Quantified Epistemic Logics with Flexible Terms

F. Belardinelli, A. Lomuscio

Keywords: Knowledge representation and Logic: knowledge representation - Mathematical Logic: modal logic

ABSTRACT. We present a family of quantified epistemic logics for reasoning about knowledge in multi-agent systems. The language enjoys flexible terms with different denotations depending on the epistemic context in which they are interpreted. We present syntax and semantics of the language formally and show completeness of an axiomatisation. We discuss the expressive features of the language by means of an example.

1 Introduction

Propositional modal languages have been extensively used to reason about attitudes of multi-agent systems (MAS)[1]. In particular, a successful field of investigation in Artificial Intelligence and philosophy is the one of epistemic logic [FHMV95, MH95, PR85]. Several frameworks have been explored to reason about various notions of knowledge (implicit, explicit, algorithmic, etc.) either in isolation or in combination with time (discrete or dense, branching or linear, etc.). A wealth of results covering axiomatisability, decidability and computational complexity of various underlying semantical classes (synchronous, no-learning, perfect recall systems) have been made available, as well as model checking techniques for automatic verification [GvdM04, PL03, RL05]. However, little attention has so far been devoted to the extensions to first-order. Although quantified modal logic is ridden with technical difficulties, the power of full quantification is necessary if we wish to represent properties of individuals and relationships between objects and agents (as in "Robot a knows that all wheels of all other robots but b are faulty"). In addition, it is known that first-order modal logic allows for additional expressivity, including being able to distinguish between *de*

[1]The research described in this paper is partly supported by the European Commission Framework 6 funded project CONTRACT (IST Project Number 034418).

re and *de dicto* knowledge. A major technical challenge in using first-order modal languages for modelling MAS is axiomatisability, as many first-order modal logics are not axiomatisable [Gar84, HWZ00, WZ01]. In this paper we aim at extending the current state of the art by introducing a family of provably complete first-order epistemic logics with *global* and *local* terms. While the denotation of the former is rigid, i.e., it is the same in every computational state, the latter's depends on the state in which these expressions are evaluated. The importance of this distinction has long been recognized [FM99, MP92], but it is also well-known that local terms increase the expressive power of first-order modal languages; as a result, certain frameworks are incomplete [Gar84, Sza86]. Our first aim is to retain both local and global terms, without incurring in unaxiomatisability, by suitably restricting the range of local terms according to the substantial interpretation of quantified modal logic in [Gar84]. Further, in the language presented below we allow each agent to reason about a possibly different set of objects. This choice is motivated by the fact that, as agents are autonomous, they may be aware of only a subset of all the existing individuals, possibly different from those of the other agents. In what follows we define systems of global states fulfilling both features above. Then, we present a first-order epistemic language for describing these structures and a sound and complete axiomatisation.

2 Systems of Global States and Equivalence Frames

In this section we present systems of global states [FHMV95, HF89] and Kripke frames [BdRV01, CZ97]. We adopt the "static" perspective on the systems of global states [LR97], rather than their "dynamic" version [FHMV95]. Altough the evolution of knowledge over time is worth studying, for simplicity we do not consider transitions explicitly. We assume a set of agents $A = \{1, \ldots, n\}$ and for $i \in A$, a set L_i of local states $l_i, l'_i, \ldots$, and a set L_e of local states $l_e, l'_e, \ldots$ for the environment.

DEFINITION 1 (SGS). A system of global states is a 5-ple $\mathcal{S} = \langle S, D, \{D_i\}_{i \in A}, F, \{F_i\}_{i \in A} \rangle$ such that $S \subseteq L_e \times L_1 \times \ldots \times L_n$ is a non-empty set of global states; D is a non-empty set of individuals and for $i \in A$, D_i is a possibly empty subset of D; F is a non-empty set of functions from S to D and for $i \in A$, $F_i : S \to D$ is a possibly empty subset of F. $\mathcal{SGS}$ is the class of the systems of global states.

Remarks: This definition fulfils both features referred to in the introduction. Intuitively, we can assign a fixed meaning to global terms like "π" by including the corresponding real number in the domain D of individuals. Expressions like "the tallest man in the world" will instead be modelled by a function $f \in F$, which returns the tallest man in a particular situation.

As to the second feature, we have possibly different domains D_i and F_i, for each agent $i \in A$. Note that the various D_i and F_i are independent from global states. Although this assumption is consistent with the external account of knowledge, it can be relaxed: just assume that for $i \in A$, $s \in S$, $D_{i,s}$ is a possibly empty subset of D, and $F_{i,s} \subseteq F$ is a possibly empty domain of functions from S to D. Henceforth we refer to these structures as *varying systems of global states*. We will show that our results apply to these structures with minor changes.

In order to study the formal properties of SGSs we introduce a particular class of Kripke frames.

DEFINITION 2. An equivalence frame is a 6-tuple $\mathcal{F} = \langle W, \{\sim_i\}_{i\in A}, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}\rangle$ such that W is a non-empty set; for $i \in A$, $\sim_i$ is an equivalence relation on W; D is a non-empty set of individuals and for $i \in A$, D_i is a possibly empty subset of D; F is a non-empty set of functions from W to D and for $i \in A$, $F_i : W \to D$ is a possibly empty subset of F. The class of all equivalence frames is denoted by $\mathcal{F}_E$.

Remarks: It is straightforward to consider *varying equivalence frames*, where for $i \in A$, $w \in W$, $D_{i,w}$ and $F_{i,w}$ are possibly empty subsets of D and F respectively. Finally, note that the individuals in D of both SGSs and equivalence frames can be seen as constant functions from S (resp. W) to D itself. This remark will be useful when interpreting individual terms.

3 Maps between $\mathcal{SGS}$ and $\mathcal{F}_E$

We explore the relationship between SGSs and equivalence frames through the maps $f : \mathcal{SGS} \to \mathcal{F}_E$ and $g : \mathcal{F}_E \to \mathcal{SGS}$. By the lemma below and Theorem 16 we will show that the axiomatisation of equivalence frames in Section 8 is sound and complete also for $\mathcal{SGS}$.

We now show that every equivalence frame $\mathcal{F} = \langle W, \{\sim_i\}_{i\in A}, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}\rangle$ is isomorphic to $f(g(\mathcal{F})) = \langle W', \{\sim'_i\}_{i\in A}, D', \{D'_i\}_{i\in A}, F', \{F'_i\}_{i\in A}\rangle$, that is, there are bijections between W and W', between D and D', between F and F', between the various D_i and D'_i, and between the various F_i and F'_i. In addition, $w \sim_i w'$ iff $(f \circ g)(w) \sim'_i (f \circ g)(w')$.

We start with the map f. Let $\mathcal{S} = \langle S, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}\rangle$ be an SGS, define $f(\mathcal{S})$ as $\langle S, \{\sim_i\}_{i\in A}, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}\rangle$ where S, D, $\{D_i\}_{i\in A}$, F, and $\{F_i\}_{i\in A}$ are the same as in $\mathcal{S}$, and for $i \in A$, the relation $\sim_i$ on S such that $\langle l_e, l_1, \ldots, l_n\rangle \sim_i \langle l'_e, l'_1, \ldots, l'_n\rangle$ iff $l_i = l'_i$ is an equivalence relation. Clearly, $f(\mathcal{S})$ is an equivalence frame.

For the converse map g, let $\mathcal{F} = \langle W, \{\sim_i\}_{i\in A}, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}\rangle$ be an equivalence frame. For every equivalence relation $\sim_i$, for $w \in W$, let the equivalence class $[w]_{\sim_i} = \{w' | w \sim_i w'\}$ be a local state of agent

i; while W is the set of local states for the environment. Let $g(\mathcal{F}) = \langle S, D, \{D_i\}_{i\in A}, F', \{F'_i\}_{i\in A}, \rangle$, where S contains all the $n{+}1$-tuples $\langle w, [w]_{\sim_1}, \ldots, [w]_{\sim_n}\rangle$; D and $\{D_i\}_{i\in A}$ are the same as in $\mathcal{F}$, and each F'_i is the set of functions f' such that $f'(\langle w, [w]_{\sim_1}, \ldots, [w]_{\sim_n}\rangle) = f(w)$, for $f \in F_i$. The structure $g(\mathcal{F})$ is an SGS and the composition of maps is an isomorphism.

LEMMA 3. *Every equivalence frame $\mathcal{F}$ is isomorphic to $f(g(\mathcal{F}))$*

The proof directly extends the propositional case [LS03], so we omit it.

It is straightforward to extend the maps f, g to cover also *varying* systems of global states and equivalence frames: let $\pi_i(\langle a_1, \ldots, a_n\rangle) = a_i$, for $i \leq n$; define f as above and g' as g but $D'_{i,s} = D_{i,w}$ and $F'_{i,s} = \{f' | f'(s') = f(w')$, for $f \in F_{i,w}, w' = \pi_1(s')\}$, whenever $\pi_1(s) = w$. It is easy to check that Lemma 3 holds also for the structures with varying domains.

4 Syntax

Our first-order epistemic formulas are defined on an alphabet containing global variables $x_1, x_2, \ldots$, local variables $z_1, z_2, \ldots$, global constants $c_1, c_2, \ldots$, and local constants $d_1, d_2, \ldots$. Moreover, we have $n{+}1$-ary function symbols $f_1^{n+1}, f_2^{n+1}, \ldots$, and n-ary predicative constants $P_1^n, P_2^n, \ldots$, for $n \in \mathbb{N}$, the identity $=$, the propositional connectives $\neg$ and $\rightarrow$, the universal quantifier $\forall$, and for every $i \in A$, the epistemic operator K_i and the unary predicative constant Adm_i. Terms and formulas in the language $\mathcal{L}_n$ are formally defined as follows:

$$t ::= x \mid z \mid c \mid d \mid f^k(t_1, \ldots, t_k)$$
$$\phi ::= P^k(t_1, \ldots, t_k) \mid t = t' \mid Adm_i(t) \mid \neg\phi \mid \phi \rightarrow \phi' \mid K_i\phi \mid \forall x\phi \mid \forall z\phi$$

The symbols $\bot$, $\wedge$, $\vee$, $\leftrightarrow$, $\exists$ are defined standardly, while $y, y', \ldots$ refer to (local and global) variables in $\mathcal{L}_n$. A global term s is defined as follows:

$$s ::= x \mid c \mid f^k(s_1, \ldots, s_k)$$

otherwise, it is local. The metavariables $s, s', \ldots$ and $u, u', \ldots$ refer to global and local terms respectively; while $v, v', \ldots$ and $r, r', \ldots$ refer to global and local closed terms where no variable appears. The sign "s" can represent either a state or a global term, the context will disambiguate.

$t[\vec{y}]$ (resp. $\phi[\vec{y}]$) means that $\vec{y} = y_1, \ldots, y_n$ are all the (local and global) free variables in t (resp. ϕ); while $t[\vec{y}/\vec{t}]$ (resp. $\phi[\vec{y}/\vec{t}]$) denotes the term (resp. formula) obtained by simultaneously substituting some, possibly all, free occurrences of $y_1, \ldots, y_n$ in t (resp. ϕ) with $t_1, \ldots, t_n$, renaming bounded variables if necessary. We stress that local variables are to be substituted by local terms only, while there is no restriction on global terms.

The indexed quantifiers $\forall_i$, $\exists_i$ are defined by restricting the universal and existential quantifier through the predicate Adm_i, that is, $\forall_i y\phi$ and $\exists_i y\phi$

are shorthands for $\forall y(Adm_i(y) \to \phi)$ and $\exists y(Adm_i(y) \land \phi)$ respectively. Intuitively, the predicate Adm_i holds for the admissible individuals for agent i and the quantifiers $\forall_i$, $\exists_i$ range over the individuals considered by i.

We write $GVar$, $LVar$, $GCon$ and $LCon$ to denote the sets of global and local variables, and the sets of global and local constants in $\mathcal{L}_n$ respectively.

5 Semantics

We interpret $\mathcal{L}_n$ on an equivalence frame $\mathcal{F}$ by means of an interpretation I mapping the syntactic features of $\mathcal{L}_n$ into the elements of $\mathcal{F}$.

DEFINITION 4. An equivalence model $\mathcal{M} = \langle \mathcal{F}, I \rangle$ is such that:

- for $c \in GCon$ and $d \in LCon$, $I(c) \in D$ and $I(d) \in F$;
- $I(f^k) : F^k \to F$ is a k-ary function and $I(f^k)(\vec{g})(w) = I(f^k)(g_1(w), \ldots, g_k(w))$;
- $I(P^k, w) \subseteq D^k$; $I(Adm_i, w) = D_i \cup F_i$; $I(=, w)$ is the equality on D.

The global constants are interpreted rigidly. Instead, it can be the case that $I(d)(w)$ is different from $I(d)(w')$, for $w \neq w'$. Each $I(f^k)$ is a function from F^k to F, but if the arguments are constant functions, i.e., elements in D, then also the output belongs to D. The condition on $I(f^k)$ guarantees that it commutes. Finally, Adm_i is an intensional predicate, i.e., its interpretation is $D_i \cup F_i$, not just a subset of D.

Let σ be an assignment, i.e., a function from $GVar$ to D and from $LVar$ to F; the valuation $I^\sigma(t, w)$ of a term t at a world w is defined as follows:

$$
\begin{aligned}
&I^\sigma(x, w) = I^\sigma(x) = \sigma(x)\\
&I^\sigma(z, w) = I^\sigma(z)(w) = \sigma(z)(w)\\
&I^\sigma(c, w) = I^\sigma(c) = I(c)\\
&I^\sigma(d, w) = I^\sigma(d)(w) = I(d)(w)\\
&I^\sigma(f^k(t_1, \ldots, t_k), w) = I(f^k)(I^\sigma(t_1, w), \ldots, I^\sigma(t_k, w)) = I(f^k)(I^\sigma(t_1), \ldots, I^\sigma(t_k))(w)
\end{aligned}
$$

A variant $\sigma\binom{y}{b}$ of an assignment σ assigns $b \in D \cup F$ to y and coincides with σ on all the other variables.

DEFINITION 5. The satisfaction relation $\models$ for $\phi \in \mathcal{L}_n$, $w \in \mathcal{M}$ and an assignment σ is inductively defined as follows:

$(\mathcal{M}^\sigma, w) \models P^k(\vec{t})$	iff	$\langle I^\sigma(t_1, w), \ldots, I^\sigma(t_k, w)\rangle \in I(P^k, w)$
$(\mathcal{M}^\sigma, w) \models t = t'$	iff	$I^\sigma(t, w) = I^\sigma(t', w)$
$(\mathcal{M}^\sigma, w) \models Adm_i(t)$	iff	$I^\sigma(t) \in D_i \cup F_i$
$(\mathcal{M}^\sigma, w) \models \neg\psi$	iff	$(\mathcal{M}^\sigma, w) \not\models \psi$
$(\mathcal{M}^\sigma, w) \models \psi \to \psi'$	iff	$(\mathcal{M}^\sigma, w) \not\models \psi$ or $(\mathcal{M}^\sigma, w) \models \psi'$
$(\mathcal{M}^\sigma, w) \models K_i\psi$	iff	for all $w \in W$, $w \sim_i w'$ implies $(\mathcal{M}^\sigma, w') \models \psi$

$(\mathcal{M}^\sigma, w) \models \forall x\psi$ iff for all $a \in D$, $(\mathcal{M}^{\sigma\binom{x}{a}}, w) \models \psi$
$(\mathcal{M}^\sigma, w) \models \forall z\psi$ iff for all $f \in F$, $(\mathcal{M}^{\sigma\binom{z}{f}}, w) \models \psi$

The truth conditions for the formulas containing the symbols $\bot, \wedge, \vee, \leftrightarrow$, $\exists$ are defined from the ones above, and we can check that:

$(\mathcal{M}^\sigma, w) \models \forall_i x\psi$ iff for all $a \in D_i$, $(\mathcal{M}^{\sigma\binom{x}{a}}, w) \models \psi$
$(\mathcal{M}^\sigma, w) \models \forall_i z\psi$ iff for all $f \in F_i$, $(\mathcal{M}^{\sigma\binom{z}{f}}, w) \models \psi$

As we pointed out, the formula $Adm_i(t)$ means that $I^\sigma(t)$ is among the individuals in $D_i \cup F_i$ admissible for agent i. If we consider *varying* equivalence frames, the definition of satisfaction above is to be modified as follows:

$(\mathcal{M}^\sigma, w) \models Adm_i(t)$ iff $I^\sigma(t) \in D_{i,w} \cup F_{i,w}$
$(\mathcal{M}^\sigma, w) \models \forall_i x\psi$ iff for all $a \in D_{i,w}, (\mathcal{M}^{\sigma\binom{x}{a}}, w) \models \psi$
$(\mathcal{M}^\sigma, w) \models \forall_i z\psi$ iff for all $f \in F_{i,w}, (\mathcal{M}^{\sigma\binom{z}{f}}, w) \models \psi$

A formula $\phi \in \mathcal{L}_n$ is *true at a world* w iff it is satisfied at w by every assignment σ, ϕ is *valid on a model* $\mathcal{M}$ iff it is true at every world in $\mathcal{M}$, ϕ is *valid on a frame* $\mathcal{F}$ iff it is valid on every model on $\mathcal{F}$, ϕ is *valid on a class* $\mathcal{C}$ *of frames* iff it is valid on every frame in $\mathcal{C}$.

Let Δ be a set of formulas in $\mathcal{L}_n$, $\mathcal{M}$ is a model *for* Δ iff every formula in Δ is valid on $\mathcal{M}$. Further, $\mathcal{F}$ is a frame *for* Δ iff every model on $\mathcal{F}$ is a model for Δ. We can now introduce the quantified interpreted systems.

DEFINITION 6 (QIS). Given an SGS $\mathcal{S}$, a quantified interpreted systems is a pair $\mathcal{P} = \langle \mathcal{S}, I \rangle$ such that I is an interpretation of $\mathcal{L}_n$ in $f(\mathcal{S})$.

The notions of satisfaction, truth and validity are defined as above, i.e., let $\mathcal{P}_f = \langle f(\mathcal{S}), I \rangle$ be the equivalence model associated with the quantified interpreted system $\mathcal{P} = \langle \mathcal{S}, I \rangle$, then $(\mathcal{P}^\sigma, s) \models \phi$ iff $(\mathcal{P}^\sigma_f, s) \models \phi$. A formula $\phi \in \mathcal{L}_n$ is valid on a quantified interpreted systems $\mathcal{P}$ iff ϕ is valid on $\mathcal{P}_f$.

The definitions above apply to *varying* systems of global states and equivalence frames as well.

6 Some validities

Since the domains of quantification D_i and F_i are independent from global states, both the Barcan formula and its converse [FM99] are valid in their indexed form on the class $\mathcal{QIS}$ of all QISs, i.e., they hold in every quantified interpreted system:

$\mathcal{QIS} \models \forall_i y K_j \phi \to K_j \forall_i y \phi$ BF$_{i-j}$
$\mathcal{QIS} \models K_j \forall_i y \phi \to \forall_i y K_j \phi$ CBF$_{i-j}$

For the same reason we have also

$\mathcal{QIS} \models Adm_j(t) \rightarrow K_i Adm_j(t)$ NecAdm
$\mathcal{QIS} \models \neg Adm_j(t) \rightarrow K_i \neg Adm_j(t)$ Nec¬Adm

These validities say that each agent knows which are the individuals he and the other agents reason about.

These principles seem rather strong even for an external account of knowledge. After all, we introduced different domains of quantification for expressing that each agent has only a limited access to the totality of individuals. If they know all other agents' domains as well as theirs, the whole construction becomes questionable. Given this, we can focus on *varying* SGSs, where the formulas above fail. But this undermines the agents' knowledge of their own domains. Our solution consists in admitting BF_{i-j} and CBF_{i-j} only for $i = j$. In fact, the equivalences below hold on any *varying* SGS $\mathcal{S}$:

$\mathcal{S} \models \forall_i x K_i \phi \leftrightarrow K_i \forall_i x \phi$ iff $l_i(s) = l_i(s') \Rightarrow D_{i,s} = D_{i,s'}$
$\mathcal{S} \models \forall_i z K_i \phi \leftrightarrow K_i \forall_i z \phi$ iff $l_i(s) = l_i(s') \Rightarrow F_{i,s} = F_{i,s'}$

By restricting our attention to the SGSs satisfying the conditions above, we can model the scenario where an agent knows his domains of quantification, but not necessarily the other agents'. We call these SGSs *regular* and provide a sound and complete axiomatisation also for this class of structures.

A *varying* equivalence frame $\mathcal{F}$ is *regular* iff the corresponding system of global states $g(\mathcal{F})$ is, i.e., iff

$\mathcal{F} \models \forall_i x K_i \phi \leftrightarrow K_i \forall_i x \phi$ iff $w \sim_i w' \Rightarrow D_{i,w} = D_{i,w'}$
$\mathcal{F} \models \forall_i z K_i \phi \leftrightarrow K_i \forall_i z \phi$ iff $w \sim_i w' \Rightarrow F_{i,w} = F_{i,w'}$

For what concerns identity, the following formulas hold for global terms on every quantified interpreted system:

$\mathcal{QIS} \models (s = s') \rightarrow (\phi[x/s] \rightarrow \phi[x/s'])$ Subst
$\mathcal{QIS} \models (s = s') \rightarrow K_i(s = s')$ K_iId
$\mathcal{QIS} \models (s \neq s') \rightarrow K_i(s \neq s')$ K_iDif

but not for local terms. These (in)validities justify the names of *flexible* and *rigid* variables given in [MP92]. For local terms we have only:

$\mathcal{QIS} \models (u = u') \rightarrow (\phi[z/u] \rightarrow \phi[z/u'])$, for atomic ϕ

but

$\mathcal{QIS} \not\models (u = u') \rightarrow (Adm_i(u) \rightarrow Adm_i(u'))$,

as Adm_i is an intensional predicate and it can be that $I^\sigma(u, w) = I^\sigma(u', w)$, $I^\sigma(u) \in F_i$ but $I^\sigma(u') \notin F_i$.

7 A case study: Battlefield

In this paragraph we present a MAS modelled as a quantified interpreted systems, and describe it by means of the language $\mathcal{L}_n$. We start by considering the set of agents $A = \{1, 2, 3, 4\}$, each agent is assigned a quadrant in $\mathbb{Z} \times \mathbb{Z}$ clockwise:

$$D_4 = \{(x, y) \mid x \in \mathbb{Z}^-, y \in \mathbb{Z}^+\} \qquad D_1 = \{(x, y) \mid x, y \in \mathbb{Z}^+\}$$
$$D_3 = \{(x, y) \mid x, y \in \mathbb{Z}^-\} \qquad D_2 = \{(x, y) \mid x \in \mathbb{Z}^+, y \in \mathbb{Z}^-\}$$

Intuitively, the set D_i is the country of agent i. We assume that each agent has 5 military units, whose positions are recorded in his local state. Further, we consider couples $(x, y) \in D_i$ and triples (k, x, y), for $1 \leq k \leq 5$, to express that there is a military unit at (x, y) and that the k^{th} military unit is at (x, y) respectively.

The local state l_i of agent i is a 4-tuple $\langle \alpha_1, \alpha_2, \alpha_3, \alpha_4 \rangle$ such that:

- α_i is a 5-tuple $\langle (1, x, y), \ldots, (5, x', y') \rangle$ with the positions of i's units.
- for $j \neq i$, α_j is a possibly empty sequence of (k, x, y) and (x, y) recording the positions and identities of i's enemies' units.

Finally, l_e is the local state of the environment recording the positions and identities of all the units.

The set S contains the global states $s = \langle l_e, l_1, l_2, l_3, l_4 \rangle$ such that if either $(x, y) \in \alpha_j(l_i)$ or $(k, x, y) \in \alpha_j(l_i)$, then $(k, x, y) \in \alpha_i(l_i)$. So, an agent may not know the position or the identity of an enemy unit, but if she does, she cannot be wrong. Each F_i is the set of functions $mu_{i.k}$, for $1 \leq k \leq 5$, such that $mu_{i.k}(s) = (x, y)$ iff the expression (k, x, y) appears in α_i of $l_i(s)$.

We assume that our language has global and local constants for denoting the individuals in the various D_i and F_i. We use the same notation for syntactic and semantic elements as the former mirror the latter, the context will disambiguate. Finally, D and F contain the real numbers and functions on them.

Let us suppose that the initial state $s = \langle l_e, l_1, l_2, l_3, l_4 \rangle$ - describing the position of the military units at the beginning - is defined as follows:

- $l_1(s) = \langle \langle (1, 2, 2), (2, 6, 5), (3, 2, 7), (4, 4, 12), (5, 7, 9) \rangle, \langle\rangle, \langle\rangle, \langle\rangle \rangle$
- $l_2(s) = \langle \langle\rangle, \langle (1, 3, -3), (2, 7, -2), (3, 6, -5), (4, 3, -6), (5, 8, -9) \rangle, \langle\rangle, \langle\rangle \rangle$
- $l_3(s) = \langle \langle\rangle, \langle\rangle, \langle (1, -3, -3), (2, -3, -6), (3, -6, -3), (4, -6, -8), (5, -8, -5) \rangle, \langle\rangle \rangle$
- $l_4(s) = \langle \langle\rangle, \langle\rangle, \langle\rangle, \langle (1, -4, 4), (2, -3, 9), (3, -7, 7), (4, -5, 12), (5, -8, 11) \rangle \rangle$

The system of global states containing s describes a situation in which the first military unit of agent 1 is positioned at (2,2). In particular, agent 1 knows this fact while agent 3 is uncertain about it:

$(\mathcal{P}, s) \models K_1 \exists_1 z(z = (2,2))$ and also $(\mathcal{P}, s) \models \exists_1 z K_1(z = (2,2))$
$(\mathcal{P}, s) \models \exists_1 z \neg K_3(z = (2,2))$ and also $(\mathcal{P}, s) \models \forall_1 z \neg K_3(z = (2,2))$

Consider now a function $dist$ returning the distance between two points in $\mathbb{Z} \times \mathbb{Z}$ as a real number: $I(dist)((x,y),(x',y')) = \sqrt{(x-x')^2 + (y-y')^2}$. It is easy to check that agent 2 starts with all his military units within a distance of less than 8, and he knows this fact, which is ignored by agent 4:

$$(\mathcal{P}, s) \models \forall_2 z, z' K_2(dist(z,z') < 8) \text{ but } (\mathcal{P}, s) \not\models \forall_2 z, z' K_4(dist(z,z') < 8)$$

Consider a global constant $firedist$ representing the maximum range of fire of the military units and set it to 8, i.e. $I(firedist) = 8$. We can express the fact that units z, z' are within fire range by the formula $dist(z,z') < firedist$, that we abbreviate as $FDist(z,z')$. Further, suppose hostilities break out in our scenario and agent 3 somehow acquires the knowledge that an unidentified 1's military unit is at (2,2). The resulting state s' differs from s only for the local state of agent 3:

- $l_3(s') = \langle\langle(2,2)\rangle, \langle\rangle, \langle(1,-3,-3),(2,-3,-6),(3,-6,-3),(4,-6,-8),(5,-8,-5)\rangle, \langle\rangle\rangle$

As a consequence, agent 3 has *de dicto* knowledge that the first of his units is within the range of enemy fire, even if she does not know this *de re*:

$$(\mathcal{P}, s') \models \exists_3 z K_3 \exists_1 z' FDist(z,z') \quad \text{but} \quad (\mathcal{P}, s') \not\models \exists_1 z' \exists_3 z K_3 FDist(z,z')$$

while agent 1 ignores this fact, let alone that agent 3 knows this:

$$(\mathcal{P}, s') \models \neg K_1 \exists_3 z K_3 \exists_1 z' FDist(z,z')$$

Now suppose that agent 1 discovers unit $mu_{3.1}$'s position; the change in his local state is recorded in the global state s'' differing from s' as to l_1:

- $l_1(s'') = \langle\langle(1,2,2),(2,6,5),(3,2,7),(4,4,12),(5,7,9)\rangle, \langle\rangle, \langle(-3,-3)\rangle, \langle\rangle\rangle$

Then, also agent 1 knows that one of his units is within the range of enemy fire, even if she is uncertain whether agent 3 is aware of this:

$$(\mathcal{P}, s'') \models \exists_1 z K_1 \exists_3 z' FDist(z,z') \text{ but } (\mathcal{P}, s'') \not\models \exists_1 z K_1 \exists_3 z' K_3 FDist(z,z')$$

Obviously, when one of agent i's units is within the enemy fire range, agent i is in a dangerous situation. Thus, our QIS validates the following *de dicto* specification:

$$\forall_i z(K_i \exists_j z' FireDist(z,z') \to Danger(z)), \text{ for } j \neq i$$

which is stronger than this *de re* specification on danger:

$$\forall_i z(\exists_j z' K_i FireDist(z,z') \to Danger'(z)), \text{ for } j \neq i$$

In fact, both agent 1 and agent 3 know to be in $Danger$, but not in $Danger'$:

$(\mathcal{P}, s'') \models \exists_1 z K_1 Danger(z)$ but $(\mathcal{P}, s'') \not\models \exists_1 z K_1 Danger'(z)$
$(\mathcal{P}, s'') \models \exists_3 z K_3 Danger(z)$ but $(\mathcal{P}, s'') \not\models \exists_3 z K_3 Danger'(z)$

They can try to find a way out of this situation either by attacking or by withdrawing. In order to analyse these alternatives, we introduce a knowledge based protocol [FHMV95]. First of all, we define a predicate $Access$ such that $Access((x, y)(x', y'))$ iff $|x - x'| \leq 1$ and $|y - y'| \leq 1$. Intuitively, the set $\{p' | Access(p, p')\}$ contains the points recheable from p in a single move. Moreover, we consider two actions: $ATTACK$ and $MOVE$. The protocol for agent $i \neq j, k, l$ can be written in pseudo-code as follows:

if $K_i Danger(z)$ then
  if $\exists_i x(Access(z, x) \wedge \forall_j z' \forall_k z'' \forall_l z''' K_i(\neg FDist(x, z') \wedge \neg FDist(x, z'') \wedge$
  $\wedge \neg FDist(x, z''')))$ then
    $MOVE(z, x)$
else if $K_i \exists_j z' FireDist(z, z')$ then
  $ATTACK(z')$

This protocol says that if agent i knows that the unit z is in danger, then he has to move it to an area known to be out of the enemy fire range. If it is not possible, then he has to attack first the enemy unit threatening his unit. Note that this protocol is extremely strict, as it requires knowledge of safety before moving. In the present case

$$(\mathcal{P}, s'') \models \exists_1 z(K_1 Danger(z) \wedge \forall_1 x(Access(z, x) \rightarrow K_1 \exists_3 z' FDist(x, z')))$$

thus agent 1 is bound to attack the first unit of agent 3.

By suitably extending our language, we can express interesting topological relationships on the various D_i, like the presence of obstacles. Moreover, we can introduce intensional predicate for describing in detail the characteristics of the military units. We conclude that our language and structures are a sound formalism for modelling agents moving units on a grid.

8 System $Q.S5_n$

The system $Q.S5_n$ on the language $\mathcal{L}_n$ is a first-order multi-modal version of the propositional system $S5$. While resolution and natural deduction systems are more natural when dealing with automated reasoning, for the purpose of the completeness proof Hilbert-style calculi are easier to handle. Moreover, a natural deduction version of $Q.S5_n$ can be easily obtained from the system presented in [FM99] for instance. Hereafter we list the postulates of $Q.S5_n$, note that $\Rightarrow$ is the inference relation between formulas.

DEFINITION 7. The system $Q.S5_n$ on $\mathcal{L}_n$ contains the following schemes of axioms and inference rules:

Taut	every classic propositional tautology
Dist	$K_i(\phi \to \psi) \to (K_i\phi \to K_i\psi)$
T	$K_i\phi \to \phi$
4	$K_i\phi \to K_iK_i\phi$
5	$\neg K_i\phi \to K_i\neg K_i\phi$
MP	$\phi \to \psi, \phi \Rightarrow \psi$
Nec	$\phi \Rightarrow K_i\phi$
Ex	$\forall y\phi \to \phi[y/t]$
Gen	$\phi \to \psi[y/t] \Rightarrow \phi \to \forall y\psi$, for y not free in ϕ
BF_{i-j}	$\forall_i yK_j\phi \to K_j\forall_i y\phi$
CBF_{i-j}	$K_j\forall_i y\phi \to \forall_i yK_j\phi$
Id	$t = t$
Func	$t = t' \to (t''[y/t] = t''[y/t'])$
Subst	$t = t' \to (\phi[y/t] \to \phi[y/t'])$, for atomic ϕ
K_iId	$s = s' \to K_i(s = s')$
K_iDif	$s \neq s' \to K_i(s \neq s')$

The first group of postulates is an axiomatisation of the propositional multi-modal system $S5_n$. Then we have the classic postulates for quantification for both global and local terms. The Barcan formula and its converse guarantee that the domains of admissible individuals are independent from global states. Finally, we have the axioms Id, Func and Subst for all terms, while K_iId and K_iDif hold only for global terms.

We define *proofs* and *theorems* as standard: $\vdash \phi$ means that $\phi \in \mathcal{L}_n$ is a theorem in $Q.S5_n$. Moreover, we say that $\phi \in \mathcal{L}_n$ is *derivable* in $Q.S5_n$ from the set Δ of formulas in $\mathcal{L}_n$ - $\Delta \vdash \phi$ in short - iff there are $\phi_1, \ldots, \phi_n \in \Delta$ such that $\vdash \phi_1 \wedge \ldots \wedge \phi_n \to \phi$.

Among the theorems and derived rules of $Q.S5_n$ we have:

NecAdm	$Adm_i(t) \to K_jAdm_i(t)$
Nec¬Adm	$\neg Adm_i(t) \to K_j\neg Adm_i(t)$
ExAdm	$\forall_i y\phi \to (Adm_i(t) \to \phi[y/t])$
GenAdm	$\phi \to (Adm_i(t) \to \psi[y/t]) \Rightarrow \phi \to \forall_i y\psi$, for y not free in ϕ

For reasons of space we omit the proofs. We can easily check that every equivalence frame is a frame for $Q.S5_n$. As a consequence, we have the following soundness result:

LEMMA 8 (Soundness). *The system $Q.S5_n$ is sound with respect to the class $\mathcal{F}_E$ of equivalence frames.*

By this lemma and the definition of validity on the systems of global states, the following implications hold: $Q.S5_n \vdash \phi \Rightarrow \mathcal{F}_E \models \phi \Rightarrow \mathcal{SGS} \models \phi$. Thus, we have soundness for the systems of global states:

COROLLARY 9 (Soundness). *The system $Q.S5_n$ is sound with respect to the class $\mathcal{SGS}$ of the systems of global states.*

By *varying* and *regular* $Q.S5_n$ we denote the systems obtained from $Q.S5_n$ respectively by eliminating BF_{i-j}, CBF_{i-j} and by restricting these postulates to $i = j$. We have only restricted versions of NecAdm and Nec¬Adm for *regular* $Q.S5_n$. On the other hand, *varying* and *regular* $Q.S5_n$ are sound for *varying* and *regular* equivalence frames, therefore also for *varying* and *regular* SGSs:

LEMMA 10 (Soundness). Varying *$Q.S5_n$ is sound for* varying *equivalence frames, therefore it is sound also for* varying *systems of global states.*

LEMMA 11 (Soundness). Regular *$Q.S5_n$ is sound for* regular *equivalence frames, therefore it is sound also for* regular *systems of global states.*

9 Completeness

The completeness of $Q.S5_n$ with respect to equivalence frames is proved by means of the canonical model method. For the case in hand this technique basically consists in showing the following fact:

If $Q.S5_n$ does not prove a formula $\phi \in \mathcal{L}_n$, then the canonical model *$\mathcal{M}^{Q.S5_n}$ for $Q.S5_n$ does not validate ϕ.*

This result relies on two lemmas: the *saturation* and the *truth lemmas*. Their proofs need the following definitions, where Λ is a set of formulas:

Λ is *consistent* iff $\Lambda \nvdash \bot$;
Λ is *maximal* iff for every $\phi \in \mathcal{L}_n$, $\phi \in \Lambda$ or $\neg\phi \in \Lambda$;
Λ is *max-cons* iff Λ is consistent and maximal;
Λ is *rich* iff $\exists x\phi \in \Lambda \Rightarrow \phi[x/c] \in \Lambda$, for some $c \in GCon$, and $\exists z\phi \in \Lambda \Rightarrow \phi[z/d] \in \Lambda$, for some $d \in LCon$;
Λ is *saturated* iff Λ is max-cons and rich.

We observe that the following lemma holds, we refer to [Gar84] for a proof:

LEMMA 12 (Saturation lemma). *If Δ is a consistent set of the formulas in $\mathcal{L}_n$, then it can be extended to a saturated set Π of formulas on some expansion $\mathcal{L}_n^+$ of $\mathcal{L}_n$.*

If $\nvdash \phi$ then the set $\{\neg\phi\}$ is consistent and by the lemma above we obtain a saturated set $\Pi \supseteq \{\neg\phi\}$. By this remark, the set W of saturated sets $w, w', \ldots$ of formulas in $\mathcal{L}_n^+$ is non-empty.

To introduce the other elements in the canonical model for $Q.S5_n$ we need few more definitions. For closed global terms v, v', define $v \sim_w v'$ iff

$(v = v') \in w$. This is an equivalence relation and $[v]_w = \{v' | v \sim_w v'\}$ is the equivalence class of v in w. Since the accessibility relation in $\mathcal{M}^{Q.S5_n}$ is defined so that wR_iw' iff $\{\phi | K_i\phi \in w\} \subseteq w'$, by K_iId and K_iDif we can show that the definition of $[v]_w$ is independent from w - i.e. wR_iw' implies $[v]_w = [v]_{w'}$ - so we simply write $[v]$. We define $D_{i,w}$ as the set $\{[v] | Adm_i(v) \in w\}$. By Subst, Func, NecAdm and Nec¬Adm we can show that this definition is independent from v and w, therefore we simply write D_i. Further, for every closed local term r define a function f_r such that

$$f_r(w) \quad = \quad \begin{cases} [v] & \text{if there is a } v \text{ such that } (r = v) \in w; \\ \{r' | (r' = r) \in w\} & \text{otherwise.} \end{cases}$$

Each $F_{i,w}$ is the set $\{f_r | Adm_i(r) \in w\}$; by NecAdm and Nec¬Adm this definition is provably independent from w, so we simply write F_i. Finally, the canonical model for $Q.S5_n$ is defined as follows:

DEFINITION 13. The canonical model $\mathcal{M}^{Q.S5_n}$ for $Q.S5_n$ on the language $\mathcal{L}_n$, with an expansion $\mathcal{L}_n^+$, is the 5-tuple $\langle W, \{R_i\}_{i\in A}, D, \{D_i\}_{i\in A}, F, \{F_i\}_{i\in A}, I\rangle$ such that:

- W is the set of saturated sets of formulas in $\mathcal{L}_n^+$;
- wR_iw' iff $\{\phi | K_i\phi \in w\} \subseteq w'$;
- $D = \{[v] | v \in \mathcal{L}_n^+\} \cup \{f_r(w) | r \in \mathcal{L}_n^+, w \in W\}$ and $F = \{f_r | r \in \mathcal{L}_n^+\}$, while D_i, F_i are defined as above;
- I is an interpretation such that:
  - $I(c) = [c]$ and $I(d) = f_d$;
  - for $a_1, \ldots, a_k \in D \cup F$, $I(f^k)(\vec{a})$ is a function such that

$$I(f^k)(\vec{a})(w) \quad = \quad \begin{cases} [f^k(\vec{v})] & \text{if each } a_i = [v_i]; \\ f_{f^k(\vec{e})}(w) & \text{for } e_i = v_i \text{ if } a_i = [v_i], \text{ or } e_i = r_i \text{ if } a_i = f_{r_i} \end{cases}$$

  - for $a_1, \ldots, a_k \in D$, $\langle\vec{a}\rangle \in I(P^k, w)$ iff $P^k(\vec{e}) \in w$, for $e_i = v_i$ and $a_i = [v_i]$ or $e_i = r_i$ and $a_i = f_{r_i}(w)$.

Note that the interpretation of functions and predicates is well defined by the axioms Func and Subst. Moreover, the definition of D guarantees that the symbol $=$ is interpreted as identity. Thus, we can prove the following lemma by extending the propositional case and the remarks above.

LEMMA 14. *The canonical model $\mathcal{M}^{Q.S5_n}$ for $Q.S5_n$ exists and satisfies the constraints on equivalence models.*

If a formula $\phi \in \mathcal{L}_n$ is not provable in *varying* $Q.S5_n$, then we can still construct the canonical model as above. Since neither NecAdm nor Nec¬Adm are theorems in *varying* $Q.S5_n$, we cannot show that $D_{i,w}$ and $F_{i,w}$ are independent from w. But this is not a problem as our canonical model is a *varying* Kripke model anyway. Similarly, the canonical model for *regular* $Q.S5_n$ satisfies the conditions on *regular* models by the restricted versions BF_{i-i} and CBF_{i-i}.

Now let σ be an assignment to local and global variables, we can show that for every $w \in W$, $I^\sigma(t[\vec{y}], w) = I(t[\vec{y}/\vec{e}])(w)$, whenever $\sigma(y_i) = I(e_i)$. By this result the base case of the truth lemma below hold. In what follows we simply write $\mathcal{M}$ for $\mathcal{M}^{Q.S5_n}$.

LEMMA 15 (Truth lemma). *For every* $w \in \mathcal{M}$, $\phi[\vec{y}] \in \mathcal{L}_n^+$, *for* $\sigma(y_i) = I(e_i)$

$$(\mathcal{M}^\sigma, w) \models \phi[\vec{y}] \quad \text{iff} \quad \phi[\vec{y}/\vec{e}] \in w$$

The proof of this lemma relies on the Barcan formula for showing that if $K_i\psi[\vec{y}/\vec{e}] \notin w$, then the consistent set $\{\phi | K_i\phi \in w\} \cup \{\neg\psi[\vec{y}/\vec{e}]\}$ can be extended to a saturated set w' on $\mathcal{L}_n^+$ such that wR_iw' and $(\mathcal{M}^\sigma, w') \models \neg\psi[\vec{y}]$ by the induction hypothesis.

By the truth lemma we conclude that the canonical model is a model for $Q.S5_n$, based on an equivalence frame, falsifying any unprovable formula ϕ. Thus, we state the following completeness result.

THEOREM 16 (Completeness). *The system* $Q.S5_n$ *is complete with respect to the class* $\mathcal{F}_E$ *of equivalence frames.*

We note without proof that Lemma 15 holds also for *varying* and *regular* $Q.S5_n$, therefore these systems are complete with respect to the classes of *varying* and *regular* equivalence frames respectively.

Further, we have completeness also with respect to the systems of global states. In fact, if $\nvdash \phi$ then by Theorem 16 there exists a model $\mathcal{M} = \langle \mathcal{F}, I \rangle$ based on an equivalence frame $\mathcal{F}$, which falsifies ϕ. Define the quantified interpreted system $\mathcal{P}$ as $\langle g(\mathcal{F}), I \rangle$: by definition $\mathcal{P} \models \phi$ iff $\mathcal{P}_f = \langle f(g(\mathcal{F})), I \rangle$ models ϕ, but by Lemma 3 $f(g(\mathcal{F}))$ is isomorphic to $\mathcal{F}$. Hence $\mathcal{P} \not\models \phi$. As a result, we can state the main result of this paper.

COROLLARY 17 (Soundness and Completeness). *For every formula* $\phi \in \mathcal{L}_n$, $\mathcal{SGS} \models \phi$ *iff* $Q.S5_n \vdash \phi$.

Finally, by analogous reasoning we can prove the following results:

COROLLARY 18. Varying $Q.S5_n$ *is sound and complete for the class of* varying *systems of global states.*

COROLLARY 19. Regular $Q.S5_n$ *is sound and complete for the class of* regular *systems of global states.*

10 Conclusions

In this paper we presented a framework for quantified epistemic logics with flexible terms based on quantified interpreted systems, an extension to first-order of interpreted systems, the popular formalism for MAS. The language is very expressive and particularly suited for representing relationships, it also supports full quantification over infinite sets of objects. In Section 6 we pointed out some problems related to agents' knowledge of their domains and of other agents' domains. It is worth noting that these difficulties arise from the specific epistemic interpretation of the modality. We put forward a solution by suitably restricting the validity of the Barcan formula and its converse. By outlining completeness results we have shown that this expressiveness can be achieved while still retaining axiomatibility.

BIBLIOGRAPHY

[BdRV01] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2001.

[CZ97] A. Chagorov and M Zakharyaschev. *Modal Logic*. Clarendon Press, Oxford, 1997.

[FHMV95] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, 1995.

[FM99] M. Fitting and R. L. Mendelsohn. *First-order Modal Logic*. Kluwer Academic Publishers, Dordrecht, 1999.

[Gar84] J. Garson. Quantification in modal logic. In *Handbook of Philosophical Logic*, volume 2, pages 249–307. D. Gabbay and F. Guenthner Eds., 1984.

[GvdM04] P. Gammie and R. van der Meyden. MCK: Model checking the logic of knowledge. In *CAV'04*, volume 3114 of *LNCS*, pages 479–483. Springer-Verlag, 2004.

[HF89] J. Halpern and R. Fagin. Modelling knowledge and action in distributed systems. *Distributed Computing*, 3(4):159–179, 1989.

[HWZ00] I. Hodkinson, F. Wolter, and M. Zakharyaschev. Decidable fragment of first-order temporal logics. *Ann. Pure Appl. Logic*, 106(1-3):85–134, 2000.

[LR97] A. Lomuscio and M. Ryan. On the relation between interpreted systems and Kripke models. In *ESSLLI97*, Aix en Provence, August 1997.

[LS03] A. Lomuscio and M. Sergot. Deontic interpreted systems. *Studia Logica*, 75(1):63–92, 2003.

[MH95] J.-J. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995.

[MP92] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems*, volume 1. Springer-Verlag, 1992.

[PL03] W. Penczek and A. Lomuscio. Verifying epistemic properties of multi-agent systems via bounded model checking. *Fundamenta Informaticae*, 55(2):167–185, 2003.

[PR85] R. Parikh and R. Ramanujam. Distributed processes and the logic of knowledge. In *Logic of Programs*, volume 193 of *LNCS*, pages 256–268. Springer, 1985.

[RL05] F. Raimondi and A. Lomuscio. Automatic verification of multi-agent systems by model checking via OBDDs. *Journal of Applied Logic*, 2005. To appear.

[Sza86] A. Szalas. Concerning the semantic consequence relation in first-order temporal logic. *Theor. Comput. Sci.*, 47(3):329–334, 1986.

[WZ01] F. Wolter and M. Zakharyaschev. Decidable fragments of first-order modal logics. *J. Symb. Log.*, 66(3):1415–1438, 2001.