Common Knowledge in Email Exchanges

Floor Sietsma[1] and Krzysztof R. Apt[1 2]

[1] Centre Mathematics and Computer Science (CWI), Amsterdam
[2] University of Amsterdam

Abstract. We consider a framework in which a group of agents communicates by means of emails, with the possibility of replies, forwards and blind carbon copies (BCC). We study the epistemic consequences of such email exchanges by introducing an appropriate epistemic language and semantics. Then we clarify when a group of agents acquires common knowledge of the formula expressing that an email was sent.

1 Introduction

1.1 Motivation

Email is by now a prevalent form of communication. Its advantages speak for themselves. However, we rarely pause to reflect on its undesired consequences. Just to mention a few.

One occasionally reads about scandals caused by email leaks, see, e.g., [2]. On a smaller scale, users of the *blind carbon copy* feature (BCC) are sometimes confronted with an undesired situation in which a BCC recipient of an email reveals his status to others by using the *reply-all* feature.

Recently, a main Dutch daily, NRC Handelsblad, reported, see [7], that Wouter Bos, the Deputy Prime minister in the previous Dutch government, revealed the extensive network of his contacts by sending out his new email address to hundreds of influential recipients whose email addresses were erronously put in the CC list instead of the BCC list. The list was leaked to the newspaper.

So when studying email exchanges a natural question arises: what are their knowledge-theoretic consequences? To put it more informally: after an email exchange took place, who knows what? To answer this question we study email exchanges by focusing on relevant features that we encounter in most email systems.

More specifically, we study the following form of email communication:

- each email has a sender, a non-empty set of regular recipients and a (possibly empty) set of blind carbon copy (BCC) recipients. Each of the recipients receives a copy of the message and is only aware of the regular recipients and not of the BCC recipients,
- in the case of a reply to or a forward of a message, the *unaltered* original message is included,
- in a reply or a forward, one can append new information to the original message one replies to or forwards.

As a result, the email exchanges, as studied here, are essentially different from other forms of communication, in particular from multicasting, i.e., sending a message to a group of recipients. Also, the resulting model of email communication differs from the ones that were studied in other papers in which only limited aspects of emails have been considered. These papers are discussed below.

1.2 Related work

The study of the epistemic effects of communication in distributed systems originated in the eighties and led to the seminal book [5]. The relevant literature, including [4], deals only with the customary form of communication, notably asynchronous send.

The epistemic effects of other forms of communication were studied in numerous papers. In particular, in [9] the communicative acts are assumed to consist of an agent j 'reading' an arbitrary propositional formula from another agent i. The idea of an epistemic contents of an email is implicitly present in [10], where a formal model is proposed that formalizes how communication changes the knowledge of a recipient of the message. In [3] a dynamic epistemic logic modelling effects of communication and change is introduced and extensively studied. [8] surveys these and related approaches and discusses the used epistemic, dynamic epistemic and doxastic logics. Further, in [12] an epistemic logic was proposed to reason about information flow w.r.t. underlying communication channels.

Most related to the work here reported are the following two references. [1] studied knowledge and common knowledge in a set up in which the agents send and forward propositional formulas in a social network. However, the forward did not include the original message which limited the scope of the resulting analysis. More recently, in [11] explicit messages are introduced in a dynamic epistemic logic to analyze a very similar setting, though it is assumed that the number of messages is finite and BCC is simulated as discussed in Section 6.

Finally, the concept of a causal relation between messages in distributed systems is due to [6].

1.3 Contributions

To study the relevant features of email communication we introduce in the next section a carefully chosen set of emails. We make a distinction between a *message*, which is sent to a public recipient list, and an *email*, which consists of a message and a set of BCC recipients. This distinction is relevant because a *forward* email contains only a message, without the list of BCC recipients. We also introduce the notion of a legal state that captures the fact that there is a *causal ordering* on the emails. For example, an email needs to precede any forward of it.

To reason about the knowledge of the agents after an email exchange has taken place we introduce in Section 3 an appropriate epistemic language. Its semantics takes into account the uncertainty of the recipients of an email about its set of BCC recipients and the ignorance about the existence of emails that one neither sent nor received. This semantics allows us to evaluate epistemic

formulas in legal states, in particular the formulas that characterize the full knowledge-theoretic effect of an email.

In Section 4 we present the main result of the paper, that clarifies when a group of agents can acquire common knowledge of the formula expressing the fact that an email has been sent. This characterization in particular sheds light on the epistemic consequences of BCC. The proof this result is given in Section 5. Then in Section 6 we show how BCC can be simulated using only messages without BCC recipients.

2 Preliminaries

2.1 Messages

In this section we define the notion of a message. We assume non-empty and finite sets of ***agents*** $Ag = \{1, \ldots, n\}$ and of ***notes*** P. Each agent has a set of notes he knows initially.

We make a number of assumptions. Firstly, we assume that the agents do not know which notes belong to the other agents. Furthermore, we assume that the agents only exchange emails about the notes. In particular, they cannot communicate epistemic formulas. We also assume that an agent can send a message to other agents containing a note only if he knows it initially or has learned it through an email he received earlier.

We inductively define ***messages*** as follows, where in each case we assume that $G \neq \emptyset$:

- $m := s(i, l, G)$; the message containing note l, sent by agent i to the group G,
- $m := f(i, l.m', G)$; the forwarding by agent i of the message m' with added note l, sent to the group G.

So the agents can send a message with a note or forward a message with a note appended. In the examples we assume that there exists a note **true** that is known by all agents and we identify **true**.m with m.

If m is a message, then we denote by $S(m)$ and $R(m)$, respectively, the singleton set consisting of the agent sending m and the group of agents receiving m. So for the above messages m we have $S(m) = \{i\}$ and $R(m) = G$. We do allow that $S(m) \subseteq R(m)$, i.e., that one sends a message to oneself.

Special forms of the forward messages can be used to model reply messages. Given $f(i, l.m, G)$, using $G = S(m)$ we obtain the customary *reply* message and using $G = S(m) \cup R(m)$ we obtain the customary *reply-all* message. (In the customary email systems there is syntactic difference between a forward and a reply to these two groups of agents, but the effect of both messages is exactly the same, so we ignore this difference.) In the examples we write $s(i, l, j)$ instead of $s(i, l, \{j\})$, etc.

2.2 Emails

An interesting feature of most email systems is that of the blind carbon copy (BCC). We would like to study the epistemic effects of sending an email with BCC recipients and will now include this feature in our presentation.

In the previous subsection we defined messages that have a sender and a group of recipients. Now we define the notion of an email which allows the additional possibility of sending a BCC of a message. The BCC recipients are not listed in the list of recipients, therefore we have not included them in the definition of a message. Formally, by an ***email*** we mean a construct of the form m_B, where m is a message and $B \subseteq Ag$ is a possibly empty set of BCC recipients. Given a message m we call each email m_B a ***full version*** of m.

Since the set of BCC recipients is 'secret', it does not appear in a forward. That is, the forward of an email m_B with added note l is the message $f(i, l.m, G)$ or an email $f(i, l.m, G)_C$, in which B is not mentioned. However, this forward may be sent not only by a sender or a regular recipient of m_B, but also by a BCC recipient. Clearly, the fact that an agent was a BCC recipient of an email is revealed at the moment he forwards its message.

A natural question arises: what if someone is both a regular recipient and a BCC recipient of an email? In this case, no one (not even this BCC recipient himself) would ever notice that this recipient was also a BCC recipient since everyone can explain his knowledge of the message by the fact that he was a regular recipient. Only the sender of the message would know that this agent was also a BCC recipient. This fact does not change anything and hence we assume that for any email m_B we have $(S(m) \cup R(m)) \cap B = \emptyset$.

2.3 Legal States

Our goal is to analyze knowledge of agents after some email exchange took place. To this end we need to define a possible collection of sent emails.

First of all, we shall assume that every message is used only once. In other words, for each message m there is at most one full version of m, i.e., an email of the form m_B. The rationale behind this decision is that a sender of m_B and $m_{B'}$ might equally well send a single email $m_{B \cup B'}$. This assumption can be summarized as a statement that the agents do not have 'second thoughts' about the recipients of their emails. It also simplifies subsequent considerations.

One could argue that there is a total ordering on the emails entailed by the time at which they were sent. However, the fact that an email was sent at a certain time does not imply that it was also read at that time. All what we can assert it that the email was read after it was sent. Further, the order in which an agent reads the emails he received is undetermined. This explains why we do not impose a linear ordering on the emails and we do not give the messages time stamps.

However, we have to impose some ordering on the sets of emails. For example, we need to make sure that the agents only send information they actually know.

Moreover, a forward can only be sent after the original email was sent. We will introduce the minimal partial ordering that takes care of such issues.

First, we define by structural induction the ***factual information*** $FI(m)$ contained in a message m as follows:

$$FI(s(i,l,G)) := \{l\},$$
$$FI(f(i,l.m,G)) := FI(m) \cup \{l\}.$$

We will use the concept of a ***state*** to model the effect of an email exchange. A state $s = (E, L)$ is a tuple consisting of a finite set E of emails that took place and a sequence $L = (L_1, \ldots, L_n)$ of sets of notes for all agents. The idea of these sets is that each agent i initially knows the notes in L_i. We use E_s and L_s to denote the corresponding elements of a state s, and $L_1, \ldots, L_n$ to denote the elements of L.

We say that a state $s = (E, L)$ is ***legal*** w.r.t. a strict partial ordering (in short, an spo) $\prec$ on E if it satisfies the following conditions:

L.1: for each email $f(i,l.m,G)_B \in E$ an email $m_C \in E$ exists such that $m_C \prec f(i,l.m,G)_B$ and $i \in S(m) \cup R(m) \cup C$,
L.2: for each email $s(i,l,G)_B \in E$, where $l \notin L_i$, an email $m_C \in E$ exists such that $m_C \prec s(i,l,G)_B$, $i \in R(m) \cup C$ and $l \in FI(m)$,
L.3: for each email $f(i,l.m',G)_B \in E$, where $l \notin L_i$, an email $m_C \in E$ exists such that $m_C \prec f(i,l.m',G)_B$, $i \in R(m') \cup C$ and $l \in FI(m')$.

Condition L.1 states that the agents can only forward messages they previously received. Conditions L.2 and L.3 state that if an agent sends, a note that he did not initially know, then he must have learned it by means of an earlier email.

We say that a state s is legal iff it is legal w.r.t. some spo. Given a legal state s, by its ***causality ordering*** we mean the smallest (so the least constraining) spo w.r.t. which s is legal.

So a state is legal if every forward was preceded by its original message, and for every note sent in an email there is an explanation how the sender of the email learned this note.

3 Epistemic language and its semantics

We want to reason about the knowledge of the agents after an email exchange has taken place. For this purpose we use a language $\mathfrak{L}$ of communication and knowledge defined as follows:

$$\varphi ::= m \mid i \blacktriangleleft m \mid \neg\varphi \mid \varphi \wedge \varphi \mid C_G\varphi$$

The formula m expresses the fact that m has been sent in the past, with some unknown group of BCC recipients. The formula $i \blacktriangleleft m$ expresses the fact that agent i was involved in a full version of the message m, i.e., he was either

the sender, a recipient or a BCC recipient. The formula $C_G\varphi$ denotes common knowledge of the formula φ in the group G.

We use the usual abbreviations $\vee$, $\rightarrow$ and $\leftrightarrow$ and we use $K_i\varphi$ as an abbreviation of $C_{\{i\}}\varphi$. The fact that an email with a certain set of BCC recipients was sent can be expressed in our language by the following abbreviation:

$$m_B ::= m \wedge \bigwedge_{i \in S(m) \cup R(m) \cup B} i \blacktriangleleft m \wedge \bigwedge_{i \notin S(m) \cup R(m) \cup B} \neg i \blacktriangleleft m$$

Note that this formula expresses the fact that the message m was sent with exactly the group B as BCC recipients, which captures precisely the intended meaning of m_B.

We now provide a semantics for this language interpreted on legal states, inspired by the epistemic logic and the history-based approaches of [9] and [10]. For every agent i we define an indistinguishability relation $\sim_i$, where we intend $s \sim_i s'$ to mean that agent i cannot distinguish between the states s and s'. We first define this relation on the level of emails as follows (recall that we assume that senders and regular recipients are not BCC recipients):

$$m_B \sim_i m'_{B'}$$

iff one of the following contingencies holds:

(i) $i \in S(m)$, $m = m'$ and $B = B'$,
(ii) $i \in R(m) \setminus S(m)$ and $m = m'$,
(iii) $i \in B \cap B'$, and $m = m'$,
(iv) $i \notin S(m) \cup R(m) \cup B$ and $i \notin S(m') \cup R(m') \cup B'$.

Condition (i) states that the sender of an email confuses it only with the email itself. In turn, condition (ii) states that each regular recipient of an email who is not a sender confuses it with any email with the same message but possibly sent to a different BCC group. Next, condition (iii) states that each BCC recipient of an email confuses it with any email with the same message but sent to a possibly different BCC group of which he is also a member. Finally, condition (iv) states that each agent confuses any two emails in which he is not involved.

As an example consider the emails $e := s(i, l, j)_\emptyset$ and $e' := s(i, l, j)_{\{k\}}$. We have then $e \not\sim_i e'$, $e \sim_j e'$ and $e \not\sim_k e'$. Intuitively, agent j cannot distinguish between these two emails because he cannot see whether k is a BCC recipient. In contrast, agents i and k can distinguish between these two emails.

Next, we extend the indistinguishability relation to legal states by defining

$$(E, L) \sim_i (E', L')$$

iff the following holds:

- $L_i = L'_i$,
- for any $m_B \in E$ such that $i \in S(m) \cup R(m) \cup B$ there is $m_{B'} \in E'$ such that $m_B \sim_i m_{B'}$,

– for any $m_{B'} \in E'$ such that $i \in S(m) \cup R(m) \cup B'$ there is $m_B \in E$ such that $m_B \sim_i m_{B'}$.

So two states cannot be distinguished by agent i if they agree on his notes and their email sets look the same to him. Since we assume that the agents do not know anything about the other notes, we do not refer to the sets of notes of the other agents. Note that $\sim_i$ is an equivalence relation.

As an example consider the legal states s_1 and s_2 which are identical apart from their sets of emails:

$$E_{s_1} := \{s(i,l,j)_\emptyset, f(j, s(i,l,j), o)_\emptyset\},$$
$$E_{s_2} := \{s(i,l,j)_{\{k\}}, f(j, s(i,l,j), o)_\emptyset, f(k, s(i,l,j), o)_\emptyset\}.$$

We assume here that $l \in L_i$. The corresponding causality orderings clarify that in the first state agent i sends a message with note l to agent j and then j forwards this message to agent o. Further, in the second state agent i sends the same message but with a BCC to agent k, and then both agent j and agent k forward the message to agent o.

From the above definition it follows that $s_1 \not\sim_i s_2$, $s_1 \sim_j s_2$, $s_1 \not\sim_k s_2$ and $s_1 \not\sim_o s_2$. For example, the first claim holds because, as noticed above, $s(i,l,j)_\emptyset \not\sim_i s(i,l,j)_{\{k\}}$. Intuitively, in state s_1 agent i is aware that he sent a BCC to nobody, while in state s_2 he is aware that he sent a BCC to agent k.

In order to express common knowledge, we define for a group of agents G the relation $\sim_G$ as the reflexive, transitive closure of $\bigcup_{i \in G} \sim_i$. Then we define the truth of a formula from our language in a state inductively as follows, where $s = (E, L)$:

$$\begin{array}{ll}
s \models m & \text{iff } \exists B : m_B \in E \\
s \models i \blacktriangleleft m & \text{iff } \exists B : m_B \in E \text{ and } i \in S(m) \cup R(m) \cup B \\
s \models \neg\varphi & \text{iff } s \not\models \varphi \\
s \models \varphi \wedge \psi & \text{iff } s \models \varphi \text{ and } s \models \psi \\
s \models C_G\varphi & \text{iff } s' \models \varphi \text{ for any legal state } s' \text{ such that } s \sim_G s'
\end{array}$$

We say that φ is ***valid*** (and often just write 'φ' instead of 'φ is valid') if for all legal states s, $s \models \varphi$.

The following lemma clarifies when specific formulas are valid. In the sequel we shall use these observations implicitly.

Lemma 1.

(i) $m \to m'$ *is valid iff* $m = m'$ *or* m' *is part of the message* m.
(ii) $m \to i \blacktriangleleft m'$ *is valid iff* $i \in S(m') \cup R(m')$ *or for some note* l *and group* G, $f(i, l.m', G)$ *is part of the message* m.

The second item states that $m \to i \blacktriangleleft m'$ is valid either if i is a sender or a receiver of m' (in that case actually $i \blacktriangleleft m'$ is valid) or i forwarded the message m'. The latter is also possible if i was a BCC receiver of m'. The claimed equivalence holds thanks to condition L.1.

To illustrate this definition let us return to the above example. In state s_2 agent j does not know that agent k received the message $s(i,l,j)$ since he cannot distinguish s_2 from the state s_1 in which agent k did not receive this message. So $s_2 \models \neg K_j k \blacktriangleleft s(i,l,j)$ holds.

On the other hand, in every legal state s_3 such that $s_2 \sim_o s_3$ both an email $f(k, s(i,l,j), o)_C$ and a 'justifying' email $s(i,l,j)_B$ have to exist such that $s(i,l,j)_B \prec f(k, s(i,l,j), o)_C$ and $k \in B$. Consequently $s_3 \models k \blacktriangleleft s(i,l,j)$, so $s_2 \models K_o k \blacktriangleleft s(i,l,j)$ holds, so by sending the forward agent k revealed himself to o as a BCC recipient.

We leave to the reader checking that both $s_2 \models C_{\{k,o\}} k \blacktriangleleft s(i,l,j)$ and $s_2 \models \neg C_{\{j,o\}} k \blacktriangleleft s(i,l,j)$ holds. In words, agents k and o have common knowledge that agent k was involved in a full version of the message $s(i,l,j)$, while the agents j and o don't.

4 Common knowledge

We now clarify when a group of agents acquires common knowledge of the formula expressing that an email was sent. This shows how we can use our framework to investigate epistemic consequences of email exchanges.

Given a set of emails E and a group of agents A, let

$$E_A := \{m_B \in E \mid A \subseteq S(m) \cup R(m) \text{ or } \exists j \in B : (A \subseteq S(m) \cup \{j\})\}.$$

When $e \in E_A$ we shall say that the email e is ***shared by the group*** A. Note that when $|A| \geq 3$, then $e \in E_A$ iff $A \subseteq S(m) \cup R(m)$. When $|A| = 2$, then $e \in E_A$ also when $\exists j \in B : A = S(m) \cup \{j\}$, and when $|A| = 1$, then $e \in E_A$ also when $A = S(m)$ or $\exists j \in B : A = \{j\}$.

The following theorem summarizes our results.

Main Theorem Consider a legal state $s = (E, L)$ and a group of agents A.

(i) $s \models C_A m$ iff there is $m'_{B'} \in E_A$ such that $m' \to m$ is valid.
(ii) Suppose that $|A| \geq 3$. Then $s \models C_A m_B$ iff the following hold:
C1 $Ag = S(m) \cup R(m) \cup B$,
C2 for each $i \in B$ there is $m'_{B'} \in E_A$ such that $m' \to i \blacktriangleleft m$ is valid,
C3 there is $m'_{B'} \in E_A$ such that $m' \to m$ is valid.

In words, $s \models C_A m_B$ iff

- the email m_B involves all agents,
- there is an email shared by the group A that proves the existence of the message m,
- for every agent i that is on the BCC list of m_B there is an email shared by the group A that proves that i forwarded message m.

As an aside let us mention that there is a corresponding result for the case when $|A| < 3$, as well. However, it involves a tedious case analysis concerning the relation between $A, S(m), R(m)$ and B, so we do not present it here.

5 Proof of the Main Theorem

We establish first a number of auxiliary lemmas. We shall use a new strict partial ordering on emails. We define

$$m_B < m'_{B'} \text{ iff } m \neq m' \text{ and } m' \to m.$$

Note that $m' \to m$ precisely if m' is a forward, or a forward of a forward, etc, of m. Then for two emails m_B and $m_{B'}$ from a legal state s with the causality ordering $\prec$, $m_B < m_{B'}$ implies $m_B \prec m_{B'}$ on the account of condition L.1. However, the converse does not need to hold since $m_B \prec m_{B'}$ can hold on the account of L.2 or L.3. Further, note that the $<$-maximal elements of E are precisely the emails in E that are not forwarded.

Given a set of emails E and $E' \subseteq E$ we then define the ***downward closure*** of E' by

$$E'_{\leq} := E' \cup \{e \in E \mid \exists e' \in E' : e < e'\}.$$

The set of emails E on which the downward closure of E' depends will always be clear from the context.

Next, we introduce two operations on states. Assume a state (E, L) and an email $m_B \in E$.

We define the state

$$s \setminus m_B := (E \setminus \{m_B\}, L'),$$

with

$$L'_i := \begin{cases} L_i \cup FI(m) \text{ if } i \in R(m) \cup B \\ L_i \qquad\qquad \text{otherwise} \end{cases}$$

Intuitively, $s \setminus m_B$ is the result of removing the email m_B from the state s, followed by augmenting the sets of notes of its recipients in such a way that they initially already had the knowledge they would have acquired from m_B. Note that $s \setminus m_B$ is a legal state if m_B is an $<$-maximal element of E.

Next, given $C \subseteq B$ we define the state

$$s[m_{B \mapsto C}] := (E \setminus \{m_B\} \cup \{m_C\}, L'),$$

with

$$L'_i := \begin{cases} L_i \cup FI(m) \text{ if } i \in B \setminus C \\ L_i \qquad\qquad \text{otherwise} \end{cases}$$

Intuitively, $s[m_{B \mapsto C}]$ is the result of shrinking the set of BCC recipients of m from B to C, followed by an appropriate augmenting of the sets of notes of the agents that no longer receive m.

Note that $s[m_{B \mapsto C}]$ is a legal state if there is no forward of m by an agent $i \in B \setminus C$, i.e., no email of the form $f(i, l.m, G)_D$ exists in E such that $i \in B \setminus C$.

We shall need the following lemma that clarifies the importance of the set E_A of emails.

Lemma 2. *Consider a legal state $s = (E, L)$ and a group of agents A. Then for some L' the state $s' := ((E_A)_{\leq}, L')$ is legal and $s \sim_A s'$.*

Proof. We prove that for all $<$-maximal emails $m_B \in E$ such that $m_B \notin E_A$ (so neither $A \subseteq S(m) \cup R(m)$ nor $\exists i \in B : (A \subseteq S(m) \cup \{i\})$) we have $s \sim_A s \setminus m_B$. Iterating this process we get the desired conclusion.

Suppose m_B is a $<$-maximal email in E such that $m_B \notin E_A$. Take some $j \in A \setminus (S(m) \cup R(m))$. Suppose first $j \notin B$. Then $s \sim_j s \setminus m_B$ so $s \sim_A s \setminus m_B$.

Suppose now $j \in B$. Define

$$s_1 := s[m_{B \mapsto \{j\}}].$$

Then s_1 is a legal state and $s \sim_j s_1$. Next, define

$$s_2 := s[m_{B \mapsto \emptyset}].$$

Now take some $k \in A \setminus (S(m) \cup \{j\})$. Then $s_1 \sim_k s_2 \sim_j s \setminus m_B$ so $s \sim_A s \setminus m_B$. Note that both s_1 and s_2 are legal states since m_B is $<$-maximal. □

Using the above lemma we now establish two auxiliary results concerning common knowledge of the formula $i \blacktriangleleft m$ or of its negation.

Lemma 3.

(i) $s \models C_A i \blacktriangleleft m$ *iff* $\exists m'_B \in E_A : (m' \rightarrow i \blacktriangleleft m)$
or $(A \subseteq S(m) \cup \{i\}$ *and* $\exists m_B \in E_A : (i \in B))$.
(ii) $s \models C_A \neg i \blacktriangleleft m$ *iff* $s \models \neg\, i \blacktriangleleft m$ *and* $(A \subseteq S(m) \cup \{i\}$ *or* $s \models C_A \neg m)$.

To illustrate various alternatives listed in (i) note that each of the following emails in E ensures that $s \models C_{\{j\}} i \blacktriangleleft m$, where in each case m is the corresponding send message:

$$s(i, l, G)_{\{j\}}, f(k, q.s(i, l, G), H)_{\{j\}},$$
$$s(k, l, i)_{\{j\}},\ f(i, q.s(k, l, G), H)_{\{j\}},\ s(j, l, G)_{\{i\}}.$$

The first four of these emails imply $s \models C_{\{j\}} i \blacktriangleleft m$ by the first clause of (i), the last one by the second clause.

Proof. (i) $(\Rightarrow)$ Suppose $s \models C_A i \blacktriangleleft m$. Take the legal state s' constructed in Lemma 2. Then $s \sim_A s'$, so $s' \models i \blacktriangleleft m$.

Hence for some group B we have $m_B \in (E_A)_{\leq}$ and $i \in S(m) \cup R(m) \cup B$. Three cases arise.

Case 1. $i \in S(m) \cup R(m)$.

Then $m \rightarrow i \blacktriangleleft m$. So if $m_B \in E_A$, then the claim holds. Otherwise some email $m'_{B'} \in E_A$ exists such that $m_B < m'_{B'}$. Consequently $m' \rightarrow m$ and hence $m' \rightarrow i \blacktriangleleft m$. So the claim holds as well.

Case 2. $i \notin S(m) \cup R(m)$ and $A \subseteq S(m) \cup \{i\}$.

Then $i \in B$ since $i \in S(m) \cup R(m) \cup B$. Then by the definition of E_A, $m_B \in E_A$ so the claim holds.

Case 3. $i \notin S(m) \cup R(m)$ and $\neg(A \subseteq S(m) \cup \{i\})$.

If for some note l and groups G and C we have $f(i, l.m, G)_C \in (E_A)_{\leq}$, then either $f(i, l.m, G)_C \in E_A$ or for some $m'_{B'} \in E_A$ we have $f(i, l.m, G)_C < m'_{B'}$. In the former case we use the fact that the implication $f(i, l.m, G) \rightarrow i \blacktriangleleft m$ is valid. In the latter case $m' \rightarrow f(i, l.m, G)$ and hence $m' \rightarrow i \blacktriangleleft m$. So in both cases the claim holds.

Otherwise let $s'' = s'[m_{B \mapsto B \setminus \{i\}}]$. Note that s'' is a legal state because i does not forward m in s'. Take some $j \in A \setminus (S(m) \cup \{i\})$. Then $s' \sim_j s''$, so $s \sim_A s''$. Moreover, $s'' \models \neg i \blacktriangleleft m$, which yields a contradiction. So this case cannot arise.

($\Leftarrow$) The claim follows directly by the definition of semantics. We provide a proof for one representative case. Suppose that for some email $m'_B \in E_A$ both $A \subseteq S(m') \cup R(m')$ and $m' \rightarrow i \blacktriangleleft m$. Take some legal state s' such that $s \sim_A s'$. Then for some group B' we have $m'_{B'} \in E_{s'}$. So $s' \models m'$ and hence $s' \models i \blacktriangleleft m$. Consequently $s \models C_A i \blacktriangleleft m$.

(*ii*) Let $s = (E, L)$.
($\Rightarrow$) Suppose $s \models C_A \neg i \blacktriangleleft m$. Then $s \models \neg i \blacktriangleleft m$. Assume $A \not\subseteq S(m) \cup \{i\}$ and $s \not\models C_A \neg m$. Then there is some legal state $s' = (E', L')$ such that $s \sim_A s'$ and $s' \models m$. Then there is some group B such that $m_B \in E'$. Let $j \in A \setminus (S(m) \cup \{i\})$ and let $s'' = (E' \setminus \{m_B\} \cup \{m_{B \cup \{i\}}\}, L')$. Then $s' \sim_j s''$ so $s \sim_A s''$. But $s'' \models i \blacktriangleleft m$ which contradicts our assumption.

($\Leftarrow$) Suppose that $s \models \neg i \blacktriangleleft m$ and either $A \subseteq S(m) \cup \{i\}$ or $s \models C_A \neg m$. We first consider the case that $A \subseteq S(m) \cup \{i\}$. Let s' be any legal state such that $s \sim_A s'$. Assume $s' \models i \blacktriangleleft m$. Then $m_B \in E_{s'}$ for some group B such that $i \in B$. Since $A \subseteq S(m) \cup \{i\}$, any legal state s'' such that $s' \sim_A s''$ contains an email $m_C \in E_{s''}$ for some group C such that $i \in C$. So $s'' \models i \blacktriangleleft m$. In particular, this holds for the state s, which contradicts our assumption. So $s' \models \neg s(i, n, G)$ and hence $s \models C_A \neg s(i, n, G)$.

Now we consider the case that $s \models C_A \neg m$. Let s' be such that $s \sim_A s'$. Then $s' \models \neg m$. Since $i \blacktriangleleft m \rightarrow m$ is valid, we get $s' \models \neg i \blacktriangleleft m$. So $s \models C_A \neg i \blacktriangleleft m$. $\square$

We are now ready to prove the Main Theorem.

Proof

(*i*) ($\Rightarrow$) Suppose $s \models C_A m$. Take the legal state s' constructed in Lemma 2. Then $s \sim_A s'$, so $s' \models m$. So for some group B we have $m_B \in (E_A)_{\leq}$.

Hence either $m_B \in E_A$ or some email $m'_{B'} \in E_A$ exists such that $m_B < m'_{B'}$. In both cases the claim holds.

($\Leftarrow$) Suppose that for some email $m'_B \in E_A$ we have $m' \rightarrow m$. Take some legal state s' such that $s \sim_A s'$. Then by the form of E_A and the definition of semantics for some group B' we have $m'_{B'} \in E_{s'}$. So $s' \models m'$ and hence $s' \models m$. Consequently $s \models C_A m$.

(ii) By the definition of m_B, the fact that the C_A operator distributes over the conjunction, part (i) of the Main Theorem and Lemma 3 we have

$$s \models C_A m_B \text{ iff } \mathbf{C3}\text{-}\mathbf{C6},$$

where

C4 $\bigwedge_{i \in S(m) \cup R(m) \cup B}$ $((A \subseteq S(m) \cup \{i\}$ and $\exists B' : (m_{B'} \in E_A$ and $i \in B'))$ or $\exists m'_{B'} \in E_A : (m' \to i \blacktriangleleft m))$,
C5 $\bigwedge_{i \notin S(m) \cup R(m) \cup B}$ $(A \subseteq S(m) \cup \{i\}$ or $s \models C_A \neg m)$,
C6 $s \models \bigwedge_{i \notin S(m) \cup R(m) \cup B} \neg i \blacktriangleleft m$.

$(\Rightarrow)$ Suppose $s \models C_A m_B$. Then properties **C3**-**C6** hold. But $|A| \geq 3$ and $s \models C_A m$ imply that no conjunct of **C5** holds. Hence property **C1** holds.

Further, since $|A| \geq 3$ the first disjunct of each conjunct in **C4** does not hold. So the second disjunct of each conjunct in **C4** holds, which implies property **C2**.

$(\Leftarrow)$ Suppose properties **C1**-**C3** hold. It suffices to establish properties **C4**-**C6**.

For $i \in S(m) \cup R(m)$ we have $m \to i \blacktriangleleft m$. So **C2** implies property **C4**. Further, since **C1** holds, properties **C5** and **C6** hold vacuously. □

6 Analysis of BCC

In our framework we built emails out of messages using the BCC feature. So it is natural to analyze whether and in what sense the emails can be reduced to messages without BCC recipients.

Given a send email $s(i, l, G)_B$, where $B = \{j_1, \ldots, j_k\}$, we can simulate it by the following sequence of messages:

$$s(i, l, G), f(i, s(i, l, G), j_1), \ldots, f(i, s(i, l, G), j_k).$$

Analogous simulation can be formed for the forward email $f(i, l.m, G)_B$.

In what follows we clarify in what sense this simulation is correct. Below, given a message m we write $f(S(m), m, j)$ for $f(i, m, \{j\})$, where $S(m) = \{i\}$.

Definition 1. *Given a state $s = (E, L)$ such that there is no forward of the message m by the agent j in E, we define $rem_j^m(s)$ as follows:*

- *if $m_B \in E$ for some group B and $j \in B$ and $f(S(m), m, j)_C \notin E$ for any group C then*

$$rem_j^m(s) := (E \backslash m_B \cup \{m_{B \backslash \{j\}}, f(S(m), m, j)_\emptyset\}, L),$$

- *if $m_B \in E$ for some group B and $j \in B$ and $f(S(m), m, j)_C \in E$ for some group C then*

$$rem_j^m(s) := (E \backslash m_B \cup \{m_{B \backslash \{j\}}\}, L),$$

– *otherwise* $rem_j^m(s) := s$.

So, assuming $m_B \in s$ and $j \in B$, we form $rem_j^m(s)$ by replacing the email m_B by $m_{B\setminus\{j\}}$ when for some group C the forward $f(S(m), m, j)_C$ is present in E, or by $m_{B\setminus\{j\}}, f(S(m), m, j)_\emptyset$ when no such forward is present in E.

We assume that in E there is no forward of m by agent j, as otherwise the removal of j from the list of the BCC recipients would yield an illegal state. Indeed, for such a forward of the message m condition L.1 would not hold. In the remainder of this section we assume that such forwards by former BCC recipients are not present.

We are currently working on a formal analysis of a simulation of BCC that does allow such forwards. It is obtained by replacing each such forward $f(j, m, G)$ by $f(j, f(i, m, j), G)$.

We now show that using the above operation $rem_j^m(s)$ we obtain a legal state that is almost equivalent to the original one. We establish first two lemmas concerning the relation between $rem_j^m(s)$ and the knowledge relation of some agent k.

Lemma 4. *For any two legal states s and t, message m and agent j, if $s \sim_k t$ then $rem_j^m(s) \sim_k rem_k^m(t)$.*

Proof. Omitted for the reasons of space. □

Lemma 5. *For any legal state s, message m and agent j, if there is some t' such that $rem_j^m(s) \sim_k t'$ then either $s \sim_k rem_j^m(s)$ or there is some t such that $s \sim_k t$ and $t' = rem_j^m(t)$.*

Proof. Let $s' = rem_j^m(s)$ and suppose $s' \sim_k t'$. If $s = s'$ then $s \sim_k s'$. Suppose otherwise. Then by the definition of $rem_j^m(s)$ we know that there is some group B such that $j \in B$, $m_B \in E_s$ and $m_{B\setminus\{j\}} \in E_{s'}$. Define $B' = B\setminus\{j\}$.

Suppose there is no full version of $f(S(m), m, j)$ in $E_{t'}$. By the definition of $rem_j^m(s)$, there is a full version of $f(S(m), m, j)$ in $E_{s'}$ so then we know that $k \notin S(m) \cup \{j\}$ because $s' \sim_k t'$. Clearly then $s \sim_k s'$.

Suppose there is a full version of $f(S(m), m, j)$ in $E_{t'}$. Then there is some group C such that $m_C \in E_{t'}$. Suppose $j \in C$. Since $m_{B'} \in E_{s'}$, $j \notin B'$ and $s' \sim_k t'$ this means $k \notin S(m) \cup \{j\}$. So $s \sim_k s'$.

Finally, suppose that $m_C \in E_{t'}$, $j \notin C$ and $f(S(m), m, j)_{C'} \in E_{t'}$. Suppose there is no full version of $f(S(m), m, j)$ in E_s. Define t as the state which is like t' but with $E_t = E_{t'}\setminus\{m_C, f(S(m), m, j)_{C'}\} \cup \{m_{C\cup\{j\}}\}$. Clearly, $rem_j^m(t) = t'$. We claim $s \sim_k t$. The condition on the notes is satisfied since the sets of notes in s and s' and in t and t' are identical, and $s' \sim_k t'$. We will to show that for any $m'_D \in E_s$ such that $k \in S(m') \cup R(m') \cup D$ there is some $m'_{D'} \in E_t$ such that $m'_D \sim_k m'_{D'}$. The proof in the other direction is very similar. Take such an m'_D.

Suppose $m'_D = m_B$. We know $m_{B'} \in E_{s'}$ so $m_{B'} \sim_k m_C$. Since $B = B' \cup \{j\}$ then clearly $m_B \sim_k m_{C\cup\{j\}}$ and we know $m_{C\cup\{j\}} \in E_t$ so let $m'_{D'} = m_{C\cup\{j\}}$.

Suppose otherwise. Then $m'_D \in E_{s'}$ so there is $m'_{D'} \in E_{t'}$ such that $m'_D \sim_k m'_{D'}$. We know that $m' \neq m$ and $m' \neq f(S(m), m, j)$ because no full version of $f(S(m), m, j))$ is in E_s so then $m'_{D'} \in E_t$.

Finally, suppose that for some group E, $f(S(m), m, j)_E \in E_s$. Let $E_t = E_{t'} \backslash \{m_C\} \cup \{m_{C \cup \{j\}}\}$. The proof is very similar.
For the case that $m'_D = f(S(m), m, j)_E$, note that $f(S(m), m, j)_E \in E_{s'}$ so $f(S(m), m, j)_E \sim_k f(S(m), m, j)_{C'}$, so let $m'_{D'} = f(S(m), m, j)_{C'}$. □

The theorem below shows that our operation of removing a BCC recipient results in a state that is equivalent for all formulas that do not explicitly mention the newly added forward or the fact that this BCC recipient received the original message.

Theorem 1. *For any state s, message m, agent j and formula φ that does not mention $j \blacktriangleleft m$ or $f(i, m, j)$, $s \models \varphi$ iff $rem^m_j(s) \models \varphi$.*

Proof. We proceed by induction on the structure of φ. The only interesting case is when $\varphi = C_G\psi$.

Suppose $rem^m_j(s) \models C_G\psi$. Let $s \sim_G t$ for some group of agents G. Then there must be a path $s \sim_{j_1} s_1 \sim_{j_2} ... \sim_{j_n} t$, with $j_1, ..., j_n \in G$. Then by Lemma 4, $rem^m_j(s) \sim_{j_1} rem^m_j(s_1) \sim_{j_2} ... \sim_{j_n} rem^m_j(t)$. Hence $rem^m_j(s) \models C_G\psi$ implies that $rem^m_j(t) \models \psi$. By the induction hypothesis, $t \models \psi$. So $s \models C_G\psi$.

Suppose $s \models C_G\psi$. If $s \sim_G rem^m_j(s)$ then clearly $rem^m_j(s) \models C_G\psi$. Suppose otherwise. Let $rem^m_j(s) \sim_G t'$ for some state t'. Then there is a path $rem^m_j(s) = s'_0 \sim_{k_1} s'_1 \sim_{k_2} ... \sim_{k_n} s'_n = t'$, with $k_1, ..., k_n \in G$. We claim that for any s'_i there is a state s_i such that $s \sim_G s_i$ and $rem^m_j(s_i) = s'_i$. We will proceed by induction. Clearly the claim holds for $s'_0 = rem^m_j(s)$. Suppose it holds for s'_{i-1}, so $s \sim_G s_{i-1}$ and $rem^m_j(s_{i-1}) = s'_{i-1}$ for some state s_{i-1}. By Lemma 5 either $s_{i-1} \sim_{k_i} s'_{i-1}$ or there is s_i such that $s_{i-1} \sim_{k_i} s_i$ and $rem^m_j(s_i) = s'_i$. In the first case, since $s \sim_G s_{i-1}$ and $k_i \in G$ we have $s \sim_G s'_{i-1}$ and since $rem^m_j(s) \sim_G s'_{i-1}$ we have $s \sim_G rem^m_j(s)$ which contradicts our assumption. In the second case, $s \sim_G s_i$ so our claim holds. So then it also holds for $s'_n = t'$, and there is some t such that $s \sim_G t$ and $rem^m_j(t) = t'$. But then by assumption $t \models \psi$ and by the induction hypothesis $t' \models \psi$. So $rem^m_j(s) \models C_G\psi$.

Clearly, by repeatedly applying above construction we obtain the simulation of BCC given above. The corollary below shows that in the original and the resulting state the status of the statement that there is common knowledge of the underlying message is the same.

Definition 2. *For a state s, a message m and a group of agents $B = \{j_1, ..., j_n\}$ such that $m_B \in E_s$, we define*

$$rem^m_B(s) := rem^m_{j_1}(rem^m_{j_2}(...rem^m_{j_n}(s))).$$

Corollary 1. *For any legal state s, a group of agents A and an email $m_B \in E_s$ such that $rem^m_B(s)$ is a legal state*

$$s \models C_A m \text{ iff } rem^m_B(s) \models C_A m.$$

7 Conclusions and future work

Email is by now one of the most common forms of group communication. This motivates the study here presented. The language we introduced allowed us to discuss various fine points of email communication, notably forwarding and the use of BCC. The epistemic semantics we proposed aimed at clarifying the knowledge-theoretic consequences of this form of communication. Our presentation focused on the issue of common knowledge aimed at clarifying when a group of agents has a common knowledge of an email.

This framework also leads to natural questions concerning axiomatization of the language and decidability of the semantics. Currently we work on

- a sound and complete axiomatization of the epistemic language $\mathfrak{L}$ of Section 3; at this stage we have such an axiomatization for the epistemic free formulas,
- the problem of decidability of the truth definition given in Section 3; at this stage we have a decidability result for positive formulas.

Acknowledgements

We acknowledge helpful discussions with Jan van Eijck and Rohit Parikh and useful referee comments.

References

1. K. R. Apt, A. Witzel, and J. A. Zvesper. Common knowledge in interaction structures. In *Proceedings of TARK XII*, pages 4–13. The ACM Digital Library, 2009.
2. E-mail leak of degree inflation. BBC News, 2008. Available at `http://news.bbc.co.uk/2/hi/uk_news/education/7483330.stm`.
3. J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(11):1620–1662, 2006.
4. K. M. Chandy and J. Misra. How processes learn. *Distributed Computing*, 1(1):40–52, March 1986.
5. R. Fagin, J. Halpern, M. Vardi, and Y. Moses. *Reasoning about knowledge*. MIT Press, Cambridge, MA, USA, 1995.
6. L. Lamport. Time, clocks, and the ordering of events in a distributed system. *Commun. ACM*, 21(7):558–565, 1978.
7. Wouter Bos e-mailt per ongeluk zijn netwerk rond. NRC Handelsblad, 7th October 2010, 2010. In Dutch.
8. E. Pacuit. Logics of informational attitudes and informative actions. Manuscript, University of Tilburg, 2010.
9. E. Pacuit and R. Parikh. Reasoning about communication graphs. *Interactive Logic. Proceedings of the 7th Augustus de Morgan Workshop*, pages 135–157, 2007.
10. R. Parikh and R. Ramanujam. A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12(4):453–467, 2003.
11. J. van Eijck and F. Sietsma. Message passing in a dynamic epistemic logic setting. In *Proceedings of TARK XIII*. The ACM Digital Library, 2011. To appear.
12. Y. Wang, F. Sietsma, and J. van Eijck. Logic of information flow on communication channels. In *Proceedings of AAMAS-10*, pages 1447–1448, 2010.