The Logic of Communication Graphs

Eric Pacuit[1] and Rohit Parikh[2]

[1] Computer Science Department,
The Graduate Center of CUNY,
365 5th Avenue, New York City, NY 10016
epacuit@cs.gc.cuny.edu
www.cs.gc.cuny.edu/∼epacuit*
[2] CS, Math and Philosophy,
Brooklyn College** and The Graduate Center of CUNY,
365 5th Avenue, New York City, NY 10016
rparikh@gc.cuny.edu
www.sci.brooklyn.cuny.edu/∼rparikh

Abstract. In 1992, Moss and Parikh studied a bimodal logic of knowledge and effort called *Topologic*. In this current paper, *Topologic* is extended to the case of many agents who are assumed to have some private information at the outset, but may refine their information by acquiring information possessed by other agents, possibly via yet other agents.

Let us assume that the agents are connected by a *communication graph*. In the communication graph, an edge from agent i to agent j means that agent i can directly receive information from agent j. Agent i can then refine its own information by learning information that j has, including information acquired by j from another agent, k. We introduce a multi-agent modal logic with knowledge modalities and a modality representing communication among agents. We show that the validities of *Topologic* remain valid and that the communication graph is completely determined by the validities of the resulting logic. Applications of our logic to current political dilemmas are obvious.

1 Introduction

In [13], Moss and Parikh introduce a bimodal logic intended to formalize reasoning about points and sets. This new logic called *Topologic* can also be understood as an epistemic logic with an effort modality. Formally, the two modalities are: K and $\Diamond$. The intended interpretation of $K\phi$ is that ϕ is known; and the intended interpretation of $\Diamond\phi$ is that after some amount of effort ϕ becomes true. For example, the formula

$$\phi \rightarrow \Diamond K\phi$$

* Both authors would like to thank Hans van Ditmarsch and the Knowledge, Games and Beliefs Group of CUNY for their comments.

** 2900 Bedford Avenue, Brooklyn, NY 11210. Research of both authors supported under the PSC-CUNY FRAP program.

J. Leite et al. (Eds.): DALT 2004, LNAI 3476, pp. 256–269, 2005.
© Springer-Verlag Berlin Heidelberg 2005

means that if ϕ is true, then after some "work", $K\phi$ becomes true, i.e., ϕ is known. In other words, the formula says that if ϕ is true, then ϕ can be known with some effort. What exactly is meant by "effort" depends on the application. For example, we may think of effort as meaning taking a measurement, performing a calculation or observing a computation. In this paper we will think of effort as meaning consulting some agent's database of known formulas.

There is a temptation to think that the effort modality can be understood as (only) a temporal operator, reading $\Diamond\phi$ as "ϕ is true some time in the future". While there is a connection between the logics of knowledge and time and logics of knowledge and effort (see [8, 9] and references therein for more on this topic), following [13] we will assume that such effort leaves the base facts about the world unchanged. In particular, in any topologic model, if ϕ does not contain any modalities, then $\phi \leftrightarrow \Box\phi$ is valid. Thus, effort will not change the base facts about the world – it can only change knowledge of these facts.

The family of logics introduced in [13] and later studied by Dabrowski, Moss and Parikh, Georgatos, Heinemann, and Weiss ([3, 4, 5, 6, 8, 21]) has a semantics in which the acquisition of knowledge is explicitly represented. Familiar mathematical structures such as subset spaces, topologies, intersection spaces and complete lattices of subsets corresponding to natural notions of knowledge acquisition are attached to standard Kripke structures.

Given a set W, a subset space is a pair $\langle W, \mathcal{O}\rangle$, where $\mathcal{O}$ is a collection of subsets of W. A point $x \in W$ represents a complete description of the world in which all ground facts are settled, whereas a set $U \in \mathcal{O}$ represents an *observation.* The pair (x, U), called a *neighborhood situation*, can be thought of as an actual situation together with an observation made about the situation. Formulas are interpreted at neighborhood situations. Thus the knowledge modality K represents movement within (consistent with) the current observation, while the effort modality $\Diamond$ represents a refining of the current observation.

Formally,

1. $x, U \models K\phi$ iff $(\forall y \in U)(y, U \models \phi)$
2. $x, U \models \Diamond\phi$ iff $(\exists V \in \mathcal{O})((x \in V \subseteq U)$ and $(x, V \models \phi))$

[13] provides a sound and complete axiomatization for all subset spaces. In [4] and [5], Georgatos provides a sound and complete axiomatization for subset spaces that are topological spaces and complete lattices. Dabrowski, Moss, and Parikh prove the same result using an embedding into **S4** ([3]). [6] provides a sound and complete axiomatization for treelike spaces, and Weiss ([21]) has provided a sound and complete axiomatization for intersection-spaces. Interestingly, it is shown in [21] that an infinite number of axiom schemes are necessary for any complete axiomatization of intersection spaces. More recently, Heinemann [8, 9] has looked at subset spaces and logics of knowledge and time, and the connection between hybrid logic and subset spaces [10, 12].

In this paper, we present a multi-agent topologic in which the effort modality $\Diamond$ is intended to mean communication among agents. In order for any communication to take place, we must assume that the agents understand a common language. Thus we assume a set At of propositional variables, understood by all the agents, but with only specific agents knowing their actual values at the start. Letters p, q, etc, will denote elements of At. The agents will have some information – knowledge of the truth values of some elements of At, but refine that information by acquiring information possessed by other agents, possibly via yet other agents. This implies that if agents are restricted in whom they can communicate with, then this fact will restrict the knowledge they can acquire.

Consider the current situation with Bush and Porter Goss, the director of the CIA. If Bush wants some information from a particular CIA operative, say Bob, he must get this information through Goss. Suppose that ϕ is a formula representing the exact whereabouts of Bin Laden, and that Bob, the CIA operative in charge of maintaining this information knows ϕ. In particular, $K_{\mathrm{Bob}}\phi$, but suppose that at the moment, Bush does not know the exact whereabouts of Bin Laden ($\neg K_{\mathrm{Bush}}\phi$). Presumably Bush *can* find out the exact whereabouts of Bin Laden ($\Diamond K_{\mathrm{Bush}}\phi$) by going through Goss, but of course, *we* cannot find out such information ($\neg\Diamond K_{\mathrm{e}}\phi \wedge \neg\Diamond K_{\mathrm{r}}\phi$) since we do not have the appropriate security clearance. Clearly, then, as a *pre-requisite* for Bush learning ϕ, Goss will also have come to know ϕ. We can represent this situation by the following formula:

$$\neg K_{\mathrm{Bush}}\phi \wedge \Box(K_{\mathrm{Bush}}\phi \rightarrow K_{\mathrm{Goss}}\phi)$$

where $\Box$ is the dual of diamond.

Let $\mathcal{A}$ be a set of agents. A **communication graph** is a directed graph $\mathcal{G}_{\mathcal{A}} = (\mathcal{A}, E)$ where $E \subseteq \mathcal{A} \times \mathcal{A}$. Intuitively $(i, j) \in E$ means that i can directly receive information from agent j, but *without* j knowing this fact. Thus an edge between i and j in the communication graph represents a one-sided relationship between i and j. Agent i has access to any piece of information that agent j knows. We have introduced this 'one sidedness' restriction in order to simplify our semantics, but also because such situations of one sided learning occur naturally. A common situation that is helpful to keep in mind is accessing a website. We can think of agent j as creating a website in which everything he *currently* knows is available, and if there is an edge between i and j then agent i can access this website without j being aware that the site is being accessed. Another important application is spying where one person accesses another's information without the latter being aware that information is being leaked. Naturally j may have been able to access some other agent k's website and had updated some of her own information. Therefore, it is important to stress that when i accesses j's website, he is accessing j's current information which may include what k knew initially.

The assumption that i can access all of j's information is a significant idealization from these common situations, but becomes more realistic if we think of this information as being confined to facts expressible as truth functional combinations of some small set of basic propositions. Thus our idealization rests on two assumptions:

1: All the agents share a common language, and
2: The agents make available all possible pieces of information which they know *and which are expressible in this common language.*

2 The Logic of Communication Graphs

In this section we will describe the logic of communication graphs, $\mathcal{K}(\mathcal{G})$. The language will be a multi-agent modal language with a communication modality. The formula $K_i\phi$ will be interpreted as "according to i's current information, i knows ϕ", and $\Diamond\phi$ will be interpreted as "after some communications (which respect the communication graph), ϕ becomes true". Thus for example, the multi-agent version of the formula $\phi \rightarrow \Diamond K\phi$, expressing that if ϕ is true then with some effort ϕ can be known, is

$$K_j\phi \rightarrow \Diamond K_i\phi$$

This formula expresses that if agent j (currently) knows ϕ, then after some communication agent i can come to know ϕ. Let At be a finite set of propositional variables. A well-formed formula of $\mathcal{K}(\mathcal{G})$ has the following syntactic form

$$\phi \ := p \mid \neg\psi \mid \phi \wedge \psi \mid K_i\phi \mid \Diamond\phi$$

where $p \in \mathsf{At}$. We abbreviate $\neg K_i\neg\phi$ and $\neg\Diamond\neg\phi$ by $L_i\phi$ and $\Box\phi$ respectively, and use the standard abbreviations for the propositional connectives ($\vee$, $\rightarrow$, and $\bot$). Let $\mathcal{L}_{\mathcal{K}(\mathcal{G})}$ denote the set of well-formed formulas of $\mathcal{K}(\mathcal{G})$. We also define $\mathcal{L}_0(\mathsf{At})$, (or simply $\mathcal{L}_0$ if At is fixed or understood), to be the set of ground formulas, i.e., the set of formulas constructed from At using $\neg, \wedge$ only.

2.1 Semantics

The semantics presented here combines ideas both from the subset models of [13] and the history based models of Parikh and Ramanajum (see [16, 17]). Suppose that $\mathcal{G} = (\mathcal{A}, E)$ is a fixed communication graph. Given that the agents are initially given some private information and assumed to communicate according to the communication graph $\mathcal{G}$, the semantics in this section is intended to formalize what agents know and may come to know after some communication.

Initially, each agent i knows or is informed (say by nature) of the truth values of a certain subset At_i of propositional variables, and the At_i *as well as this fact are common knowledge.* Thus the other agents know that i knows the truth values of elements of At_i, but, typically, not what these values actually are. We do not need to assume that the At_i are disjoint, nor that the At_i together add up to all of At, although such sub-cases will be of interest. Thus if At_i and At_j intersect then agents i, j will share information at the very beginning. Let W be the set of boolean valuations on At. An element $v \in W$ is called a *state*. We use 1 for the truth value *true*. Initially each agent i is given a boolean valuation $v_i : \mathsf{At}_i \rightarrow \{0, 1\}$. This initial distribution of information among the agents can be represented by a vector $\boldsymbol{v} = (v_1, \ldots, v_n)$. Of course, since we are modeling knowledge and not belief, these initial boolean valuations must be compatible. I.e., for each i, j, v_i and v_j agree on $\mathsf{At}_i \cap \mathsf{At}_j$. Call any vector of partial boolean valuations $\boldsymbol{v} = (v_1, \ldots, v_n)$ **consistent** if for each $p \in dom(v_i) \cap dom(v_j)$, $v_i(p) = v_j(p)$ for all $i, j = 1, \ldots, n$. We shall assume that only such consistent vectors arise as initial information. All this information is common knowledge and only the precise values of the v_i are private.

Definition 1. *Let* At *be a finite set of propositional variables and* $\mathcal{A} = \{1, \ldots, n\}$ *a finite set of agents. Given the distribution of sublanguages* $\mathbf{At} = (\mathsf{At}_1, \ldots, \mathsf{At}_n)$, *an* **initial information vector for At** *is any consistent vector* $\boldsymbol{v} = (v_1, \ldots, v_n)$ *of partial boolean valuations such that for each* $i \in \mathcal{A}$, $dom(v_i) = \mathsf{At}_i$.

We assumed that all initial vectors are consistent, although if we were dealing with beliefs rather than knowledge, then very interesting questions about *in*-consistent initial vectors could arise.

We assume that the only communications that take place are about the physical world. But we do allow agents to learn objective facts which are not atomic, but may be complex, like $p \vee q$ where $p, q \in \mathsf{At}$. Now note that if agent i learned some *literal* from agent j, then there is a simple way to update i's valuation v_i with this new information by just adding the truth value of another propositional symbol. However, if i learns a more general ground formula from agent j, then the situation will be more complex. For instance if the agent knows p and learns $q \vee r$ then the agent now has three valuations on the set $\{p, q, r\}$ which cannot be described in terms of a partial valuation on a subset of At.

Fix a communication graph $\mathcal{G}$ and suppose that agent i learns some ground fact ϕ from agent j. Of course, there must be an edge from agent i to agent j in $\mathcal{G}$. This situation will be represented by the tuple (i, j, ϕ) and will be called a **communication event**. Let $\Sigma_{\mathcal{G}}$ be the set of all possible events. Formally,

Definition 2. *Let* $\mathcal{G} = (\mathcal{A}, E_{\mathcal{G}})$ *be a communication graph. A tuple* (i, j, ϕ), *where* $\phi \in \mathcal{L}_0(\mathsf{At})$ *and* $(i, j) \in E_{\mathcal{G}}$ *is called a* **communication event**. *Then* $\Sigma_{\mathcal{G}} = \{(i, j, \phi) \mid \phi \in \mathcal{L}_0, (i, j) \in E_{\mathcal{G}}\}$ *is the set of all possible communication events (given the communication graph* $\mathcal{G}$*).*

Given the set of events $\Sigma_{\mathcal{G}}$, a **history** is a finite sequence of events. I.e., $H \in \Sigma_{\mathcal{G}}^*$. The empty history will be denoted ϵ. The following notions are standard (see [16, 17] for more information). Given two histories H, H', say $H \preceq H'$ iff $H' = HH''$ for some history H'', i.e., H is an initial segment of H'. Obviously, $\preceq$ is a partial order. If H is a history, and (i, j, ϕ) is a communication event, then H followed by (i, j, ϕ) will be written $H; (i, j, \phi)$. Given a history H, let $\lambda_i(H)$ be i's local history corresponding to H. I.e., $\lambda_i(H)$ is a sequence of events that i can "see". Formally, λ_i maps each event of the form (i, j, ϕ) to itself, and maps other events (m, j, ψ) with $m \neq i$ to the null character while preserving the order among events.

Fix a finite set of agents $\mathcal{A} = \{1, \ldots, n\}$ and a finite set of propositional variables At along with subsets $(\mathsf{At}_1, ..., \mathsf{At}_n)$. A **communication graph frame** is a pair $\langle \mathcal{G}, \mathbf{At} \rangle$ where $\mathcal{G}$ is a communication graph, and $\mathbf{At} = (\mathsf{At}_1, ..., \mathsf{At}_n)$ is an assignment of sub-languages to the agents. A **communication graph model** based on a frame $\langle \mathcal{G}, \mathbf{At} \rangle$ is a triple $\langle \mathcal{G}, \mathbf{At}, \boldsymbol{v} \rangle$, where $\boldsymbol{v}$ is a consistent vector of partial boolean valuations for $\mathbf{At}$.

Now we address two issues. One is that not all histories are legal. For an event (i, j, ϕ) to take place after a history H, it must be the case that after H, j knows ϕ. Clearly i cannot learn from j something which j did not know. Whether a history is justified depends not only on the initial valuation, but also on the set of communications that have taken place prior to each communication in the history.

The second issue is that the information which an agent learns by "reading" a formula ϕ may be *more* than just the fact that ϕ is true. For suppose that i learns $p \vee q$ from j, but j is not connected, directly or indirectly, to anyone who might know the initial truth value of q. In this case i has learned *more* than $p \vee q$, i has learned p as well. For the only way that j could have known $p \vee q$ is if j knew p in which case p must be true. Our definition of the semantics below will address both these issues.

Formulas will be interpreted at pairs (w, H) where w is a state (boolean valuation) and H is a finite sequence of communication events.

We first introduce the notion of i-equivalence among histories. Intuitively, two histories are i-equivalent if those communications which i takes active part in, are the same.

Definition 3. *Let w be a state and H a finite history. Define the relation $\sim_i$ as follows: $(w, H) \sim_i (v, H')$ iff $w_{|\mathsf{At}_i} = v_{|\mathsf{At}_i}$ and $\lambda_i(H) = \lambda_i(H')$.*

Before proceeding further, we summarize the uncertainty faced by each of the agents:

1. Agents may be uncertain about the actual state of the world.
2. Agents may be uncertain about which communications have taken place.

Example: *The Valerie Plame Affair:* In an earlier version of this paper we stated that if a formula ϕ was stable, agent j knew it, and agent i was connected either directly or indirectly to agent j, then agent i could also come to know ϕ. Here a formula ϕ is said to be stable if for all legal (w, H), $(w, H) \models_{\mathcal{M}} (\phi \rightarrow \Box\phi)$.

However, we were mistaken and an abstract example as well as the Valerie Plame/Judith Miller affair shows why. Suppose that agent i is connected directly to agent j who is connected directly to agents k, m, both of whom are connected to r who knows the value of p. Now m reads p, which is true, from r's website, and j reads p from m's website and thus knows not only that p but also $K_m(p)$. Now the formula $K_m(p)$ is stable, it will never again become false. But i cannot know this although i can know p. For *just by reading j's web page*, i cannot rule out the possibility that j learned about p from k.

The way in which this applies to the Plame-Miller affair is that the fact that Plame was a CIA covert operative was revealed by columnist Robert Novak in July 2003, possibly endangering her life, and this information seems to have come from Miller who is under a federal sentence for refusing to reveal who leaked the name of Valerie Plame to Novak. The point here is that while we know *what* Miller and Novak knew about Plame, we do not know *how* they knew it.

To deal with the notion of legal or justified history we introduce a propositional symbol L which is satisfied only by legal pairs (w, H). (We may also write $L(w, H)$ to indicate that the pair (w, H) is legal.) Since L can only be defined in terms of knowledge, and knowledge in turn requires quantification over legal histories we shall need mutual recursion.
Given a communication graph and the corresponding model $\mathcal{M} = \langle \mathcal{G}, \mathsf{At}, v \rangle$, and pair (w, H), we define the legality of (w, H) and the truth $\models_\mathcal{M}$ of a formula as follows:

- $w, \epsilon \models_\mathcal{M} L$
- $w, H; (i, j, \phi) \models_\mathcal{M} L$ iff $w, H \models_\mathcal{M} L$ and $w, H \models_\mathcal{M} K_j\phi$
- $w, H \models_\mathcal{M} p$ iff $w(p) = 1$, where $p \in \mathsf{At}$
- $w, H \models_\mathcal{M} \neg\phi$ iff $w, H \not\models_\mathcal{M} \phi$
- $w, H \models_\mathcal{M} \phi \wedge \psi$ iff $w, H \models_\mathcal{M} \phi$ and $w, H \models_\mathcal{M} \psi$
- $w, H \models_\mathcal{M} \Diamond\phi$ iff $\exists H',\ H \preceq H',\ L(w, H')$, and $w, H' \models_\mathcal{M} \phi$
- $w, H \models_\mathcal{M} K_i\phi$ iff $\forall (v, H')$ if $(w, H) \sim_i (v, H')$, and $L(v, H')$, then $v, H' \models_\mathcal{M} \phi$

Unless otherwise stated, we will only consider legal pairs (w, H), i.e., pairs (w, H) such that $w, H \models L$. We say ϕ is **valid in** $\mathcal{M}$, $\models_\mathcal{M} \phi$ if for all (w, H), $w, H \models_\mathcal{M} \phi$. ϕ is **valid in the communication graph frame** $\mathcal{F}$ if ϕ is valid in all models based on $\mathcal{F}$.

2.2 Surface Knowledge

Except for each agent's initial information, one may suspect that all information acquired by the agent i is just the sum of the ϕ which i learned from communications (i, j, ϕ). But we saw that this is not true. Given the assumption that both **At** and the structure of the communication graph are common knowledge, agents can come to know facts that are not explicitly contained in the communications.[1] We might still be interested in this 'surface' knowledge which the agents acquire.

Define the sets $X_i(w, H)$ as follows:

1. $X_i(w, \epsilon) = \{v \mid v_{|\mathsf{At}_i} = w_{|\mathsf{At}_i}\}$
2. $X_i(w, H; (i, j, \phi)) = X_i(w, H) \cap \hat{\phi}$
3. $i \neq m$ then $X_i(w, H; (m, j, \phi)) = X_i(w, H)$

[1] Here is an amusing story involving one of us, Parikh. Parikh had published a paper on pumping lemmas and regular sets jointly with A. Ehrenfeucht and G. Rozenberg. At some conference someone asked Parikh, *where* this paper would appear and Parikh did not remember. At this point Rao Kosaraju of Johns Hopkins who was standing by said, it was the *SIAM Journal of Computing.* Parikh then turned to Kosaraju and said, "*you* were the referee!" The point was that Kosaraju's information revealed the existence of an *edge* between him and the editor of *the SIAM journal.*

Intuitively, if $X_i(w,H) \subseteq \widehat{\phi}$, then ϕ is implied (for i) by the sequence of communications. We first show a preliminary lemma which is needed to show that at (w,H), agents know at least the formulas implied by $X_i(w,H)$.

Lemma 1. *If $(w,H) \sim_i (v,H')$, then $X_i(w,H) = X_i(v,H')$.*

Proof. The proof is by induction on $\lambda_i(H) = \lambda_i(H')$. If $\lambda_i(H)$ was empty then H itself might as well be ϵ, and then we use the fact that $X_i(w,\epsilon) = \{u \mid u_{|\mathsf{At}_i} = w_{|\mathsf{At}_i}\}$ is the same as $X_i(v,\epsilon) = \{u \mid u_{|\mathsf{At}_i} = v_{|\mathsf{At}_i}\}$ since $w_{|\mathsf{At}_i} = v_{|\mathsf{At}_i}$. Otherwise we use the fact that since $\lambda_i(H) = \lambda_i(H')$, the initial set $X_i(w,\epsilon) = X_i(v,\epsilon)$ went through exactly the same intersections with various $\hat{\phi}$ when the ground facts ϕ were learned by i. Indeed $X_i(w,H)$ depends *only* on the *set* of ϕ which i learned in H and not on their order. In particular, If (i,j,ϕ) already occurs in H, then $X_i(w,H;(i,j,\phi)) = X_i(w,H)$. □

Lemma 2. *Let $\mathcal{M} = \langle \mathcal{G}, \mathbf{At}, \boldsymbol{v}\rangle$ be any communication graph model and ϕ a ground formula. If $X_i(w,H) \subseteq \widehat{\phi}$, then $(w,H) \models_{\mathcal{M}} K_i(\phi)$.*

Proof. Let $\mathcal{M} = \langle \mathcal{G}, \mathbf{At}, \boldsymbol{v}\rangle$ be a communication graph model. Suppose that ϕ is a ground formula with $X_i(w,H) \subseteq \widehat{\phi}$. Let $(v,H') \sim_i (w,H)$. We must show that $v,H' \models_{\mathcal{M}} \phi$. Since ϕ is a ground formula, this is equivalent to showing that $v(\phi) = 1$. Since $(w,H) \sim_i (v,H')$ by Lemma 1 $X_i(v,H') = X_i(w,H) \subseteq \widehat{\phi}$. Thus we need only the following claim.

Claim: If $X_i(v,H') \subseteq \widehat{\phi}$, then $v(\phi) = 1$.

Proof of claim: The proof is by induction on H'. If $H' = \epsilon$, then since $X_i(v,H') = \{y \mid y_{|\mathsf{At}_i} = v_{|\mathsf{At}_i}\}$ and, of course, $v_{\mathsf{At}_i} = v_{|\mathsf{At}_i}$, we have $v \in X_i(v,H') \subseteq \widehat{\phi}$. Hence $v(\phi) = 1$. Suppose that $m \neq i$ and $H' = H_1;(m,j,\psi)$. Then by construction $X_i(v,H') = X_i(v,H_1)$, and so, since $X_i(v,H_1) = X_i(v,H') \subseteq \widehat{\phi}$, by the induction hypothesis we have $v(\phi) = 1$.

Finally suppose that $H' = H_1(i,j,\psi)$. Then $X_i(v,H') = X_i(v,H_1) \cap \widehat{\psi}$. Since we only consider justified state-history pairs, $X_j(v,H_1) \subseteq \widehat{\psi}$. Hence, by the induction hypothesis $v(\psi) = 1$. Let θ be any formula such that $X_i(v,H_1) = \widehat{\theta}$ (such a formula must exist since At is finite and so every set of states can be defined by a formula). By the induction hypothesis since $X_i(v,H_1) = \widehat{\theta}$, $v(\theta) = 1$. Hence $\widehat{\theta} \cap \widehat{\psi} = X_i(v,H_1;(i,j,\psi)) \subseteq \widehat{\phi}$. Since $v(\theta) = v(\psi) = 1$, $v(\phi) = 1$. This completes the proof of the claim and of the lemma. □

But as we saw, the converse is not true. That is, there are ground formulas that the agents may come to know that are not explicitly contained in their communications. Essentially, these are facts that the agents can derive given their knowledge of the structure of the communication graph and the initial distribution of facts. The sets $X_i(w,H)$ represent the knowledge which agents i would acquire after communication *if* they did not know the structure of the graph.

2.3 Axioms and Decidability

The following axioms and rules are known to be sound and complete with respect to the set of all subset spaces ([13]). Thus they represent the core set of axioms and rules for any topologic.

1. All propositional tautologies
2. $(p \to \Box p) \wedge (\neg p \to \Box\neg p)$, for $p \in \mathsf{At}$.
3. $\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$
4. $\Box\phi \to \phi$
5. $\Box\phi \to \Box\Box\phi$
6. $K_i(\phi \to \psi) \to (K_i\phi \to K_i\psi)$
7. $K_i\phi \to \phi$
8. $K_i\phi \to K_iK_i\phi$
9. $\neg K_i\phi \to K_i\neg K_i\phi$
10. (Cross axiom) $K_i\Box\phi \to \Box K_i\phi$

We include the following rules: modus ponens, K_i and $\Box$ necessitation. We write $\vdash \phi$ if ϕ can be derived from any of the above schemes and rules. The soundness of axioms 1-9 and the rules are easy to verify also for our framework.

We now show that the cross axiom $K_i\Box\phi \to \Box K_i\phi$ is sound. It is easier to consider it in its contrapositive form: $\Diamond L_i\phi \to L_i\Diamond\phi$. This is interpreted as follows: if there is a sequence of updates that lead agent i to consider ϕ possible, then i already thinks it possible that there is a sequence of updates after which ϕ becomes true.

Proposition 1. *$\Diamond L_i\phi \to L_i\Diamond\phi$ is valid in all communication graph models.*

Proof. Let $\mathcal{M} = \langle \mathcal{G}, \mathbf{At}, \boldsymbol{v}\rangle$ be a communication graph model and (w, H) any justified state-history pair. Suppose that $w, H \models \Diamond L_i\phi$. Then there exists H' with $H \preceq H'$ such that $w, H' \models L_i\phi$. Hence there is a pair (v, H'') such that $(v, H') \sim_i (w, H'')$ and $v, H'' \models_{\mathcal{M}} \phi$. Let H''' be any sequence such that $\lambda_i(H) = \lambda_i(H''')$ and $H''' \preceq H''$. Such a history must exist since $H \preceq H'$ and $H' \sim_i H''$. Since $H \preceq H'$, $\lambda_i(H) \preceq \lambda_i(H') = \lambda_i(H'')$. Therefore, we need only let H''' be any initial segment of H'' containing $\lambda_i(H)$. By definition of L, all initial sequences of a legal history are legal. Therefore, since $v, H'' \models_{\mathcal{M}} \phi$, $v, H''' \models \Diamond\phi$; and since $H \sim_i H'''$, $w, H \models_{\mathcal{M}} L_i\Diamond\phi$. □

We leave the problem of finding a complete axiomatization for a future paper, and move to decidability. We show that the satisfiability problem is decidable by showing that a satisfiable formula has a model of bounded size. The main idea is to show that for any history H in which an event of the form (i, j, ϕ) occurs twice is "equivalent" to another history in which that event only occurs once. Here "equivalent" means satisfies the same formulas. We first need a definition. Given any history H, let $c(H)$ be the sequence of events of H generated by the order: e comes before e' iff the first occurrence of e in H occurred before the first occurrence of e' in H. Thus $c(H)$ is the compressed history obtained from H by deleting the second and subsequent occurrences of any event. Thus, for instance, if $H = e_2e_1e_2e_1e_3$ then $c(H) = e_2e_1e_3$.

Definition 4. *Let $w \in W$ be any state and suppose that H and H' are justified histories (for w). We say that H and H' are C-equivalent, written $C(H, H')$, iff $c(H) = c(H')$.*

Intuitively, for two histories H and H', $C(H, H')$ holds if their compressed versions are the same.

Lemma 3. *Fix a state w and suppose that H and H' are justified histories. Then*

1. *If $C(H, H')$ and $L(w, HH_1)$, then $L(w, H'H_1)$ and $C(HH_1, H'H_1)$. In particular, taking H_1 to be empty, $L(w, H)$ iff $L(w, H')$.*
2. *If $C(H, H')$ and $H \sim_i H_1$ for some i, then there is a legal history H_1' such that $C(H_1, H_1')$ and $H' \sim_i H_1'$.*

Proof. Let w be a state and H and H' two justified histories such that $C(H, H')$. To prove part 1, Let H_1 be any history such that HH_1 is legal. Now the legality of an event (i, j, ϕ) in H_1 as part of HH_1 depended on the fact that j knew ϕ. Now every (j, m, ψ) which occurred in H also occurred in H' and if it occurred in H_1 as part of HH_1 it would also occur in H_1 as part of $H'H_1$. Thus the same justifications for H_1 events are available in both cases and $H'H_1$ must also be legal. Clearly, $c(HH_1) = c(H'H_1)$. Therefore $C(HH_1, H'H_1)$.

For part 2, suppose that $H \sim_i H_1$ for some agent i and legal history H_1. Since $H \sim_i H_1$, $\lambda_i(c(H)) = \lambda_i(c(H_1))$. Also, since $c(H) = c(H')$, $\lambda_i(c(H)) = \lambda_i(c(H'))$. Therefore, $\lambda_i(c(H')) = \lambda_i(c(H_1))$.

That is, the sequence of first occurrence of i events in H' is the same as the sequence of first occurrence of i events in H_1. Thus, by adding extra i events to or removing excess i events from H_1, a history H_1' can be constructed such that $H' \sim_i H_1'$. Clearly by construction $c(H_1) = c(H_1')$. □

Corollary 1.
1. *Let the relation D between state history pairs be defined by $D((w, H), (w, H'))$ iff $C(H, H')$. Then $L(w, H)$ iff $L(w, H')$ and D is a bisimulation.*
2. *with the same assumptions, for all formulas ϕ, $w, H \models \phi$ iff $w, H' \models \phi$.*
3. *For all formulas ϕ, $w, H \models \phi$ iff $w, c(H) \models \phi$.*
4. *If H contains (i, j, ψ) and $L(w, H)$ holds, then also $L(H; (i, j, \psi))$, and for all ϕ, $(w, H) \models \phi$ iff $(w, H; (i, j, \psi)) \models \phi$*

Corollary 2. *If a formula ϕ is satisfiable in some graph model $(\mathcal{G}, \textbf{At})$ then it is satisfiable in a history in which no communication (i, j, ϕ) occurs twice.*

This last result immediately gives us a decision procedure as we can limit the length of the history which might satisfy some given formula ϕ. Now there are only a finite number of ground formulas ϕ, thus only a finite number of learnings (i, j, ϕ), and hence only a finite number of histories we need to look at. Alas, this number is quite large and we hope to find a better decision procedure. Note that if we limited the agents to read *only* atomic formulas, a very natural restriction, then the number of possible communications would be smaller and the decision procedure would be faster, and indeed would be in non-deterministic exponential time. The logic *would* change as the formulas $K_i(p \vee q) \rightarrow K_i(p) \vee K_i(q)$ would be valid with such a restriction, but are not valid if non-atomic formulas can be read from another agent's website.

We now define a maximal history (relative to some w) as a history in which all possible (finitely many) communication events have taken place at least once. If H is a maximal history, then we will have, for all H', $C(H, HH')$ and hence for all H', all w, ϕ, $w, H \models \phi$ iff $w, HH' \models \phi$. In other words, a maximal w, H satisfies, for all ϕ, $\phi \leftrightarrow \Box\phi$.

Theorem 1. *The axiom $\Box\Diamond\phi \rightarrow \Diamond\Box\phi$ is valid in Logic of Communication Graphs.*

Proof. Fix w compatible with some history H which satisfies $\Box\Diamond\phi$. Let H' be a maximal history extending H, then w, H' satisfies $\Diamond\phi$ and hence ϕ and hence $\Box\phi$. Since H' extends H, w, H satisfies $\Diamond\Box\phi$. □

We strongly suspect that if H and H' are maximal histories (relative to w), then w, H and w, H' satisfy the same formulas. In this case, $\Diamond\Box\phi \rightarrow \Box\Diamond\phi$ would be valid. This and other issues related to a complete axiomatization will be left for another paper.

3 Connection with Communication Graphs

In this section we will investigate the close connection between formulas valid in a model based on the communication graph and the communication graph. We will prove that the valid formulas characterize the communication graph.

Theorem 2. *Let $\mathcal{G} = (\mathcal{A}, E)$ be a communication graph. Then $(i, j) \in E$ if and only if, for all $l \in \mathcal{A}$ such that $l \neq i$ and $l \neq j$ and all ground formulas ϕ, the scheme*

$$K_j\phi \wedge \neg K_l\phi \rightarrow \Diamond(K_i\phi \wedge \neg K_l\phi)$$

is valid in all communication graph models based on $\mathcal{G}$.

Proof. Suppose that $w, H \models_M K_j\phi \wedge \neg K_l\phi$. Then j knows ϕ and hence i can read ϕ directly from j's website. l is none the wiser as $\lambda_l(H) = \lambda_l(H; (i, j, \phi))$. Therefore, $w, H; (i, j, \phi) \models K_i\phi \wedge \neg K_l\phi$. □

4 Conclusions and Further Work

In this paper we have introduced a logic of knowledge and communication. Communication among agents is restricted by a communication graph, and idealized in the sense that the agents are unaware when their knowledge base is being accessed. We have shown that the communication graph is characterized by the validities of formulas in models based on that communication graph, and that our logic is decidable.

Related Work: This paper fits in with a growing body of work on social software ([14]). One of the main goals of the social software research program is to develop mathematical tools that can be used to study social procedures. Other work that falls into this category is [17] which studies the semantics of messages, [2] which studies voting strategies in the presence of knowledge, and [15] which studies a logic of knowledge with obligation.

Logics of knowledge acquisition through communication have been studied earlier, starting with [18] and more recently in [1, 11, 19, 7]. In chapter 4 of [11], Kooi provides an excellent overview of the current state of affairs of these dynamic epistemic logics. These logics use **PDL** style operators to represent an epistemic update. For example, if $!\phi$ is intended to meain a public announcement of ϕ, then $\langle !\phi \rangle K_i \phi$ is intended to mean that after ϕ is publically announced, agent i knows ϕ. From this point of view, the communication modality $\diamond$ can be understood as existentially quantifying over a sequence of private epistemic updates. However, there are some important differences between the semantics presented in this paper and the semantics found in the dynamic epistemic logic literature. First of all, in our semantics communication is limited by the communication graph. Secondly, we do not consider general epistemic updates as is common in the literature, but rather study a specific type of epistemic update and its connection with a communication graph. Most important is the fact that the history of communications plays a key role in the deninition of knowledge in this paper. The general approach of dynamic epistemic semantics is to define update operations mapping Kripke structures to other Kripke structures intended to represent the effect of an epistemic update on the first Kripke structure. For example, a public announcement of ϕ selects the submodel of a Kripke structure in which ϕ is true at every state. The definition of knowledge after an epistemic update is the usual definition, i.e., ϕ is known by i at state w if ϕ is true in all states that i considers possible from state w in the updated Kripke structure. A closer analysis of the similarities and differences between these two approaches is an interesting topic for further study.

Further Work: We showed that the logic of communication graphs has the finite model property and so is decidable. Other standard questions such as finding an elegant complete axiomatization will also be studied. Another interesting extension would be to allow different types of updates, such as lying, conscious updates (where j is aware that his website is being read), updating to subgroups (creating common knowledge) and so on.

Another natural extension is to consider situations in which agents have a preference over which information they will read from another agent's website. Thus for example, if one hears that an English Ph.D. student and his advisor recently had a meeting, then one is justified in assuming that they probably did not discuss the existence of non-recursive sets, even though the advisor may conceivably know this fact. I.e., the advisor may have the fact, that there exists a non-recursive set, on her website, but there is a very good chance that the Ph.D. student did not ask about this particular fact. Given that this preference over the formulas under discussion among different groups of agents is common knowledge, each agent can regard some (legal) histories as being more or less likely than other (legal) histories. From this ordering over histories, we can define a defeasible knowledge operator for each agent. The operator is defeasible in the sense that agents may be wrong, i.e., it *is* after all possible that the English student and his advisor actually spent the meeting discussing the fact that there must be a non-recursive set.

Finally we remark that our framework and the logic can be seen as a demonstration of the need for cryptographic protocols. Two issues are important here. The first is that an agent may only want part of its knowledge base to be accessible by the public. This may be modeled in our framework by restricting for each agent j the set of formulas that the agent makes available, and so when i is directly connected to j, i can only update by facts in the accessible domain. The second issue is that we may not know the exact structure of the communication graph. For example, if Ann accesses some information from Bob's website, but unknown to Ann, Charles is listening in, then the communication graph has an edge between Charles and Bob, whose presence is not known to Ann or to Bob. Then clearly as a condition for Ann learning some information from Bob, Charles must be able to be informed of that same piece of information. Thus cryptographic protocols essentially intended to ensure that there are no undesired edges between agents in the communication graph. Thus, in that version of our model where the entire graph is not common knowledge, inferring the existence of edges *from* knowledge (as the Kosaraju example showed) is yet another, potentially important extension.

References

1. Baltag, A. and Moss, L., Logics for Epistemic Programs, *Knowledge, Rationality, and Action* section of *Synthese*, 139:**2**, pgs. 165-224, 2004,
2. Samir Chopra, Eric Pacuit and Rohit Parikh, *Knowledge-theoretic Properties of Strategic Voting*, Proceedings of 9th European Conference on Logics in Artifical Intelligence, Jos Jlio Alferes and Joo Leite editors, Lecture Notes in Artificial Intelligence, Springer, pgs. 18-30, 2004.
3. Dabrowski, A, Moss, L, and Parikh, R. Topolgical reasoning and the logic of knowledge. *Annals of Pure and Applied Logic*, **78**, (1996), pp. 73 - 110.
4. Georgatos, K, Modal Logics for Topological Spaces. PhD Dissertation. Graduate School and University Center. City University of New York, 1993.
5. Georgatos, K, Knowledge Theoretic Properties of Topological Spaces. In *Knowledge Representation and Uncertainty.* M. Masuch and L. Polos, Eds. Lecture Notes in Artificial Intelligence, vol. 808, pages 147-159, Springer-Verlag, 1994.
6. Georgatos, K, Knowledge on Treelike Spaces. *Studia Logica*, **59**, (1997), pp. 271 - 231.
7. Gerbrandy, J., *Bisimulations on Planet Kripke*, Ph.D. dissertation, University of Amsterdam, 1999.
8. Heinemann, B., Temporal Aspects of the Modal Logic of Subset Spaces, *Theoretical Computer Science*, **224(1-2)**:135-155, 1999.

9. Heinemann, B., Extending Topological Nexttime Logic. In S. D. Goodwin, A. Trudel, editors, *Temporal Representation and Reasoning*, TIME-00, Cape Breton, Nova Scotia, Canada, pages 87-94, IEEE Computer Society Press, Los Alamitos, CA, 2000.
10. Heinemann, B., A Hybrid Treatment of Evolutionary Sets. In C. A. Coello Coello, A. de Albornoz, L. E. Sucar, O. Cair Battistutti, editors, MICAI'2002: *Advances in Artificial Intelligence*, Mrida, Yucatn, Mexico. Volume 2313 of Lecture Notes in Artificial Intelligence, pages 204-213, Springer, Berlin, 2002.
11. Kooi, B., *Knowledge, Chance, and Change*, Ph.D. dissertation, University of Groningen, 2003.
12. Heinemann, B., A Hybrid Logic of Knowledge Supporting Topological Reasoning. In *Algebraic Methodology and Software Technology*, AMAST 2004, Stirling, United Kingdom. Lecture Notes in Computer Science, Springer, Berlin, 2004. *To appear.*
13. Moss, L. and Parikh, Topological Reasoning and the Logic of Knowledge, *TARK IV*, Ed. Y. Moses, Morgan Kaufmann, 1992.
14. Parikh, R., Social Software, *Synthese*, **132: 3**, Sep 2002, pp. 187-211.
15. Parikh, R., Pacuit, E. and Cogan, E., The logic of knowledge based obligation. Early version presented at DALT '04. Forthcoming in *Knowledge Rationality and Action*: Special Issue on the Knowledge and Games Workshop, 2005.
16. Parikh, R., and R. Ramanujam "Distributed Processing and the Logic of Knowledge", in *Logics of Programs*, Proceedings of a Conference at Brooklyn College, June 1985, Springer Lecture Notes in Computer Science #193., pp. 256-268.
17. Parikh, R. and Ramanujam, R., A knowledge based semantics of messages, in *J. Logic, Language, and Information*, **12**, pp. 453 - 467, 2003.
18. Plaza, J., Logics of public communications, *Proceedings, 4th International Symposium on Methodologies for Intelligent Systems*, 1989.
19. van Ditmarsch, H., *Knowledge Games*, Ph.D. dissertation, University of Groningen, 2000.
20. Vickers, S. *Topology Via Logic*, Cambridge University Press. 1989.
21. Weiss, M. A. and Parikh, R., "Completeness of Certain Bimodal Logics of Subset Spaces", *Studia Logica*, **71:1**, pp. 1 - 30, 2002.