal{A}}(n)$ is equivalent to the interpretation of ϕ^i on $\mathcal{S}_{\mathcal{A}}(n)$, where ϕ^i is obtained from ϕ by replacing each epistemic modality K_{A_1} with K_{A_i}.

2. for each E_Γ and C_Γ appearing within, Γ denotes the set of all concrete agents in the system on which the modalities are interpreted.

3. PARAMETERISED VERIFICATION FOR OPINION FORMATION SYSTEMS

In this section we put forward a verification procedure for the analysis of OFSs independently of the number of agents in the system. Our technique is based on previous work in the literature aimed to solve the *parameterised model checking problem* [4, 9, 14]; but it is extended and adapted to opinion formation protocols. In the context of OFSs, we define the decision problem as follows.

Definition 3.1 (PMCP). *Given an agent template $\mathcal{A}$ and an ACTLK formula ϕ, the parameterised model checking problem (PMCP) is the decision problem of determining whether the following holds:*

$$\mathcal{S}_{\mathcal{A}}(n) \models \phi \text{ for every } n \geq \alpha.$$

Obviously the PMCP involves checking an unbounded number of systems. Consequently, the problem cannot by solved by traditional model checking techniques for finite state systems. Indeed, the problem is known to be undecidable in general [1]. However, we observe that any real-world swarm system obeys the *small neighbourhood property* defined below; this enables us not only to show the problem is decidable but also to give a finite abstraction that allows us to solve the PMCP for the case under analysis.

We begin by formulating the small neighbourhood property. Given an agent template $\mathcal{A}$ defined on a neighbourhood size $\alpha \in \mathbb{N}$, we say that an OFS $\mathcal{S}_{\mathcal{A}}(n)$ satisfies the small neighbourhood property if at each time step the number of agents in latent state is much greater than the neighbourhood size. By "much greater" we mean that for any given

global state g and opinion o, we have that $\left(\frac{\#^{false}(g,o)}{\#^{false}(g)}\right)^{\alpha} = \left(\frac{\#^{false}(g,o)-\alpha}{\#^{false}(g)-\alpha}\right)^{\alpha} \pm \epsilon$, for some small constant ϵ. We write $n \gg \alpha$ to denote this. Since swarms are typical made of very large numbers of agents each interacting with very few neighbours, all systems of interest satisfy the small neighbourhood property. Given this we formally restate the PCMP defined above as one that assumes small neighbourhoods.

DEFINITION 3.2 (PMCP FOR SMALL NEIGHBOURHOODS). *The PMCP for OFSs with small neighbourhoods of size α concerns establishing whether the following holds:*

$$\mathcal{S}_{\mathcal{A}}(n) \models \phi_i \text{ for every } n \gg \alpha.$$

In the following we will assume that the OFSs we consider obey the small neighbourhood property and present a solution for the PMCP for small neighbourhoods.

3.1 Weighted abstraction

To solve the PMCP we introduce the notion of *weighted abstraction*. By means of a weighted abstraction we build an abstract model that represents every concrete system. An abstract state is a set of pairs of weights and template local states in $\mathbb{R} \times L$; it represents every concrete state in which the ratio of agents in each local state to all the agents approximates the weight associated with the state. Note that every local state that does not appear in an abstract state is assumed to be associated with a weight equal to 0.

We now describe the construction of the abstract model. Given an abstract state γ and an opinion o, we write $\%(\gamma, o)$ to denote the sum of the weights associated with o, i.e, $\%(\gamma, o) = \sum_{(w,l)\in\gamma, opinion(l)=o} w$. By $\%^{false}(\gamma)$, we mean the sum of weights corresponding to local states that have the latent status set to *false*. We use $\%^{false}(\gamma, o)$ to express the sum of weights corresponding to local states that have the latent status set to *false* and have opinion o, and $\%^{true}(\gamma, o)$ for the sum of weights corresponding to local states that have the latent status set to *true* and have opinion o.

A concrete agent in a global state represented by γ updates its current opinion to the opinion o if there are at least $\hat{mj}(\gamma)$ agents with the opinion o in its neighbourhood, where $\hat{mj}(\gamma)$ is equal to the following:

$$\hat{mj}(\gamma) = \left\lceil \frac{\alpha}{|\{o : \exists (w,l) \in \gamma \text{ with } opinion(l) = o\}|} \right\rceil$$

In the previous section we have defined the probability $P(g, o)$ that a concrete agent in a state g will update its current opinion to the opinion o. We now calculate $P(g, o)$ from an abstract state γ that represents g. If g is a state of n agents, then $P(g, o)$ can be expanded as follows:

$$\sum_{r=\hat{mj}(\gamma)}^{\min(\alpha,\#^{false}(g,o))} \binom{\alpha}{r} \frac{\#^{false}(g,o)}{\#^{false}(g)} \frac{\#^{false}(g,o)-1}{\#^{false}(g)-1} \cdots \frac{\#^{false}(g,o)-r+1}{\#^{false}(g)-r+1} \frac{n-\#^{false}(g,o)}{\#^{false}(g)-r} \cdots \frac{n-\#^{false(g,o)}-\alpha+r+1}{\#^{false}(g)-\alpha+1}$$

From the above and the small neighbourhood assumption, it is easy to show that $\hat{P}(\gamma, o) = P(g, o) \pm \epsilon$, where $\hat{P}(\gamma, o)$ is given by the following:

$$\hat{P}(\gamma, o) = \sum_{r=\hat{mj}(\gamma)}^{\alpha} \left(\binom{\alpha}{r} \left(\%^{false}(\gamma, o)\right)^r \left(\%^{false}(\gamma) - \%^{false}(\gamma, o)\right)^{\alpha-r} \right)$$

$\hat{P}$ provides a means to represent concrete transitions that are enabled from any concrete state represented by a given abstract state. The following definition makes this precise.

DEFINITION 3.3 (WEIGHTED ABSTRACTION). *Given an agent template $\mathcal{A} = (O, h, \alpha, t)$, assume a finite uniformly discrete set W in the metric space $[0, 1]$. The abstract model $\hat{\mathcal{S}}_{\mathcal{A}}$ is a tuple $\hat{\mathcal{S}}_{\mathcal{A}} = \left(\hat{G}, \hat{I}, \hat{R}, \hat{V}\right)$ where*

- *$\hat{G} \subseteq \mathcal{P}(W \times L)$ is the set of abstract states;*
- *$\hat{I}$ is the set of initial abstract states:*

$$\hat{I} = \left\{ X : X \subseteq \{(w, (o, h(o), false) : o \in O\} \text{ and } \sum_{(w,l)\in X} w = 1 \right\}$$

- *$\hat{R} \subseteq \hat{G} \times \hat{G}$ is the abstract transition relation that is defined as $(\gamma, \gamma') \in \hat{R}$ iff the following hold:*
  - *for all $o \in O$, we have that*
    $$\%(\gamma', o) = [\%^{false}(\gamma).\hat{P}(\gamma, o)] + \%^{true}(\gamma, o)$$
    where $[x]$ denotes the nearest weight to x such that $\sum_{o\in O} \%(\gamma', o) = 1$;
  - *for every $(w, l) \in \gamma$ there is $(w', l') \in \gamma'$ with $t(l, o') = l'$, where $o' = opinion(l')$;*
  - *for every $(w', l') \in \gamma'$ there is $(w, l) \in \gamma$ with $t(l, o') = l'$, where $o' = opinion(l')$.*
- *$\hat{V} : \hat{G} \to \mathcal{P}(AP)$ is the abstract labelling function defined as $p(o) \in \hat{V}(\gamma)$ iff for all $(w, l) \in \gamma$ we have that $opinion(l) = o$.*

Thus, the set of initial abstract states represents any possible initial density of opinions in a concrete system. The abstract transition relation is such that whenever $(\gamma, \gamma') \in \hat{R}$, the density of each opinion in γ' corresponds to the probability that a concrete agent in a state represented by γ will have said opinion in the next time step. Finally, the abstract labelling function assigns an atomic proposition $p(o)$ on a state iff the opinion o is encoded in the state.

While weighted abstraction provides a natural way to interpret temporal formulae built from global atomic propositions, it does not allow for the interpretation of epistemic modalities. This is because individual agents' behaviours are not encoded in the abstract model. Therefore, although, as we show below, the abstract model can be used to check temporal specifications, the verification of epistemic specifications is thus far problematic.

To circumvent this, we perform weighted abstraction on the concrete space modulo one agent. In other words, we

compose the abstract model with one concrete agent. In this setting, an abstract state is built from a concrete component and an abstract component. The abstract component is an abstract state as given in Definition 3.3; it represents the local states for the agents $A_2, \ldots, A_n$ in a concrete state g with n agents. The concrete component corresponds to the local state of agent A_1 in g. Given an abstract state γ, we write $\gamma.c$ for the concrete component in γ, and $\gamma.\hat{a}$ for the abstract component in γ.

Definition 3.4 (Partial Weighted Abstraction). *Given an agent template $\mathcal{A} = (O, h, \alpha, t)$, the composition of the abstract model with one concrete agent is a tuple $\hat{\mathcal{S}_\mathcal{A}}(1) = \left(\hat{G}(1), \hat{I}(1), \hat{R}(1), \hat{V}(1)\right)$ where*

- $\hat{G}(1) = L_1 \times \hat{G}$;
- $\hat{I}(1) = I_1 \times \hat{I}$;
- $\hat{R}(1) \subseteq \hat{G}(1) \times \hat{G}(1)$ *is defined as* $(\gamma, \gamma') \in \hat{R}(1)$ *iff* $(\gamma.\hat{a}, \gamma'.\hat{a}) \in \hat{R}$ *and* $(\gamma.c, opinion(\gamma'.c)) = \gamma'.c$;
- $\hat{V}(1) : \hat{G}(1) \to \mathcal{P}(AP)$ *is defined as* $p(o) \in \hat{V}(1)(\gamma)$ *iff* $opinion(\gamma.c) = o$ *and* $p(o) \in \hat{V}(\gamma.\hat{a})$.

Finally we consider group and common knowledge. Since the abstract model is composed of exactly one agent and a concrete system is composed of arbitrarily many agents, it is easy to see that the group and common knowledge modalities are not necessarily preserved from a concrete system to the abstract model. To alleviate this problem, we abstract the satisfaction relation for $\sim_{E_\Gamma}$ and $\sim_{C_\Gamma}$ as follows.

Definition 3.5 (Abstract $\sim_{\hat{E_\Gamma}}$). *The abstract relation for group knowledge $\sim_{\hat{E_\Gamma}} \subseteq \hat{G}(1) \times \hat{G}(1)$ is defined as $(\gamma, \gamma') \in \sim_{\hat{E_\Gamma}}$ iff either one of the following holds:*

- $\gamma \sim_{A_1} \gamma'$;
- *there is a template local state $l \in L$ such that $(w, l) \in \gamma$ and $(w', l) \in \gamma'$ for some weights w, w'.*

Intuitively, two abstract states are $\sim_{\hat{E_\Gamma}}$-related iff they have concrete representatives that are $\sim_i$-related for an arbitrary agent A_i.

Definition 3.6 (Abstract $\sim_{\hat{C_\Gamma}}$). *The abstract relation for common knowledge $\sim_{\hat{C_\Gamma}} \subseteq \hat{G}(1) \times \hat{G}(1)$ is defined as the transitive closure of $\sim_{\hat{E_\Gamma}}$.*

The abstract satisfaction relation $\models_{ab}$ is defined for group knowledge as $(\hat{\mathcal{S}_\mathcal{A}}, \gamma) \models_{ab} E_\Gamma \phi$ iff for all γ' with $\gamma \sim_{\hat{E_\Gamma}} \gamma'$, we have that $(\hat{\mathcal{S}_\mathcal{A}}, \gamma') \models_{ab} \phi$; for common knowledge it is defined as $(\hat{\mathcal{S}_\mathcal{A}}, \gamma) \models_{ab} C_\Gamma \phi$ iff for all γ' with $\gamma \sim_{\hat{C_\Gamma}} \gamma'$, we have that $(\hat{\mathcal{S}_\mathcal{A}}, \gamma') \models_{ab} \phi$; $\models_{ab}$ is defined for the other cases as in Definition 2.4.

We now establish a correspondence between the concrete systems and the abstract model. Specifically, we show that the abstract model simulates every concrete system. Additionally, we show that there is a concrete system that simulates the abstract model. By means of the former result, the satisfaction of an ACTLK formula on the abstract model entails the satisfaction of the formula on every concrete system. Conversely, by means of both results, the falsification of an ACTLK formula on the abstract model implies the existence of a concrete system that falsifies the formula. Consequently, the PMCP is reduced to checking the abstract model against a given specification. We begin by defining the notion of simulation between a concrete system and the abstract model.

Definition 3.7 (Simulation). *A relation $\mathcal{R} \subseteq G(n) \times \hat{G}(1)$ is a simulation between a concrete system $\mathcal{S}_\mathcal{A}(n)$ and the abstract model $\hat{\mathcal{S}_\mathcal{A}}(1)$ if the following conditions hold:*

1. *For every $g \in I(n)$, there is a $\gamma \in \hat{I}(1)$ with $(g, \gamma) \in \mathcal{R}$;*

   Whenever $(g, \gamma) \in \mathcal{R}$, then

2. $V(n)(g) = \hat{V}(1)(\gamma)$;
3. *If $(g, g') \in R(n)$ for some $g' \in G(n)$, then there is a $\gamma' \in \hat{G}(1)$ such that $(\gamma, \gamma') \in \hat{R}(1)$ and $(g', \gamma') \in \mathcal{R}$;*
4. *If $g \sim_{A_1} g'$ for some $g' \in G(n)$, then there is a $\gamma' \in \hat{G}(1)$ such that $\gamma \sim_{A_1} \gamma'$ and $(g', \gamma') \in \mathcal{R}$.*
5. *If $g \sim_{A_i} g'$ for some i with $2 \leq i \leq n$ and some $g' \in G(n)$, then there is a $\gamma' \in \hat{G}(1)$ such that $\gamma \sim_{\hat{E_\Gamma}} \gamma'$ and $(g', \gamma') \in \mathcal{R}$.*

We say that the abstract model $\hat{\mathcal{S}_\mathcal{A}}(1)$ simulates a concrete system $\mathcal{S}_\mathcal{A}(n)$ if there is a simulation relation between $\mathcal{S}_\mathcal{A}(n)$ and $\hat{\mathcal{S}_\mathcal{A}}(1)$. ACTLK formulae are preserved from the abstract model to the concrete system being simulated.

Lemma 3.8. *Assume that $\hat{\mathcal{S}_\mathcal{A}}(1)$ simulates $\mathcal{S}_\mathcal{A}(n)$. Then, $\hat{\mathcal{S}_\mathcal{A}}(1) \models_{ab} \phi$ implies $\mathcal{S}_\mathcal{A}(n) \models \phi$, for any ACTLK formula ϕ.*

Proof Sketch. Assume a simulation relation $\mathcal{R}$ between $\mathcal{S}_\mathcal{A}(n)$ and $\hat{\mathcal{S}_\mathcal{A}}(1)$. We show that

$$(g, \gamma) \in \mathcal{R}, \gamma \models_{ab} \phi \text{ implies } g \models \phi$$

by induction on ϕ. $\phi \in AP$: from simulation requirement 2; $\phi = AX\psi$, $\phi = A(\phi U \psi)$, $\phi = A(\phi R \psi)$: from requirement 3 [3]; $\phi = K_{A_1}\phi$: from requirement 4; $\phi = E_\Gamma \psi$: from requirements 4 and 5. Let $\phi = C_\Gamma \psi$ and assume $\gamma \models_{ab} \phi$. We have to show that $g \models \phi$. Let g' with $g \sim_{C_\Gamma} g'$. Then, there is a sequence $g^1 g^2 \ldots g^k$ such that $g = g^1, g' = g^k$ and for all i with $1 \leq i < k$, there is an agent $A_j \in \Gamma$ such that $g^i \sim_{A_j} g^{i+1}$. By requirements 4 and 5, there is sequence $\gamma^1 \gamma^2 \ldots \gamma^k$ such that $\gamma = \gamma^1$ and for all i with $1 \leq i < k$, $\gamma^i \sim_{\hat{E_\Gamma}} \gamma^{i+1}$. Hence, $\gamma \sim_{\hat{C_\Gamma}} \gamma^k$, and therefore $\gamma^k \models_{ab} \psi$. By the inductive hypothesis, $g' \models \psi$, and thus $g \models \phi$.

By the conclusion of the above induction and by simulation requirement 1, the lemma is entailed. □

A simulation relation between the abstract model and a concrete system is similarly defined to Definition 3.7, but swapping the LHS with the RHS in each of the clauses. We say that a concrete system $\mathcal{S}_\mathcal{A}(n)$ simulates the abstract model $\hat{\mathcal{S}_\mathcal{A}}(1)$ if there is a simulation relation between $\hat{\mathcal{S}_\mathcal{A}}(1)$ and $\mathcal{S}_\mathcal{A}(n)$.

Lemma 3.9. *Assume that $\mathcal{S}_\mathcal{A}(n)$ simulates $\hat{\mathcal{S}_\mathcal{A}}(1)$. Then, $\mathcal{S}_\mathcal{A}(n) \models \phi$ implies $\hat{\mathcal{S}_\mathcal{A}}(1) \models_{ab} \phi$, for any ACTLK formula ϕ.*

PROOF. The proof is similar to the proof of Lemma 3.8. □

A concrete system $\mathcal{S}_{\mathcal{A}}(n)$ and the abstract model $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ are said to be simulation equivalent if $\mathcal{S}_{\mathcal{A}}(n)$ simulates $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ and $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ simulates $\mathcal{S}_{\mathcal{A}}(n)$.

LEMMA 3.10. *Assume that $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ simulates $\mathcal{S}_{\mathcal{A}}(n)$ and $\mathcal{S}_{\mathcal{A}}(n)$ simulates $\hat{\mathcal{S}}_{\mathcal{A}}(1)$. Then, $\hat{\mathcal{S}}_{\mathcal{A}}(1) \models_{ab} \phi$ iff $\mathcal{S}_{\mathcal{A}}(n) \models \phi$, for any ACTLK formula ϕ.*

PROOF. ($\Rightarrow$) Lemma 3.8. ($\Leftarrow$) Lemma 3.9. □

We now show that the abstract model simulates every concrete system and we prove the existence of a concrete system that simulates the abstract model.

LEMMA 3.11. *Given an agent template $\mathcal{A}$ and an ACTLK formula ϕ, the following hold:*

1. *for all $n \gg \alpha$, $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ simulates $\mathcal{S}_{\mathcal{A}}(n)$.*
2. *there is $n \gg \alpha$ such that $\mathcal{S}_{\mathcal{A}}(n)$ simulates $\hat{\mathcal{S}}_{\mathcal{A}}(1)$.*

PROOF SKETCH.
(1) Assume $n \gg \alpha$. Define a mapping $\delta_n : G(n) \to \hat{G}(1)$ from concrete states to abstract states as follows: $\delta_n(g) = (g.1, X)$, where

$$X = \left\{ (w, l) : \exists i.\, 2 \leq i \leq n, g.i = l \text{ and } w \approx \frac{\#(g,l)}{n} \right\}$$

and $w \approx x$ whenever w is the nearest weight to x. Assume the relation $\mathcal{R} = \{(g, \gamma) : \delta_n(g) = \gamma\}$. We show that $\mathcal{R}$ is a simulation relation between $\mathcal{S}_{\mathcal{A}}(n)$ and $\hat{\mathcal{S}}_{\mathcal{A}}(1)$. Simulation requirement 1 follows from the definitions of the initial states for the two models. Let $(g, \gamma) \in \mathcal{R}$ be arbitrary. We show simulation requirements 2,3,4,5.

- Requirement 2. From the definition of $\mathcal{R}$.
- Requirement 3. Assume $(g, g') \in R(n)$ for some $g' \in G(n)$. From the small neighbourhood assumption we have that $\hat{P}(\gamma, o) = P(g, o) \pm \epsilon$, for each opinion o. Therefore,

  $$\frac{\#(g', o)}{n} \approx [\%^{false}(\gamma).\, \hat{P}(\gamma, o)] + \%^{true}(\gamma, o)$$

  for each opinion o. The latter entails $(\gamma, \delta_n(g')) \in \hat{R}(1)$. Also, $(g', \delta_n(g')) \in \mathcal{R}$. Therefore the requirement is satisfied.
- Simulation requirement 4. Assume $g \sim_{A_1} g'$ for some $g' \in G(n)$. Let π be a path in $\mathcal{S}_{\mathcal{A}}(n)$ such that $\pi(i) = g$, for some $i \geq 0$. By simulation requirements 1 and 3, there is a path $\hat{\pi}$ in $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ with $\hat{\pi}(i) = \delta_n(\pi(i))$. As such, we have that $\gamma \sim_{A_1} \hat{\pi}(i)$. Consequently the requirement is satisfied.
- Simulation requirement 5. Assume $g \sim_{A_i} g'$ for some i with $2 \leq i \leq n$ and some $g' \in G(n)$. From simulation conditions 1 and 3, $\delta_n(g') \in \hat{G}(1)$. By definition of $\sim_{\hat{E_\Gamma}}$, $\gamma \sim_{\hat{E_\Gamma}} \delta_n(g')$. Obviously, $(g', \delta_n(g')) \in \mathcal{R}$. Hence the requirement is satisfied.

(2) Pick $n \gg \alpha$ such that $\mathcal{S}_{\mathcal{A}}(n)$ admits every initial density of opinions that is represented by the set of abstract initial states. Define $\mathcal{R}$ as above. Then the proof proceeds along the same lines with the proof in (1). □

A consequence of the above is the following.

THEOREM 3.12. *Given an agent template $\mathcal{A}$ and an ACTLK formula ϕ, the following hold:*

1. *$\hat{\mathcal{S}}_{\mathcal{A}}(1) \models \phi$ implies $\forall n \gg \alpha.\, \mathcal{S}_{\mathcal{A}}(n) \models \phi$.*
2. *$\hat{\mathcal{S}}_{\mathcal{A}}(1) \not\models \phi$ implies $\exists n \gg \alpha.\, \mathcal{S}_{\mathcal{A}}(n) \not\models \phi$.*

PROOF. **(1)** By (1) of Lemma 3.11 and by Lemma 3.8. **(2)** From Lemma 3.11 there is $n \gg \alpha$ such that $\mathcal{S}_{\mathcal{A}}(n)$ and $\hat{\mathcal{S}}_{\mathcal{A}}(1)$ are simulation equivalent. Therefore the thesis follows from Lemma 3.10. □

Theorem 3.12 is our main theoretical result. The theorem provides a constructive methodology for solving the PMCP by checking the abstract model against a given specification.

4. APPLICATIONS

We implemented the weighted abstraction methodology presented earlier in `MCMAS-OFP`, an experimental toolkit that we built from `MCMAS-P`, an open-source model checker for the verification of unbounded multi-agent systems [14]. We designed the input language for `MCMAS-OFP`, called `ISPL-OFP`, to allow for the semantic structures considered here. The language closely follows the modular structure of an agent template. In particular, a template's declaration includes declarations of the template's opinions and their qualities, its transition function, and its neighbourhood size. Figure 4.1 exemplifies `ISPL-OFP` on the protocol described below.

Given an input description for an agent template, `MCMAS-OFP` constructs the abstract model which it encodes symbolically. The base model-checker `MCMAS` [19] is then called to verify the abstract model against the given specifications. Following this, the user can conclude as per Theorem 3.12 whether the specifications hold on a swarm of any size that satisfies the small neighbourhood assumption. We refer to [20] for more details.

4.1 A majority rule protocol

To evaluate our approach, we consider a majority rule protocol put forward to solve the best-of-n decision problem [30]. The protocol assumes two opinions where each opinion corresponds to a spatial area associated with certain resources that determine its quality. Upon exploring a site, an agent determines the site's quality. Then the agent returns to the *nest* where it engages in the opinion formation protocol. According to the protocol, the agents can either be in a *dissemination state* or in an *exploration state*. In the former case, the agents move around the nest while maintaining a well-mixed spatial distribution. Additionally, they broadcast their opinions by means of wireless sensors of limited range. The time an agent spends in this state is proportional to the quality of the opinion it currently holds. Before an agent goes into the exploration state, it updates its opinion according to the majority rule. In the exploration state, the agent leaves the nest to explore the site associated with its current opinion. The site is explored for a period of time that is proportional to its quality. Afterwards, the agent returns to the nest.

We encode the above scenario in the formalism of OFSs. We represent the dissemination state by means of template states that are not in latent state, and we express the exploration state using template states that are in latent state.

```
Agent Template
 Opinions = {A,B};
 Qualities = {A->8, B->4};
 NeighbourhoodSize = 25;
 Evolution:
  opinion = majority and lvalue = lvalue-1 and
                    lvalue>0 and latent=false;
  opinion = majority and lvalue = 8 and
            latent=true if opinion=A and lvalue=0
            and latent=false;
  opinion = majority and lvalue = 4 and
            latent=true if opinion=B and lvalue=0
            and latent=false;
  lvalue = lvalue-1 if lvalue>0 and latent=true;
  lvalue = 8 and latent = false if opinion=A and
                       lvalue=0 and latent=true;
  lvalue = 4 and latent = false if opinion=B and
                       lvalue=0 and latent=true;
```

Figure 1: ISPL-OFP snippet.

The ratio of the qualities for the sites, as well as the neighbourhood size given below correspond to the robot experiments performed in [30]. The agent template $\mathcal{A} = (O, h, \alpha, t)$ is defined as follows:

- $O = \{A, B\}$, where A, B represent the two sites.
- $h(A) = 8, h(B) = 4$. Thus site A is twice as good as site B.
- $\alpha = 25$.

Finally, the template transition relation $t : L \times O \to L$ is defined by:

- $t((o, v, \mathit{false}), o') = (o', v - 1, \mathit{false})$ if $v > 0$.
- $t((o, v, \mathit{false}), o') = (o', h(o'), \mathit{true})$ if $v = 0$.
- $t((o, v, \mathit{true}), o) = (o, v - 1, \mathit{true})$ if $v > 0$.
- $t((o, v, \mathit{true}), o) = (o, h(o), \mathit{false})$ if $v = 0$.

So whenever the template changes its latent status, the latent value is set to the quality of the currently held opinion. It then decreases at each time step until it reaches 0 at which point the agent switches its latent status. As such, the period of time an agent spends in the dissemination (exploration, respectively) state is here modelled by the number of time steps the template is in non-latent (latent, respectively) state.

We used MCMAS-OFP to verify the above protocol. The specifications checked and found to be true were $\phi_1, \ldots, \phi_{10}$, as introduced in Section 2. By means of these results we conclude that not only the protocol reaches consensus, but it also reaches consensus on the opinion with the highest quality, namely site A. Additionally, not only every agent knows this, but it is also common knowledge among the swarm that consensus is eventually reached.

The construction of the abstract model and its verification against all formulae took approximately 5 seconds on an Intel Core i7 machine clocked at 3.4 GHz, with 7.7 GiB cache, running 64-bit Fedora 20, kernel 3.16.6. MCMAS-OFP and the ISPL-OFP file encoding the scenario are available at [20].

5. CONCLUSIONS AND RELATED WORK

In this paper we investigated the formal verification of consensus protocols in swams. We put forward templates to model the behaviour of the agents in an opinion dynamics setting and formulated their verification in terms of the associated parameterised model checking problem. While this is undecidable in general, we built a finite abstract model that we showed to formally simulate any swarm of any size under very permissive conditions. As we proved, the abstract model encodes any concrete instantiation of the swarm and can be used to verify its properties. We presented a toolkit that can be used to generate said abstract model automatically and verify opinion formation protocols. Indeed, we used the toolkit to verify the correctness a majority rule protocol in swarms.

The key aspect of this work is that we work at protocol level and not at system level. In other words, we do not just assess a particular system instantiation; but evaluate the whole class of swarm instances following a consensus protocol. We are not aware of other work in the literature that is based on parameterised model checking. As stated in the introduction, swarm protocols, including consensus protocols, are typically analysed either via simulation [6, 30] or via optimisation methods [28, 29]. By means of optimisation techniques one can typically evaluate the behaviour of the system with very large number of components; whereas simulation approaches are limited by the size of the population under analysis. Equally, model checking techniques for swarm systems are typically limited by the number of agents in the swarm [25, 31, 7, 13]. In this paper we set out to overcome these limitations by providing a first formal, yet completely automatic approach, to the problem. The model we put forward is general enough to model all consensus protocols that follow various forms of the majority rule. While these protocols are normally defined in probabilistic terms, we showed that a purely discrete analysis that merely accounts for the possible evolutions of the system can provide considerable insight in the protocol.

The results here presented builds upon recent work in which the foundations of parameterised verification for multi-agent systems were laid out [15]. This paper differs from that work in several key aspects. Firstly, we here investigate concrete protocols and not just arbitrary multi-agent interactions. This requires the definition of novel, specialised models and appropriate templates. Secondly, given the different semantics the cut-off results presented in [15] cannot be applied; instead we used an ad-hoc construction of an abstract model to simulate all possible behaviours of the system. Thirdly, while the toolkit we released is based on MCMAS-P and we reuse its routines for parsing the files and performing symbolic model checking operations, the key aspect of the implementtaion that we presented is its support for the automatic construction of the abstract model. This has correspondences in work in counter-abstraction, including in [14], but the technical details of abstracted model are entirely different.

In the future we intend to work on other swarm protocols in order to ascertain whether they can also be analysed by means of parameterised model checking and appropriate abstractions.

REFERENCES

[1] K. Apt and D. C. Kozen. Limits for automatic verification of finite-state concurrent systems. *Information Processing Letters*, 22(6):307–309, 1986.

[2] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm intelligence*. Oxford University Press, 1999.

[3] M. C. Browne, E. M. Clarke, and O. Grumberg. Characterizing finite Kripke structures in propositional temporal logic. *Theoretical Computer Science*, 59:115–131, 1988.

[4] E. Clarke, O. Grumberg, and M. Browne. Reasoning about networks with many identical finite state processes. *Information and Computation*, 81(1):13–31, 1989.

[5] E. M. Clarke, O. Grumberg, and D. A. Peled. *Model Checking*. The MIT Press, Cambridge, Massachusetts, 1999.

[6] M. de Oca, , E. Ferrante, A. Scheidler, C. C. Pinciroli, M. Birattari, and M. Dorigo. Majority-rule opinion dynamics with differential latency: a mechanism for self-organized collective decision-making. *Swarm Intelligence*, 5(3-4):305–327, 2011.

[7] C. Dixon, A. Winfield, M. Fisher, and C. Zeng. Towards temporal verification of swarm robotic systems. *Robotics and Autonomous Systems*, 60(11):1429–1441, 2012.

[8] D. Dolev, C. Dwork, and L. Stockmeyer. On the minimal synchronism needed for distributed consensus. *Journal of the ACM (JACM)*, 34(1):77–97, 1987.

[9] E. A. Emerson and K. S. Namjoshi. Automatic verification of parameterized synchronous systems. In *Proceedings of the 8th International Conference one Computer Aided Verification (CAV96)*, volume 1102 of *Lecture Notes in Computer Science*, pages 87–98. Springer, 1996.

[10] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, 1995.

[11] M. Fischer, N. Lynch, and M. Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM (JACM)*, 32(2):374–382, 1985.

[12] S. Garnier, J. Gautrais, M. Asadpour, C. Jost, and G. Theraulaz. Self-organized aggregation triggers collective decision making in a group of cockroach-like robots. *Adaptive Behavior*, 17(2):109–133, 2009.

[13] S. Konur, C. Dixon, and M. Fisher. Analysing robot swarm behaviour via probabilistic model checking. *Robotics and Autonomous Systems*, 60(2):199–213, 2012.

[14] P. Kouvaros and A. Lomuscio. A counter abstraction technique for the verification of robot swarms. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI15)*, pages 2081–2088. AAAI Press, 2015.

[15] P. Kouvaros and A. Lomuscio. Verifying emergent properties of swarms. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI15)*, pages 1083–1089. AAAI Press, 2015.

[16] P. Kouvaros and A. Lomuscio. Parameterised verification for multi-agent systems. *Artificial Intelligence*, 234:152–189, 2016.

[17] P. Krapivsky and S. Redner. Dynamics of majority rule in two-state interacting spin systems. *Physical Review Letters*, 90(23):238701, 2003.

[18] R. Lambiotte, J. Saramäki, and V. Blondel. Dynamics of latent voters. *Physical Review E*, 79(4):046107, 2009.

[19] A. Lomuscio, H. Qu, and F. Raimondi. MCMAS: A model checker for the verification of multi-agent systems. In *Proceedings of the 21th International Conference on Computer Aided Verification (CAV09)*, volume 5643 of *Lecture Notes in Computer Science*, pages 682–688. Springer, 2009.

[20] MCMAS-OFP. Model Checking Opinion Formation Protocols `http://tinyurl.com/q4k69j5`, 2015.

[21] C. Parker and H. Zhang. Cooperative decision-making in decentralized multiple-robot systems: The best-of-n problem. *IEEE/ASME Transactions on Mechatronics*, 14(2):240–251, 2009.

[22] C. Parker and H. Zhang. Collective unary decision-making by decentralized multiple-robot systems applied to the task-sequencing problem. *Swarm Intelligence*, 4(3):199–220, 2010.

[23] W. Penczek and A. Lomuscio. Verifying epistemic properties of multi-agent systems via bounded model checking. *Fundamenta Informaticae*, 55(2):167–185, 2003.

[24] W. Penczek and A. Lomuscio. Verifying epistemic properties of multi-agent systems via bounded model checking. In *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multi-agent systems (AAMAS03)*, pages 209–216. IFAAMAS, 2003.

[25] C. R. r, A. Vanderbilt, M. Hinchey, W. Truszkowski, and J. Rash. Properties of a formal method for prediction of emergent behaviors in swarm-based systems. In *Proceedings of the 2nd International Conference on Software Engineering and Formal Methods (SEFM04)*, pages 24–33. IEEE, 2004.

[26] E. Şahin. Swarm robotics: From sources of inspiration to domains of application. In *Proceedings of the 2004 international conference on Swarm Robotics (SAB04)*, volume 3342 of *Lecture Notes in Computer Science*, pages 10–20. Springer, 2005.

[27] E. Şahin and A. Winfield. Special issue on swarm robotics. *Swarm Intelligence*, 2(2):69–72, 2008.

[28] A. Scheidler. Dynamics of majority rule with differential latencies. *Physical Review E*, 83(3):031116, 2011.

[29] G. Valentini, H. Hamann, and M. Dorigo. Self-organized collective decision making: The weighted voter model. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 45–52. IFAAMAS Press, 2014.

[30] G. Valentini, H. Hamann, and M. Dorigo. Efficient decision-making in a self-organizing robot swarm: On the speed versus accuracy trade-off. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pages 1305–1314. IFAAMAS Press, 2015.

[31] A. Winfield, J. Sa, M. Fernández-Gago, C. Dixon, and

M. Fisher. On formal specification of emergent behaviours in swarm robotic systems. *International journal of advanced robotic systems*, 2(4):363–370, 2005.