Message Passing in a Dynamic Epistemic Logic Setting

Floor Sietsma
Centrum Wiskunde & Informatica
Science Park 123
1098 XG Amsterdam
f.sietsma@cwi.nl

Jan van Eijck
Centrum Wiskunde & Informatica
Science Park 123
1098 XG Amsterdam
jve@cwi.nl

ABSTRACT

We propose a framework for message passing that combines the best properties of dynamic epistemic semantics and history-based approaches. We assume that all communication is truthful and reliable. Our framework consists of Kripke models with records of sent messages in their valuations. We introduce an update operation for message sending. With this update we can study the exact epistemic consequences of sending a message. We define a class of models that is generated from initial Kripke models by means of message updates, and we axiomatize a logic for this class of models. Next, we add an update modality and sketch a procedure for defining it by means of equivalence axioms.

Categories and Subject Descriptors

F.4.1 [**Mathematical Logic**]: Modal Logic; F.4.1 [**Mathematical Logic**]: Proof Theory; I.2.3 [**Artificial Intelligence**]: Deduction and Theorem Proving

General Terms

Theory

Keywords

Communication, Message Passing, Dynamic Epistemic Logic

1. INTRODUCTION

Message passing is used in a great number of practical situations ranging from daily talk between people to email communication within large companies or the automated communication between computers over long-distance networks.

There have been different proposals on how to model messages and their effects on the knowledge of the agents sending and receiving them. History-based message semantics was defined by Parikh and Ramanujam [10] in terms of states consisting of message sequences (see also [1]). Another approach to message passing is via the update construction in dynamic epistemic logic or DEL [2, 3, 5]. DEL defines the effect of messages (public announcements, group announcements), but without an explicit notion of agency (the message sender is 'outside' the model), and without keeping records of past communications. We propose a framework that combines the best properties of dynamic epistemic semantics and history-based approaches.

We start out by defining a rich set of factual messages, with a very strict identity criterion: messages are the same if and only if they have the same sender, the same recipient set, and the (syntactically) same body. In short: identity of messages means syntactic identity. Thus, we make a strict distinction between message *form* and message *content*. The two messages $p \vee q$ and $q \vee p$ from i to j have different syntactic form, but they have the same content, so they will have the same update effect, except for the trivial difference that syntactically distinct messages are remembered by the agents.

The basic action in our set-up is a *communication step*, an act of revealing part of one's factual knowledge to a particular audience by means of a message. Generally, a communication step is a choice that a particular agent addressing a particular recipient group makes about how much of the factual information she possesses to reveal to that audience. We will show how Kripke models can be constructed from finite sequences of communication steps, and how success of communication sequences can be measured with respect to existing Kripke models.

In this paper we only consider truthful communication, so a message can only be sent if the sender of the message knows that its content is true. Also, all messages are believed, so if an agent receives a message she gains knowledge of the contents. We also assume that the communication channel is reliable, so all messages that are sent are received immediately without any loss of information.

Our semantics allows us to model reasoning about the way the communication took place: agents remember which messages they sent or received, but they may be uncertain about the messages sent or received by other agents. Our construction models these uncertainties in a very precise way.

The system is designed for reasoning about an actual sequence of communication steps, given some initial situation represented by a Kripke model. The semantics allows us to check properties and effects of communication sequences that took place in the past, and allows a limited amount of reasoning about counterfactual situations, like "suppose instead of actual message m, another message, m' had been sent." Also, we can reason about properties and effects of new communication steps.

2. A LOGIC FOR REASONING ABOUT FACTUAL MESSAGES

In this section we will define a rich language that allows to reason about the effects of messages and the knowledge of the agents about these messages. We start out with defining a simpler language L_0 with the purpose that the content of messages is limited to formulas from this simpler language.

Let P be a set of proposition letters. Let N be a finite set of agents.

Definition 1. *Let L_0 be the following language:*

$$\begin{array}{rcl} \psi & ::= & p \mid (i,\psi,G) \mid \neg\psi \mid (\psi \vee \psi) \\ & & \textit{where } p \in P, i \in G \subseteq N. \end{array}$$

This is propositional logic enriched with factual messages: the formula (i,ψ,G) expresses that message (i,ψ,G) with body ψ was sent by agent i to the group G at some moment in the past. Let M be the set of all L_0 messages.

We adopt the convention that a sender always cc's to herself: any message (i,ψ,G) has $i \in G$. We will write $(i,\psi,\{i,j\})$ (a message with a single recipient, plus cc to the sender) as (i,ψ,j).

Note that this language does not contain any epistemic operators, so the messages cannot contain epistemic formulas. However, the messages can contain other messages, which allows us to express complex communication:

Example 1. *Acknowledgement of receipt of a message (i,p,j) can be expressed as $(j,(i,p,j),i)$. Forwarding of (i,p,j) can be expressed as $(j,(i,p,j),k)$.*

Example 2. *A message (i,p,j) with bcc-list $\{k_1,\dots,k_n\}$ can be modeled as a sequence of messages*

$$(i,p,j),(i,(i,p,j),k_1),\dots,(i,(i,p,j),k_n).$$

Each member on the bcc list gets a separate message from the sender of the message to the effect that message (i,p,j) was sent. We will prove in Theorem 4 that the order in which the list $(s_m,(i,p,j),k_1)$, …, $(s_m,(i,p,j),k_n)$ is sent does not matter.

This example already shows that notation can become a bit thick when nesting messages. Therefore we will often shorten notation by naming our messages m, m', m_1, etc. These names should be seen as pure abbreviations. If we abbreviate a message (i,ψ,G) with m then we mean with $s_m = i$ the sender of the message, $b_m = \psi$ the body of the message and $r_m = G$ the group of recipients of the message. We also use these abbreviations in the body of other messages: for example, (j,m,k) is an abbreviation for the message $(j,(i,\psi,G),k)$.

In order to evaluate the truth value of a L_0 formula we need a valuation that gives a truth value not only to the propositions in P, but also to all messages that are mentioned in the formula. Then a positive truth value for a message means that the message was sent, and a negative truth value that it was not sent. In order to know which messages we should consider we first assign a vocabulary to every formula. This is the set of all propositions and messages that are relevant to the truth value of the formula.

Definition 2 (Vocabulary of ψ).

$$\begin{array}{rcl} V_p & := & \{p\} \\ V_{(i,\psi,G)} & := & \{(i,\psi,G)\} \cup V_\psi \\ V_{\neg\psi} & := & V_\psi \\ V_{\psi_1 \vee \psi_2} & := & V_{\psi_1} \cup V_{\psi_2} \end{array}$$

Letting x, y range over vocabulary elements, we set $x \prec y$ if $x \neq y$ and $x \in V_y$. This partially orders a vocabulary by 'depth of embedding'. For example, $\{p,q\} \prec (i,p \vee q,j) \prec (j,(i,p \vee q,j),k)$. A vocabulary is a set of messages and proposition letters that is closed under taking V images. So if a vocabulary contains a message then it will also contain every proposition or message that is used in the body of that message.

Now we can give a truth definition for formulas of L_0 given some valuation of their vocabulary:

Definition 3 (Truth Definition for L_0). *Let Ψ be a set of L_0 formulas ranging over some vocabulary V_Ψ. Let v be a subset of V_Ψ. Then truth at v is defined as follows for all formulas in Ψ:*

$$\begin{array}{rcl} v \models p & \textit{iff} & p \in v \\ v \models (i,\psi,G) & \textit{iff} & (i,\psi,G) \in v \\ v \models \neg\psi & \textit{iff} & v \not\models \psi \\ v \models \psi_1 \vee \psi_2 & \textit{iff} & v \models \psi_1 \textit{ or } v \models \psi_2 \end{array}$$

Call a valuation *truthful* for a vocabulary if $v \models m$ implies $v \models b_m$, for all m in the vocabulary. We will set up our models as sets of truthful valuations.

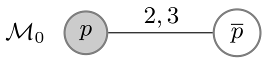
Example 3. *If m is a message, then $(i,\neg m,j)$ is a message quoting m. The* formula *$\neg m$ expresses that m was not sent. The message $(i,\neg m,j)$ expresses that i informs j that m was not sent. The* formula *$\neg(i,\neg m,j)$ expresses that the message $(i,\neg m,j)$ was not sent.*

Such messages can be used to set up a transition system with epistemic models (Kripke models) as states and messages as transition labels. Formal definitions for this will be given in Section 3; the following example illustrates the idea.

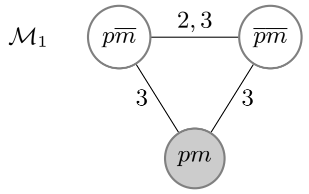
Example 4. *Suppose the initial state $\mathcal{M}_0$ is an epistemic model where agent 1 knows about p, while agents 2 and 3 do not. Let m be the message $(1,p,2)$ of agent 1 informing agent 2 that p is the case. Let m' be the message $(2,p,3)$ of agent 2 informing agent 3 that p is the case. If $\mathcal{M}_0$ is the initial state, the messages can only be sent with m preceding m', for we assume that the sender of a message has to know its contents, and 2 does not know p before the reception of m. This gives:*

$$\begin{tikzcd} \mathcal{M}_0 \arrow[r, "m"] & \mathcal{M}_1 \arrow[r, "m'"] & \mathcal{M}_2 \end{tikzcd}$$

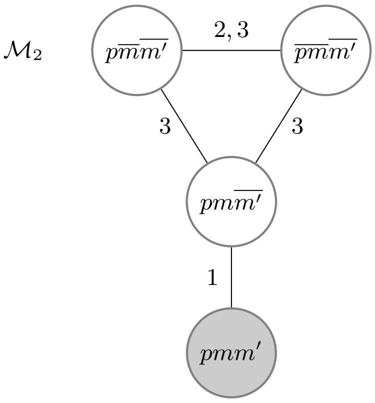
What do the models look like? $\mathcal{M}_0$ is a Kripke model where only 1 knows about p, while p is in fact true. Using grey shading for the actual world, this gives:



Sending message m will inform 2 about p, while agent 3 still considers it possible that nothing has happened:



Finally, after m' is sent, everyone knows about p. Since 1 does not know that m' was sent, 1 cannot distinguish the actual situation with the actual situation before the sending of m':



Note that in the actual situation of $\mathcal{M}_2$ both m and m' were sent. 1 does not know that m' was sent, 3 does not know that m was sent.

We will set up our formal definitions in Section 3 so that message sending actually has the epistemic effects pictured in Example 4. The example illustrates how our approach incorporates elements from message passing histories into epistemic models, and how epistemic models with history information are generated from initial epistemic models without such information.

To describe and reason about the epistemic effects of message exchange we will use epistemic PDL [9]. We start with a version without message modalities. Message modalities will be added in Section 4.

Definition 4 (Epistemic PDL language L). *Let ψ range over L_0 formulas:*

$$\begin{array}{lll} \phi & ::= & \psi \mid \neg\phi \mid \phi \vee \phi \mid \langle\alpha\rangle\phi \\ \alpha & ::= & i \mid ?\phi \mid \alpha;\alpha \mid \alpha \cup \alpha \mid \alpha^* \text{ where } i \in N. \end{array}$$

Call α an epistemic program.

We use the usual abbreviations: $\phi \wedge \psi$ for $\neg(\neg\phi \vee \neg\psi)$ and $[\alpha]\phi$ for $\neg\langle\alpha\rangle\neg\phi$. The formula $\langle i\rangle\phi$ should be read as "agent i considers it possible that ϕ" while $[i]\phi$ stands for "agent i knows that ϕ". To interpret this language we use vocabulary-based Kripke models [6].

Definition 5 (Kripke models with vocabularies). *Let a set of agents N, a set of propositional atoms P and a set of L_0 messages M be given. A vocabulary based Kripke model for N, P, Msg is a tuple $\mathcal{M} = (W, R, Val, Voc, W_0)$ where W is a set of worlds, R is a function that assigns to each $i \in N$ an equivalence relation $\sim_i$ on W, $Voc \subseteq P \cup M$ is a vocabulary, Val is a function that assigns to each world in W a subset of Voc (its valuation), and $W_0 \subseteq W$ is the set of actual worlds.*

Definition 6 (Truth definition for L). *Let Φ be a set of L formulas, and let $\mathcal{M} = (W, R, Val, Voc, W_0)$ be a vocabulary-based Kripke model. Then the truth of an L formula ϕ such that $V_\phi \subseteq Voc$ is given by:*

$$\begin{array}{rll} \mathcal{M} \models_w \psi & \textit{iff} & Val(w) \models \psi \quad \textit{(definition 3)} \\ \mathcal{M} \models_w \neg\phi & \textit{iff} & \mathcal{M} \not\models_w \phi \\ \mathcal{M} \models_w \phi_1 \vee \phi_2 & \textit{iff} & \mathcal{M} \models_w \phi_1 \textit{ or } \mathcal{M} \models_w \phi_2 \\ \mathcal{M} \models_w \langle\alpha\rangle\phi & \textit{iff} & \exists w' : w[\![\alpha]\!]^{\mathcal{M}} w' \\ & & \;\textit{and } \mathcal{M} \models_{w'} \phi \\ w[\![i]\!]^{\mathcal{M}} w' & \textit{iff} & (w, w') \in R_i \\ w[\![?\phi]\!]^{\mathcal{M}} w' & \textit{iff} & w = w' \textit{ and } \mathcal{M} \models_w \phi \\ w[\![\alpha_1;\alpha_2]\!]^{\mathcal{M}} w' & \textit{iff} & \exists w'' \in W : w[\![\alpha_1]\!]^{\mathcal{M}} w'' \\ & & \;\textit{and } w''[\![\alpha_2]\!]^{\mathcal{M}} w' \\ w[\![\alpha_1 \cup \alpha_2]\!]^{\mathcal{M}} w' & \textit{iff} & w[\![\alpha_1]\!]^{\mathcal{M}} w' \textit{ or } w[\![\alpha_2]\!]^{\mathcal{M}} w' \\ w[\![\alpha^*]\!]^{\mathcal{M}} w' & \textit{iff} & \exists w_1, \ldots, w_n \in W : \\ & & w = w_1[\![\alpha]\!]^{\mathcal{M}} w_2 \textit{ and } \ldots \\ & & \;\textit{and } w_{n-1}[\![\alpha]\!]^{\mathcal{M}} w_n = w' \end{array}$$

In each world there is for each agent a relation to the worlds he considers possible. It is not hard to see from this definition that $[i]\phi$ holds in a world w iff for all w' such that $w \sim_i w'$, ϕ holds in w'. So an agent knows something if it holds in all related possible worlds.

Using the axioms and rules for S5 and for PDL we can get a complete axiomatisation for this language:

Theorem 1. *There is a complete logic of L for the class of S5 Kripke models. This logic is decidable.*

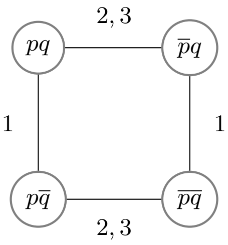
But this is not quite what we want, for the model class is too large. For example, it contains models for non-truthful communication. In section 5 we will define and axiomatize an appropriate class of models.

3. MESSAGE UPDATING

We imagine that the mention of a new message always causes an awareness expansion, by which we shall mean vocabulary expansion, in the sense of [6]. The vocabulary of a Kripke model can be viewed as a global awareness function, indicating the set of propositions and messages that the agents are aware of across the model.

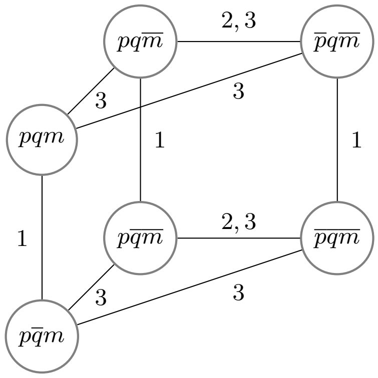
Example 5. *Assume 1 knows (only) about p, 2 and 3 have common knowledge about q.*

Given that the initial facts only mention p and q, we can assume that the initial vocabulary is the set $\{p, q\}$. Our initial Kripke model looks like this:



The vocabulary is indicated in the picture as the list of (proposition) letters that get assigned a truth value. Here p means p is true and $\overline{p}$ means p is false.

Now message $m : (1, p \vee q, 2)$ gets sent. Then the vocabulary of the updated model will be $\{m, p, q\}$. The new vocabulary element m can be either true or false at a world: true means the message was sent, false means that it was not. The non-recipient 3 of the message cannot distinguish between these two cases. m can only be true in worlds where 1 knows $p \vee q$. And if m is true, the recipient 2 knows that $p \vee q$ is true. This gives the following Kripke model:



Note that 'back plane' of the picture is a copy of the old model. Agent 3 cannot distinguish situations where m was sent from situations where m was not sent. But as a result of the message update, in the m-worlds agent 2 now knows everything there is to know about the vocabulary: that p is true, that q is false, and that m was sent.

We now give a formal definition of the update with a new message:

Definition 7 (Updating). *Let a Kripke model $\mathcal{M} = (W, R, Val, Voc, W_0)$ and a message $m : (i, \psi, G)$ be given. Assume that $V_\psi \subseteq Voc$. Then $\mathcal{M} \circ m$ equals*

$$(W', R', Val', Voc', W_0')$$

where

- $W' = \{(w, m) \mid w \in W, \mathcal{M}, w \models [i]\psi\} \cup$ $\{(w, \overline{m}) \mid w \in W\}$,
- $R'(j)$ *for* $j \in G$ *is given by* $\{((w, \overline{m}), (w', \overline{m})) \mid (w, w') \in R(j)\} \cup$ $\{((w, m), (w', m)) \mid (w, m) \in W', (w', m) \in W',$ $(w, w') \in R(j)\}$,
- $R'(j)$ *for* $j \notin G$ *is given by* $\{((w, \overline{m}), (w', \overline{m})) \mid (w, w') \in R(j)\} \cup$ $\{((w, m), (w', \overline{m}) \mid (w, m) \in W', (w, w') \in R(j)\} \cup$ $\{((w, \overline{m}), (w', m) \mid (w', m) \in W', (w, w') \in R(j)\} \cup$ $\{((w, m), (w', m)) \mid (w, m) \in W', (w', m) \in W',$ $(w, w') \in R(j)\}$,
- Val' *is given by*
  $Val'(w, m) = Val(w) \cup \{m\}$,
  $Val'(v, \overline{m}) = Val(w) \backslash \{m\}$,
- $Voc' = Voc \cup \{m\}$,
- $W_0' = \{(w, \overline{m}) \mid w \in W_0\} \cup$ $\{(w, m) \mid (w, m) \in W', w \in W_0\}$.

Intuitively, $\mathcal{M} \circ m$ is the result of extending the vocabulary of $\mathcal{M}$ to include m. We take into account the fact that m can only be sent if its sender knows its contents by checking this for each world w before creating a (w, m) world. The fact that only the recipients of the message get to know that it is sent is reflected by the relations between the $(w, \overline{m})$ and (w, m) worlds.

The following lemma shows that this update does not change any basic facts about the world:

Lemma 2. *For $\psi \in L_0$ such that m does not occur in ψ, $(M, w) \models \psi$ iff $(M \circ m, (w, m)) \models \psi$ iff $(M \circ m, (w, \overline{m}) \models \psi$.*

Proof. A simple induction on ψ. □

Also, the following theorem shows that in the case that m was actually not sent, the knowledge of the agents about basic facts does not change and even if it is sent, the knowledge of the agents who did not receive the message does not change.

Theorem 3. *For $\psi \in L_0$ that does not contain m,*

$$(M, w) \models K_i\psi \text{ iff } (M \circ m, (w, \overline{m})) \models K_i\psi.$$

Furthermore, if $i \notin r_m$ then

$$(M, w) \models K_i\psi \text{ iff } (M \circ m, (w, m)) \models K_i\psi.$$

Proof. Suppose $(\mathcal{M}, w) \models K_i\psi$. Suppose $(w, \overline{m}) \sim_i (w', x)$. Then $w \sim_i w'$ so $(\mathcal{M}, w') \models \psi$ so by Lemma 2 $(\mathcal{M} \circ m, (w', x)) \models \psi$. So $(M \circ m, (w, \overline{m})) \models K_i\psi$.
Suppose $(M \circ m, (w, \overline{m})) \models K_i\psi$. Suppose $w \sim_i w'$. Then $(w, \overline{m}) \sim_i (w', \overline{m})$ so $(\mathcal{M} \circ m, (w', \overline{m})) \models \psi$ and by Lemma 2 $(\mathcal{M}, w') \models \psi$. So $(\mathcal{M}, w) \models K_i\psi$.
Let $i \notin r_m$. Suppose $(\mathcal{M} \circ m, (w, m)) \models K_i\psi$. Let $w \sim_i w'$. Then since $i \notin r_m$, $(w, m) \sim_i (w', \overline{m})$ so $(\mathcal{M} \circ m, (w', \overline{m})) \models \psi$. Then by Lemma 2, $(\mathcal{M}, w') \models \psi$. So $(\mathcal{M}, w) \models K_i\psi$.
Suppose $(\mathcal{M}, w) \models K_i\psi$. Let $(w, m) \sim_i (w', x)$. Then $w \sim_i w'$ so $(\mathcal{M}, w') \models \psi$. Then by Lemma 2 $(\mathcal{M} \circ m, (w', x)) \models \psi$. So $(\mathcal{M} \circ m, (w, m)) \models K_i\psi$. □

We can now show formally that BCCs are unordered.

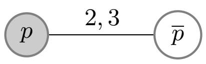
Theorem 4. *Let $\mathcal{M}, w$ be such that $M \models_w m$. Let $m' = (s_m, m, j)$ and $m'' = (s_m, m, k)$. Then*

$$\mathcal{M} \circ m' \circ m'', (w, m', m'') \leftrightarroows \mathcal{M} \circ m'' \circ m', (w, m'', m').$$

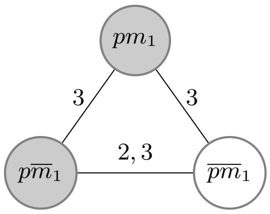
Proof. Check that
$\{((w, x, y), (w, y, x)) \mid$
$w \in W_\mathcal{M}, x \in \{m', \overline{m'}\}, y \in \{m'', \overline{m''}\}\}$
is a bisimulation. □

Theorem 4 and its (easy) proof illustrate that the framework enables us to formalize and prove subtle properties of message exchange. The next example shows that factual messages plus message quoting allows the establishment of 'common knowledge of learning'. 2 learns whether p from 1, and this becomes common knowledge, but outsiders do not learn whether p from the interaction.

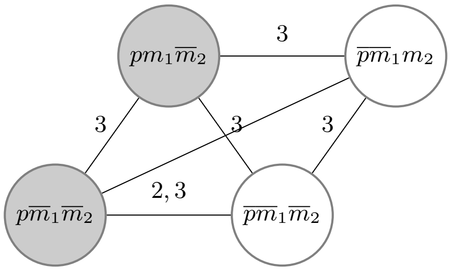
Example 6. *Let $\mathcal{M}$ be a model where 1 knows whether p, while $2, 3$ do not (and this is common knowledge). Gray shading indicates the actual world:*



Let m_1 be the message $(1, p, 2)$ and let m_2 be the message $(1, \neg p, 2)$. The result $\mathcal{M} \circ m_1$ of expanding with m_1:



The result $\mathcal{M} \circ m_1 \circ m_2$ of expanding with m_2:


$pm_1\overline{m}_2$
3
$\overline{pm}_1m_2$
3
3
3
$p\overline{m}_1\overline{m}_2$
$2, 3$
$\overline{pm}_1\overline{m}_2$

Notice that 3 confuses all worlds, since he would not receive either message m_1 or m_2 if it was sent. On the other hand, in the worlds where m_1 or m_2 was sent, 1 and 2 have common knowledge of the truth value of p. Now suppose 1 and 2 want to create common knowledge among the three of them that they know the truth value of p, without revealing that truth value to 3. Then there could be a third message m_3 of the form $(1, m_1 \vee m_2, \{1, 2, 3\})$ that informs the three of them that either m_1 or m_2 was sent without revealing which of the two was actually sent. If we expand our model with this third model, we would see in the resulting model that in those worlds where m_3 was sent, it holds that $[3]([2]p \vee [2]\neg p)$, so 3 knows that 2 knows whether p, but neither $[3]p$ nor $[3]\neg p$ hold, so 3 does not know the value of p himself.

Note that in the extended model, both worlds where the message was sent and worlds where it was not sent are actual. This means that in the extension with a message, we only make the agents aware of the possibility that this message was sent. We do not yet choose between the scenario where it was sent and the scenario where it was not sent. This allows us to go back to the situation where the message was not sent.

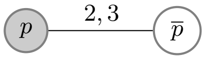
Definition 8. *Let $\mathcal{M} = (W, R, \mathit{Val}, \mathit{Voc}, W_0)$ be a Kripke model and m a message such that $m \in \mathit{Voc}$. Then the model with m removed from it is $\mathcal{M} \circ m^- = (W', R', \mathit{Val}', \mathit{Voc}', W_0)$ with*

$$W' = \{w \in W \mid m \notin \mathit{Val}(w)\}$$

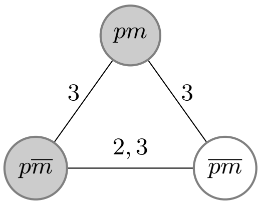
and R', Val', Voc' and W_0' equal to R, Val, Voc and W_0 restricted to W'.

So when we want to "remove" a message from the model, we restrict the model to those worlds where the message was not sent. Notice the difference between $\mathcal{M}$ and $\mathcal{M} \circ m \circ m^-$: m is not in the vocabulary of $\mathcal{M}$ but it is in the vocabulary of $\mathcal{M} \circ m \circ m^-$, even though it is false in all worlds.

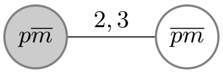
Example 7. *Let us look again at the model from the previous example:*



Updating with m gives the following result:



Now if we do the "back" update m^- we get a model which is very much like the original, but with m in the vocabulary:



Apart from the addition of m, the third model is identical to the first one.

Theorem 5. *For any model $\mathcal{M}$ such that $m \notin \mathit{Voc}_{\mathcal{M}}$, $\mathcal{M} \leftrightarrows_{\backslash\{m\}} \mathcal{M} \circ m \circ m^-$.*

Proof. Let $w \in W_{\mathcal{M}}$. Then $(w, \overline{m}) \in W_{\mathcal{M} \circ m}$ and possibly $(w, m) \in W_{\mathcal{M} \circ m}$. But since (w, m) satisfies m if it exists, $(w, m) \notin W_{\mathcal{M} \circ m \circ m^-}$. We define the relation Z between $W_{\mathcal{M}}$ and $W_{\mathcal{M} \circ m \circ m^-}$ as follows:

$$wZ(w, \overline{m})$$

Clearly, Z is a bisimulation if we do not consider m so $\mathcal{M} \leftrightarrows_{\backslash\{m\}} \mathcal{M} \circ m \circ m^-$. □

As we will show in section 5, we need this definition of "removing" a message to give an axiomatization for the class of models we consider appropriate.

4. A LOGIC FOR MESSAGE PASSING

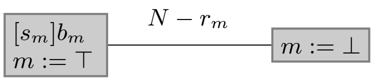
We extend the language with two message modalities $[m]\phi$ and $[\overline{m}]\phi$, with the following truth definitions:

$\mathcal{M} \models_w [\overline{m}]\phi$ iff $\mathcal{M} \circ m \models_{(w,\overline{m})} \phi$ and
$\mathcal{M} \models_w [m]\phi$ iff

$$(w, m) \in W_{\mathcal{M} \circ m} \text{ implies } \mathcal{M} \circ m \models_{(w,m)} \phi.$$

Intuitively, $[\overline{m}]\phi$ expresses that ϕ is true in the model extended with m assuming that m was not sent and $[m]\phi$ that it is true in that model assuming that m was actually sent.

The transitions from $\mathcal{M}$ to $\mathcal{M} \circ m$ can be seen as product updates in the sense of [3]. The appropriate action model is:



In [3] a technique is developed for compiling out action modalities from a language consisting of epistemic PDL plus action model modalities. We will give a brief sketch of how this can be adopted to show that adding the modalities does not increase the expressive power of L. We will give reduction axioms for the $[\overline{m}]$ and $[m]$ modalities. For the Boolean cases, these look as follows:

$$\begin{array}{rcl}
[\overline{m}]p & \leftrightarrow & p \\
[\overline{m}]m' & \leftrightarrow & m' \quad m' \neq m \\
[\overline{m}]m & \leftrightarrow & \bot \\
[\overline{m}]\neg\phi & \leftrightarrow & \neg[\overline{m}]\phi \\
[\overline{m}](\phi_1 \vee \phi_2) & \leftrightarrow & [\overline{m}]\phi_1 \vee [\overline{m}]\phi_2
\end{array}$$

$$\begin{array}{rcl}
[m]p & \leftrightarrow & [s_m]b_m \to p \\
[m]m' & \leftrightarrow & [s_m]b_m \to m' \quad m' \neq m \\
[m]m & \leftrightarrow & \top \\
[m]\neg\phi & \leftrightarrow & \neg[m]\phi \\
[m](\phi_1 \vee \phi_2) & \leftrightarrow & [m]\phi_1 \vee [m]\phi_2
\end{array}$$

The reduction axioms for formulas containing epistemic programs (the PDL modalities α) are more complicated. This is because when we take a step in the original model with an epistemic program, we can only take the same step in the updated model if this step is also possible in the action model. So we have to check whether this relation is present in the action model by doing the same step in both the original model and the action model "concurrently".

For all epistemic programs, we can compute an epistemic program that is the equivalent of the original program together with a concurrent step in the action model. This new epistemic program can be executed in the updated model only if the original epistemic program could be executed concurrently in the original model and in the action model. Call the $m := \top$ state of the action model e_0 and the $m := \bot$ state e_1. With $T^m_{ij}(\pi)$ we mean the program that is the equivalent of doing π in the original model and concurrently moving from state e_i to state e_j in the action model. We define it inductively as follows:

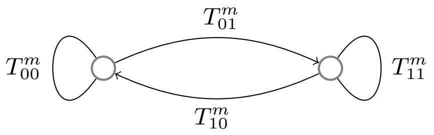
$$\begin{array}{rcl}
T^m_{00}(i) & := & ?[s_m]b_m; i; ?[s_m]b_m \\
T^m_{01}(i) & := & \left\{ \begin{array}{cl} ?\bot & \text{if } i \in G \\ ?[s_m]b_m; i & \text{otherwise} \end{array} \right. \\
T^m_{10}(i) & := & \left\{ \begin{array}{cl} ?\bot & \text{if } i \in G \\ i; ?[s_m]b_m & \text{otherwise} \end{array} \right. \\
T^m_{11}(i) & := & i
\end{array}$$

$$\begin{array}{rcl}
T^m_{00}(?\psi) & := & ?([s_m]b_m \wedge \psi) \\
T^m_{01}(?\psi) & := & ?\bot \\
T^m_{10}(?\psi) & := & ?\bot \\
T^m_{11}(?\psi) & := & ?\psi
\end{array}$$

$$\begin{array}{l}
T^m_{00}(\pi_1;\pi_2) := (T^m_{00}(\pi_1); T^m_{00}(\pi_2)) \cup (T^m_{01}(\pi_1); T^m_{10}(\pi_2)) \\
T^m_{01}(\pi_1;\pi_2) := (T^m_{00}(\pi_1); T^m_{01}(\pi_2)) \cup (T^m_{01}(\pi_1); T^m_{11}(\pi_2)) \\
T^m_{11}(\pi_1;\pi_2) := (T^m_{11}(\pi_1); T^m_{11}(\pi_2)) \cup (T^m_{10}(\pi_1); T^m_{01}(\pi_2)) \\
T^m_{10}(\pi_1;\pi_2) := (T^m_{11}(\pi_1); T^m_{10}(\pi_2)) \cup (T^m_{10}(\pi_1); T^m_{00}(\pi_2))
\end{array}$$

$$\begin{array}{rcl}
T^m_{00}(\pi_1 \cup \pi_2) & := & T^m_{00}(\pi_1) \cup T^m_{00}(\pi_2) \\
T^m_{01}(\pi_1 \cup \pi_2) & := & T^m_{01}(\pi_1) \cup T^m_{01}(\pi_2) \\
T^m_{10}(\pi_1 \cup \pi_2) & := & T^m_{10}(\pi_1) \cup T^m_{10}(\pi_2) \\
T^m_{11}(\pi_1 \cup \pi_2) & := & T^m_{11}(\pi_1) \cup T^m_{11}(\pi_2)
\end{array}$$

The final case is the reduction for π^*. Note that we can see the action model as the following automaton:



Then the epistemic program giving all finite paths through the action model starting in e_0 and ending in e_0 is:

$${T^m_{00}}^*(T^m_{01}{T^m_{11}}^*T^m_{10}{T^m_{00}}^*)^*.$$

Similarly, if we take e_0 as start state and e_1 as final state, we get:

$${T^m_{00}}^*T^m_{01}{T^m_{11}}^*(T^m_{10}{T^m_{00}}^*T^m_{01}{T^m_{11}}^*)^*.$$

For e_1 as start and as stop state:

$${T^m_{11}}^*(T^m_{10}{T^m_{00}}^*T^m_{01}{T^m_{11}}^*)^*.$$

And finally, if we take e_1 as start state and e_0 as stop state:

$${T^m_{11}}^*T^m_{10}{T^m_{00}}^*(T^m_{01}{T^m_{11}}^*T^m_{10}{T^m_{00}}^*)^*.$$

All in all we get the following recipe for transforming an epistemic expression of the form π^*:

$$T^m_{00}(\pi^*) :=$$

$$(T^m_{00}(\pi))^*; (T^m_{01}(\pi); (T^m_{11}(\pi))^*; T^m_{10}(\pi); (T^m_{00}(\pi))^*)^*.$$

$$T^m_{01}(\pi^*) := (T^m_{00}(\pi))^*; T^m_{01}(\pi); (T^m_{11}(\pi))^*;$$

$$(T^m_{10}(\pi); (T^m_{00}(\pi))^*; T^m_{01}(\pi); (T^m_{11}(\pi))^*)^*$$

$$T^m_{11}(\pi^*) :=$$

$$(T^m_{11}(\pi))^*; (T^m_{10}(\pi); (T^m_{00}(\pi))^*; T^m_{01}(\pi); (T^m_{11}(\pi))^*)^*$$

$$T^m_{10}(\pi^*) := (T^m_{11}(\pi))^*; T^m_{10}(\pi); (T^m_{00}(\pi))^*;$$

$$(T^m_{01}(\pi); (T^m_{11}(\pi))^*; T^m_{10}(\pi); (T^m_{00}(\pi))^*)^*$$

Now we can give the reduction axioms for the case of epistemic programs:

$$\begin{array}{rcl}
[m][\pi]\phi & \leftrightarrow & [T^m_{00}(\pi)][m]\phi \wedge [T^m_{01}(\pi)][\overline{m}]\phi \\
[\overline{m}][\pi]\phi & \leftrightarrow & [T^m_{11}(\pi)][\overline{m}]\phi \wedge [T^m_{10}(\pi)][m]\phi
\end{array}$$

Now all that remains to do is to give reduction axioms for the update of removing a message. Since we simply only look at the worlds where m is false, the reduction axiom is simple:

$$[m^-]\phi \leftrightarrow (\neg m \to \phi)$$

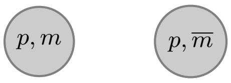
This gives:

THEOREM 6. *The language L and the language L with message modalities added have the same expressive power.*

5. PROPERLY GENERATED MODELS

As we mentioned before, we do not consider all models with messages appropriate. We assume that all communication is truthful and reliable, therefore any world where a message has been sent should also make its contents true, and the fact that it is sent should be known among its recipients. However, there are more subtle requirements for reasonable models, as the following example shows.

EXAMPLE 8. *Consider the following model with three agents and a message $m = (2, p, 3)$:*



There are two possible situations, one where m was sent and one where it was not sent, and none of the agents confuse the two situations. All communication in this model is truthful and reliable but still there is something strange about the model: agent 1 knows the message from 2 to 3 was sent, even though he should not have received it.

This example shows that simple axioms expressing properties like "if a message is true then its content should also be true" are not enough to rule out unrealistic models. The class of models we would like to consider is the class of *properly generated* models:

DEFINITION 9 (PROPERLY GENERATED). *A model $\mathcal{M}$ is properly generated iff there is some model $\mathcal{M}_0$ and a list of messages $m_1, ..., m_n$ such that there are no messages in the vocabulary of $\mathcal{M}_0$ and*

$$\mathcal{M} \underline{\leftrightarrow} \mathcal{M}_0 \circ m_1 \circ ... \circ m_n$$

So a model is properly generated if it can be built from a model containing no messages (we call such a model an *initial* model) using our message update. We can check whether a model is properly generated by using our "removing" update.

LEMMA 7. *For any sequence of messages $m_1, ..., m_n$ such that $m_1 \not\prec ... \not\prec m_n$ and for any two models $\mathcal{M}, \mathcal{N}$ such that $\mathcal{M} \underline{\leftrightarrow}_{\backslash\{m_1,...,m_n\}} \mathcal{N}$,*

$$\mathcal{M} \circ m_1 \circ ... \circ m_n \underline{\leftrightarrow} \mathcal{N} \circ m_1 \circ ... \circ m_n$$

PROOF. Let Z be a bisimulation between $\mathcal{M}$ and $\mathcal{N}$. We define a relation X between $\mathcal{M} \circ m_1 \circ ... \circ m_n$ and $\mathcal{N} \circ m_1 \circ ... \circ m_n$ as follows. For any two worlds $w \in W_{\mathcal{M}}$ and $v \in W_{\mathcal{N}}$ and any sequence $x = x_1, ..., x_n$ where $x_i = m_i$ or $x_i = \overline{m_i}$,

$$(w, x)X(v, x) \Leftrightarrow wZv$$

Note that the question of whether (w, x) exists depends on whether $(\mathcal{M}, w) \models K_{s_{m_1}} b_{m_1}$ if $x_1 = m_1$, and whether $(\mathcal{M}, w, x_1) \models K_{s_{m_2}} b_{m_2}$ if $x_2 = m_2$, etcetera. Similarly for (v, x) and $\mathcal{N}$. But because $m_1 \not\prec ... \not\prec m_n$, these things only depend on the propositions that are true in w and in v (which are the same because wZv) and the earlier messages. So (w, x) exists iff (v, x) exists. So X is total. It is clear from the definition of message update that X is a bisimulation. □

THEOREM 8. *A model $\mathcal{M}$ is properly generated iff there is an order $m_1, ..., m_n$ listing all messages in the vocabulary of $\mathcal{M}$ such that $m_1 \not\prec ... \not\prec m_n$ and*

$$\mathcal{M} \underline{\leftrightarrow} \mathcal{M} \circ m_n^- \circ ... \circ m_1^- \circ m_1 \circ ... \circ m_n$$

PROOF. ⇒: Suppose $\mathcal{M}$ is properly generated. Then there is some initial model $\mathcal{M}_0$ and a list of messages m_1, ..., m_n such that $\mathcal{M} \underline{\leftrightarrow} \mathcal{M}_0 \circ m_1 \circ ... \circ m_n$. By repeated use of Theorem 5, $\mathcal{M}_0 \underline{\leftrightarrow}_{\backslash\{m_n,...,m_1\}} \mathcal{M} \circ m_1^- \circ ... \circ m_n^-$. Then by Lemma 7, $\mathcal{M} \underline{\leftrightarrow} \mathcal{M} \circ m_n^- \circ ... \circ m_1^- \circ m_1 \circ ... \circ m_n$.

⇐: Suppose there is such an order $m_1, ..., m_n$. Let $\mathcal{N}$ be the model like $\mathcal{M} \circ m_n^- \circ ... \circ m_1^-$ but with m not in the vocabulary. Clearly, $\mathcal{M} \circ m_n^- \circ ... \circ m_1^- \underline{\leftrightarrow}_{\backslash\{m_1,...,m_n\}} \mathcal{N}$ so by Lemma 7, $\mathcal{M} \circ m_n^- \circ ... \circ m_1^- \circ m_1 \circ ... \circ m_n \underline{\leftrightarrow} \mathcal{N} \circ m_1 \circ ... \circ m_n$. But because $m_1, ..., m_n$ are all the messages in the vocabulary of $\mathcal{M}$, $\mathcal{N}$ is an initial model so $\mathcal{M} \circ m_n^- \circ ... \circ m_1^- \circ m_1 \circ ... \circ m_n$ is properly generated, and then so is $\mathcal{M}$. □

Given a set of messages M, let $\overrightarrow{M}$ be the set of all sequences listing all these messages that comply with $\prec$. Then the following theorem, which follows from Theorem 8 gives an axiom scheme that characterizes the class of properly generated models:

THEOREM 9. *Let $\mathcal{M}$ be a model and let M be the set of messages in the vocabulary of $\mathcal{M}$. Then $\mathcal{M}$ is properly generated if and only if it satisfies for any $\varphi \in L$:*

$$\bigvee_{(m_1 ... m_n) \in \overrightarrow{M}} (\varphi \leftrightarrow [m_n^-]...[m_1^-][m_1]...[m_n]\varphi)$$

6. RELATED APPROACHES

This work was inspired by the wish to incorporate explicit messages in Dynamic Epistemic Logic. We will clarify what the added value of our approach is compared to the usual DEL as in [2, 3, 5]. In the usual DEL, there is no mention of any messages and the only atoms in the models are propositions. The models can be updated with so-called action models, of which our message update is a special case. In our approach we have tailored an action model for a specific kind of group messages with a sender and a set of recipients. This is very useful in modeling since it is no longer up to the user of the framework to come up with the right action model: this is automatically "generated" when defining the message. This way, we make a step towards formalizing the modeling procedure which makes it easier and less error-prone.

We have combined DEL with the vocabulary expansion proposed in [6] and use this to introduce messages explicitly in the models. This has the great advantage that we can model agents who reason about messages that have been sent and even messages about other messages. This allows for constructions like forward, acknowledgement, BCC recipients etcetera.

A different approach to modelling email communication is presented in [?]. This work specifically investigates how common knowledge emerges in settings where emails are exchanged. In [?] the messages can only contain atomic *notes* which may be compared to propositions. Clearly this is less expressive than the messages containing any non-epistemic formula that we propose. This is partially made up for because in [?] there are explicit forward messages that include a previous message, and also each message has an explicit set of BCC recipients. In our framework we can model forward by sending a message containing another message and in fact it is shown in [?] that BCC can be simulated by a number of forwards.

Another difference is that in [**?**] the initial situation is quite limited. Each agent has a set of notes he knows initially, and the agents do not know anything about each other's notes. Of course in our framework we have a lot more possibilities for the intial situation: we allow any Kripke model without messages.

While the structure of messages and the initial situation in [**?**] is simpler than that in our approach, on another level the approach presented there is a lot more complex because there is no finite set of possible messages. Instead, every message is considered possible by the agents. The number of possible basic facts the messages are about is still finite, but the number and depth of the forwards is unlimited. This may be more realistic in some situations, for example email communication where any email may be sent by the agents. However, it introduces the great complication that if the number of possible messages is infinite, then the number of possible states is infinite too. This makes model checking very complicated: in [**?**] a decidability result for only positive formulas is mentioned.

A key concept in [**?**] is that of a *legal state*. This is comparable to our notion of properly generated models, but since there the relations are generated only from the messages that have been sent it is a condition on single states rather than on the whole (infinitely large!) model. There, a state is legal if the content of the messages that were sent were in fact true and known to the sender. In our approach, these conditions are checked automatically when doing a message update, which ensures that all models that are generated using this update (i.e. properly generated models) are legal in their sense.

There is also a difference in perspective on the states: while we consider our states as basic entities with the messages as their attributes, in [**?**] the states are defined by their messages and all relations between the states are generated by them. In practice, this does not change much because in our framework the relations are (after deciding on the initial situation) generated by the message updates. However, the point of view is clearly different and it shows that while our work is based in DEL, [**?**] leans more towards interpreted systems.

This brings us to the comparison of our work to interpreted systems as presented in e.g. [8]. There, the focus is on a global state that is constructed by combining local states of the agents. Then these states are related for an agent if that agent's local state matches. In our approach, there is no clear distinction between one agent's and another agent's information. One possible such distinction would be to say that an agen't local state is his "inbox" of messages he sent or received up to that moment. Then one would somehow also have to incorporate the messages forwarded to the agent. In our future research we want to investigate this possibility, which might even give us another axiomatization of properly generated models by checking whether an agent's relations are only between worlds that match on the agent's inbox.

The idea of time is clearly incorporated in interpreted systems. In our framework this is less explicit: we can show how the model evolves over time by doing a sequence of message updates, but once these updates have been done the only information that is preserved in the model is whether the message has been sent at some point in time, not when it was sent exactly or an ordering between them. Of course there is the vocabulary embedding relation $\prec$, but this only partially orders the messages. This has the advantage of keeping the messages simple, and in a lot of applications the exact ordering between messages is not so relevant. However, it might be interesting to incorporate timing in our framework and this is one of the points on our research agenda.

7. FURTHER WORK

It would be nice to have another axiomatization of the class of properly generated models without explicit use of the removing update m^-. For example, two sound axioms are

$$\begin{array}{lll} m \rightarrow & b_m & \text{(truthful communication)} \\ m \rightarrow & [r_m^*]m & \text{(successful communication)} \end{array}$$

but these axioms are not complete.

The definition of messages in this paper is purely syntactic. The messages $(1, p \vee q, 2)$ and $(1, q \vee p, 2)$ are different: their bodies use different syntax to express the same propositional content. We intend to explore definitions with a more semantical flavour that identify messages with the same propositional content.

The treatment of awareness in the paper may be too blunt. The global awareness notion of the paper could be replaced with a more subtle one, where different agents may be aware of different vocabularies in different worlds [7, 4].

Our system has the curious property that agents are affected by an expansion with messages that are not addressed to them: they consider the fact that such a message was sent possible. The history-based system of Parikh and Ramanujam has the same property, as does the process of updating with S5 action models for group announcements (see, e.g., [2]).

In some situations this property is perfectly realistic, for example in a game where in every new round the agents know which messages may be sent. However, if we are modeling everyday communication it is less realistic: when two people are communication and a third person does not know what they are communicating about the third person usually thinks any message is possible, and does not have a specific possible message in mind.

Using awareness expansion does not change this property. For consider an agent j that is aware of p but does not know whether p is the case, and a message m not addressed to j that informs its audience that p is the case. Then, even if j is not aware of m, something changes in the model that j can notice: after m was sent j must hold it for possible that agents have learnt something about p.

Look at this informally. How can an agent i ever know for sure that another agent j does *not* know whether p? Suppose initially $[i](\neg[j]p \wedge \neg[j]\neg p)$. Suppose i holds it for possible that some other agent k knows whether p. In other words, $\langle i \rangle([k]p \vee [k]\neg p)$ holds. How can this situation persist? How can i be sure that k does not send a secret message (k, p, j) or $(k, \neg p, j)$? One possible solution would be to always start from initial models where $[i](\neg[j]p \wedge \neg[j]\neg p)$ does not hold, for any i, j, p. This has the disadvantage of blowing up the size of the initial models. Another possibility would be to add temporal operators. Then $[i](\neg[j]p \wedge \neg[j]\neg p)$ might hold at some point in time, but such 'knowledge of ignorance' formulas would not be persistent over time, for in the next instant j might have learnt the value of p through some message.

Finally, as mentioned in the previous section we would like to investigate what happens when we add timing and timestamps to our messages, or at least explicitly define an ordering between them.

Acknowledgements

This paper has benefited from conversations and communications with Krzysztof Apt, Johan van Benthem, Giovanna D'Agostino, Hans van Ditmarsch, Tim French, Rohit Parikh, Sunil Simon, Yanjing Wang, and the participants of the Daghstuhl Seminar on Reasoning and Interaction, March 7 – 11, 2011. We also wish to thank three anonymous TARK reviewers for helpful comments.

8. REFERENCES

[1] K.R. Apt, A. Witzel, and J.A. Zvesper. Common knowledge in interaction structures. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge (TARK XII)*, pages 4–13, 2009.

[2] A. Baltag and L.S. Moss. Logics for epistemic programs. *Synthese*, 139(2):165–224, 2004.

[3] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.

[4] Hans van Ditmarsch and Tim French. Becoming aware of propositional variables. In Mohua Banerjee and Anil Seth, editors, *Logic and Its Applications*, volume 6521 of *Lecture Notes in Computer Science*, pages 204–218, Berlin/Heidelberg, 2011. Springer.

[5] H.P. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library.* Springer, 2006.

[6] Jan van Eijck, Yanjing Wang, and Floor Sietsma. Composing models. In Wiebe van der Hoek, editor, *Online Proceedings of LOFT 2010*, `http://loft2010.csc.liv.ac.uk/`, 2010.

[7] R. Fagin and J.Y. Halpern. Belief, awareness and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

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

[9] D. Kozen and R. Parikh. An elementary proof of the completeness of PDL. *Theoretical Computer Science*, 14:113–118, 1981.

[10] R. Parikh and R. Ramanujam. A knowledge based semantics of messages. *Journal of Logic Language and Information*, 12, 2003.