EPiC Series in Computing

Volume 57, 2018, Pages 36–55

LPAR-22. 22nd International Conference on Logic for Programming, Artificial Intelligence and Reasoning

When Are Two Gossips the Same?

Krzysztof R. Apt[1,2], Davide Grossi[3], and Wiebe van der Hoek[4]

[1] CWI, Amsterdam, The Netherlands
[2] MIMUW, University of Warsaw, Warsaw, Poland
[3] Bernoulli Institute, University of Groningen, Groningen, The Netherlands
[4] Department of Computer Science, University of Liverpool, Liverpool, U.K.

Abstract

We provide an in-depth study of the knowledge-theoretic aspects of communication in so-called gossip protocols. Pairs of agents communicate by means of calls in order to spread information—so-called secrets—within the group. Depending on the nature of such calls knowledge spreads in different ways within the group. Systematizing existing literature, we identify 18 different types of communication, and model their epistemic effects through corresponding indistinguishability relations. We then provide a classification of these relations and show its usefulness for an epistemic analysis in presence of different communication types. Finally, we explain how to formalise the assumption that the agents have common knowledge of a distributed epistemic gossip protocol.

1 Introduction

In the gossip problem [33, 7] a number of agents, each one knowing a piece of information (a *secret*) unknown to the others, communicate by one-to-one interactions (e.g., telephone calls). The result of each call is that the two agents involved in it learn all secrets the other agent knows at the time of the call. The problem consists in finding a sequence of calls which disseminates all the secrets among the agents in the group. It sparked a large literature in the 70s and 80s [33, 7, 18, 9, 31], typically on establishing—in the above and other variants of the problem—the minimum number of calls to achieve dissemination of all the secrets. This number has been proven to be $2n - 4$, where n, the number of agents, is at least 4.

The gossip problem constitutes an excellent toy problem to study information dissemination in distributed environments. A vast literature on distributed protocols has taken up the problem and analyzed it together with a wealth of variations including different communication primitives (e.g., broadcasting instead of one-to-one calls), as well as communication structures (networks), faulty communication channels [10], and probabilistic information transmission, where the spreading of gossips is used to model the spread of an epidemic [6, 30]. Surveys are [15, 24, 21, 25].

Background The present paper investigates a knowledge-based approach to the gossip problem in a multi-agent system. Agents perform calls following individual epistemic protocols they run in a distributed fashion. These protocols tell the agents which calls to execute depending

on what they know, or do not know, about the information state of the agents in the group. We call the resulting distributed programs *epistemic gossip protocols*, or *gossip protocols*, for short. Such protocols were introduced and studied in [5, 1]. 'Distributed' means that each agent acts autonomously, and 'epistemic' means that the gossip protocols refer to the agents' knowledge. The reliance of these protocols on epistemic properties makes them examples of so-called knowledge-based protocols, as studied in the context of distributed systems [29, 27, 20, 13].

Besides the aforementioned [5, 1], a number of papers have recently focused on epistemic gossip protocols. In [23] gossip protocols were studied that aim at achieving higher-order shared knowledge, for example knowledge of level 2 which stipulates that everybody knows that everybody knows all secrets. In particular, a protocol is presented and proved correct that achieves in $(k+1)(n-2)$ steps shared knowledge of level k. Further, in [11] gossip protocols were studied as an instance of multi-agent epistemic planning that is subsequently translated into the classical planning language PDDL. More recently, [34] presented a study of *dynamic* gossip protocols in which the calls allow the agents not only to share the secrets but also to share the communication channels (that is, who can call whom). In turn, [3] studied the computational complexity of distributed epistemic gossip protocols, while [4] showed that implementability, partial correctness, termination, and fair termination of these protocols is decidable.

More broadly, the paper positions itself within the long-standing tradition of analysis of distributed systems from the perspective of epistemic logic [14, 28]. Such a perspective has led in [29, 27, 13] to a useful level of abstraction allowing one to study a number of topics in distributed computing from the knowledge theoretic perspective, in particular protocols for the sequence transmission problem (for instance the alternating bit protocol) in [20], coordination [19], and secure communication [8], to mention some. The characteristic feature of these programs is that they use tests for knowledge.

Contributions The form of communication underpinning the epistemic gossip problem may vary from work to work, and the above papers sometimes make different assumptions on the nature of communication upon which the considered protocols are based. Little attention has been devoted to a systematic analysis, with the notable exception of [17], which singled out some of the key informational assumptions on calls—specifically observability, synchronicity and asynchronicity assumptions—and systematically studied the effects of such assumptions on the aforementioned $2n - 4$ call-length bound.

It is our claim that research on epistemic gossip protocols can at this point benefit from a systematisation of the key possible assumptions that a modeler can make on the type of communication (call) underpinning such protocols. From an epistemic logic point of view, each call type induces a specific notion of knowledge. The comparison of the resulting definitions of knowledge is of obvious importance for the study of epistemic aspects of communication.

By 'type of communication' we mean the way in which communication takes place and may be observed, and to focus on it we disregard the type of information exchanged (in particular, whether higher order knowledge, or communication links may be exchanged—matters we do not address), or the type of information the agents have initially at their disposal (e.g., whether it is common knowledge what the number of agents is).

More specifically, here are the features we focus on. First of all, a call between two agents takes place in the presence of other agents. What these other agents become aware of after the call is one natural parameter. We call it *privacy*. The second parameter, that we call *direction*, clarifies in which direction the information flows. Here we focus on three possibilities: they exchange all information, one agent passes all information to the other one, or one agent acquires all information available to the other one. The final parameter of a call is what we call

observance. It determines whether the agent(s) affected by the call learn what information was held by the other agent prior to the call.

By a *call type* we mean a combination of these three parameters. What the agents know after a call, or more generally a sequence of calls, depends on the assumed call type. This yields in total 18 possibilities. The paper provides a framework in which we model these possibilities in a unified way. This allows us to provide in Theorem 5.1 a complete classification of the resulting indistinguishability relations. This in turn makes it possible to clarify in Propositions 6.1 and 6.2 the effect of a call type on the truth of the considered formulas. Additionally, we provide in Proposition 6.7 a natural proposal on how to incorporate into this framework an assumption that the agents have common knowledge of the underlying protocol.

Paper outline Section 2 introduces gossip protocols by example, and identifies the features of calls we will focus on. Section 3 introduces the syntax and semantics of a simple epistemic language to study communication and its effects in gossip protocols, together with some motivating examples. Crucially the semantics introduced is parametrised by the indistinguishability relations which, for each call type, identify the call sequences that the agents cannot distinguish. These equivalence relations are systematically introduced and defined in Section 4, and then compared in terms of their relative informativeness in Section 5. The proposed systematisation is then applied in Section 6: first, to deliver general results on the analysis of how knowledge depends on the assumed call types (Section 6.1); second, to offer a natural approach to the problem of modelling common knowledge of protocols in the epistemic gossip setting (Section 6.2). Finally, Section 7 summarises our results and charts several directions for future research.

2 A Typology of Calls

We start by recalling the notion of a gossip protocol, moving then to introduce the formal set-up of the paper.

2.1 Gossip protocols

Gossip protocols aim at sharing knowledge between agents in a pre-described way. In a distributed protocol several agents may decide to initiate a call at the same time. Let us now consider such an epistemic protocol, so one in which the agents refer to their knowledge.

Protocol 1 (Hear my secret)**.** *Any agent a calls agent b if a does not know that b is familiar with a's secret.*

This protocol has been proven in [1] to terminate and be correct, under specific assumptions on the type of communication taking place during each call. In this paper we aim at providing a systematic presentation of such assumptions and at an analysis of their logical interdependencies.

Throughout the paper we assume a fixed finite set Ag of at least three ***agents***. We further assume that each agent holds exactly one ***secret*** and that the secrets are pairwise different. We denote by S the set of all secrets, the secret of agent a by A, the secret of agent b by B, and so on. A secret can be any piece of data, for instance birthday, salary or social security number. Furthermore, we assume that each secret carries information identifying the agent to whom this secret belongs. So once agent b learns secret A she knows that this is the secret of agent a.

2.2 Calls

In the context of gossip protocols calls constitute the sole form of knowledge acquisition the agents have at their disposal. Each ***call*** concerns two agents, the *caller* (a, below) and the *callee* (b, below). We call a the *partner* of b in the call, and vice versa. Any agent c different from a and b is called an *outsider*. We study the following properties of calls:

- ***privacy***, which is concerned with what the outsiders note about the call,
- ***direction***, which clarifies the direction of the information flow in the call,
- ***observance***, which clarifies, when an agent a is informed by b, whether a sees b's secrets before adding them to her own set, or only sees the result of the fusion of the two sets of secrets.

More specifically, we distinguish three ***privacy degrees*** of a call where agent a calls b:

- $\circ$: every agent $c \neq a, b$ notes that a calls b,
- $\boldsymbol{\ominus}$: every agent $c \neq a, b$ notes that some call takes place, though not between whom,
- $\bullet$: no agent $c \neq a, b$ notes that a call is taking place.

Intuitively, these degrees can be ordered as $\circ <_{\mathsf{p}} \boldsymbol{\ominus} <_{\mathsf{p}} \bullet$, with $\circ$ meaning no privacy at all, $\boldsymbol{\ominus}$ ensuring anonymity of the caller and callee, and $\bullet$ denoting full privacy.

We distinguish three ***direction types***, in short ***directions***, of a call:

- ***push***, written as $\triangleright$. As a result of the call the callee learns all the secrets held by the caller.
- ***pull***, written as $\triangleleft$. As a result of the call the caller learns all the secrets held by the callee.
- ***push-pull***, written as $\diamond$. As a result of the call the caller and the callee learn each other's secrets.

Depending on the direction of a call between a and b, one or both agents can learn *directly* new information thanks to it. We say that these are the agents *affected* in the call. More formally, an agent a is ***affected*** by a call c if c is one of the following forms:

$$a \diamond b, b \diamond a, b \triangleright a, \textit{or } a \triangleleft b.$$

Intuitively, a is affected by the call if it can affect the set of secrets a is familiar with. This brings us to two possible levels of ***observance*** of a call:

- α: During the call the affected agent(s) add the secrets of their partner to their own secrets, and only *after* that, inspect the result.
- β: During the call the affected agent(s) inspect the secrets of their partner *before* adding them to their own secrets.

Intuitively, the observance level α is less informative for an affected agent than β, because in the latter case she also learns which secrets were known to the other agent before adding them to the secrets she is familiar with. Let

- $\mathsf{P} = \{\circ, \boldsymbol{\circ}, \bullet\}$,
- $\mathsf{D} = \{\diamond, \triangleleft, \triangleright\}$,
- $\mathsf{O} = \{\alpha, \beta\}$.

Each call between agents a and b is of the shape ab^{τ}, where $\tau = (\mathsf{p}, \mathsf{d}, \mathsf{o}) \in \mathsf{P} \times \mathsf{D} \times \mathsf{O}$ is called its ***type***. So we defined in total 18 call types. To clarify their effect on communication we will elaborate on some representative call types in Examples 3.3 and 3.4.

The types $(\circ, \diamond, \beta)$ and $(\boldsymbol{\circ}, \diamond, \beta)$ were studied in [5] while the types $(\bullet, \diamond, \alpha)$, $(\bullet, \triangleright, \alpha)$, and $(\bullet, \triangleleft, \alpha)$, were analyzed in [1]. For a type τ like $(\circ, \diamond, \beta)$, we define $\tau(\mathsf{p}) = \circ, \tau(\mathsf{d}) = \diamond$ and $\tau(\mathsf{o}) = \beta$.

Often, the call type (or parts of it) is (are) clear from the context, and we omit it (them). In our examples, at the level of calls, we often only explicitly mention the direction type. Given a call between a and b we shall sometimes write it simply as ab for the direction type $\diamond$, $a \triangleright b$ for the direction type $\triangleright$ and $a \triangleleft b$ for the direction type $\triangleleft$.

3 Language and Semantics

In this section we introduce a modal language for epistemic gossip and its formal semantics.

3.1 Modal language

We are interested in determining agents' knowledge after a sequence of calls took place. To this end we use the following modal language $\mathcal{L}$ for epistemic logic:

$$\phi \quad ::= \quad F_a S \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid K_a \phi,$$

where $a \in Ag$ and $S \in \mathsf{S}$.

In what follows we refer to the elements ϕ of $\mathcal{L}$ as ***epistemic formulas***, or in short, just formulas. We read $F_a S$ as 'agent a is familiar with the secret S' (or 'S belongs to the set of secrets a has learned') and $K_a \phi$ as 'agent a knows that formula ϕ is true'. So $\mathcal{L}$ is an epistemic language with the atomic formulas of the form $F_a S$. The above language was introduced in [1]. It is a modification of the language introduced in [5].

Example 3.1. Consider the statement that agent a is familiar with all the secrets. This can be expressed as the formula

$$\bigwedge_{b \in Ag} F_a B$$

that we subsequently abbreviate to Exp_a ("a is an expert").

Here and elsewhere for simplicity we refer in the conjunction limits only to agents and not to their secrets. This convention allows us to write more complex statements, for instance that each agent is familiar only with her own secret. This can be expressed as the formula

$$\bigwedge_{a \in Ag} (F_a A \wedge \bigwedge_{b \in Ag, b \neq a} \neg F_a B). \tag{1}$$

□

Next, we clarify the use of the knowledge operators. In the presented reasoning we assume that the agents have the knowledge of the underlying call type. In all cases we assume that the initial situation is the one in which every agent is only familiar with her own secret, that is, we assume (1) to be true for each agent before any communication takes place. The examples provide intuitions about how agents' knowledge is influenced by the types of calls underpinning their communications. Such intuitions will then be formalised in Section 3.2.

Example 3.2. Initially, each agent is familiar with her secret and each agent knows this fact. Additionally, she does not know that any other agent is familiar with a secret different from her own. This can be expressed by means of the formula

$$\bigwedge_{a\in Ag} (\bigwedge_{b\in Ag} K_a F_b B \wedge \bigwedge_{b,c\in Ag, a\neq b, b\neq c} \neg K_a F_b C)$$

that holds initially, for all call types. □

Example 3.3. Suppose there are three agents, a, b and c. Consider the two call types $(\bullet, \diamond, \mathsf{o})$, where $\mathsf{o} \in \mathsf{O}$, and assume the call sequence ac, bc, ab. After it the agents a and b (and c too) are familiar with all the secrets, which can be expressed as the formula

$$\phi = Exp_a \wedge Exp_b,$$

and both know this fact, which can be expressed as $K_a\phi \wedge K_b\phi$.

If the observance of the calls is β, agent a also learns that prior to the call ab agent b was familiar with a's secret, i.e., with A. This allows a to conclude that agent b was involved in a call with c and hence agent c is familiar with B. We can express this as

$$K_a F_c B.$$

Contrast the above with the situation when the observance is α. Although again after the considered call sequence both agents a and b are familiar with all the secrets, now agent a cannot conclude that agents b and c communicated. Hence agent a does not know whether agent c is familiar with B, i.e., the formula $K_a F_c B$ is not true. □

Example 3.4. Assume the same call sequence as in the previous example but suppose that the call parameters are now $(\circ, \lhd, \mathsf{o})$, where $\mathsf{o} \in \mathsf{O}$. So we consider now the call sequence $\mathbf{c} = a \lhd c, b \lhd c, a \lhd b$.

Because of the assumed privacy level, after this call sequence agent a knows that agent b learned the secret C and agent c knows that agent a learned the secret B, i.e., the following holds after $\mathbf{c}$

$$K_a F_b C \wedge K_c F_a B.$$

Suppose now the privacy degree is $\boldsymbol{\ominus}$ and the observance is β. Then we only have $K_a F_b C$ as agent a cannot distinguish $\mathbf{c}$ from $a \lhd c, c \lhd b, a \lhd b$. Clearly, $K_c F_a B$ does not hold after $\mathbf{c}$ as agent c cannot distinguish $\mathbf{c}$ from $a \lhd c, c \lhd b, b \lhd a$.

Finally, if the privacy degree is $\bullet$ then for the same reason $K_c F_a B$ does not hold after $\mathbf{c}$ either. □

We conclude that what the agents know after a call sequence crucially depends on the parameters of the calls. Further, the precise effect of a single call on the agents' knowledge is very subtle, both for the agents involved in it and for the outsiders.

3.2 Semantics

We provide now a formal semantics for the modal language $\mathcal{L}$. It is parameterized by a call type τ.

Gossip situations and calls First we recall the following crucial notions introduced in [1]. A ***gossip situation*** is a sequence $\mathsf{s} = (\mathsf{Q}_a)_{a \in Ag}$, where $\mathsf{Q}_a \subseteq \mathsf{S}$ for each agent a. Intuitively, Q_a is the set of secrets agent a is familiar with in the situation s. Given a gossip situation $\mathsf{s} = (\mathsf{Q}_a)_{a \in Ag}$, we denote Q_a by s_a. The ***initial gossip situation*** is the one in which each Q_a equals $\{A\}$ and is denoted by i (for "initial"). The initial gossip situation reflects the fact that initially each agent is familiar only with her own secret.

Each call transforms the current gossip situation by possibly modifying the set of secrets the agents involved in the call are familiar with. The definition depends solely on the direction of the call.

Definition 3.5. The application of a call c to a gossip situation s is defined as follows, where $\mathsf{s} := (\mathsf{Q}_a)_{a \in Ag}$:

$\boxed{\mathsf{c} = ab}$ $\mathsf{c}(\mathsf{s}) = (Q'_a)_{a \in Ag}$, where $\mathsf{Q}'_a = \mathsf{Q}'_b = \mathsf{Q}_a \cup \mathsf{Q}_b$, $\mathsf{Q}'_c = \mathsf{Q}_c$, for $c \neq a, b$.

$\boxed{\mathsf{c} = a \triangleright b}$ $\mathsf{c}(\mathsf{s}) = (Q'_a)_{a \in Ag}$, where $\mathsf{Q}'_b = \mathsf{Q}_a \cup \mathsf{Q}_b$, $\mathsf{Q}'_a = \mathsf{Q}_a$, $\mathsf{Q}'_c = \mathsf{Q}_c$, for $c \neq a, b$.

$\boxed{\mathsf{c} = a \triangleleft b}$ $\mathsf{c}(\mathsf{s}) = (Q'_a)_{a \in Ag}$, where $\mathsf{Q}'_a = \mathsf{Q}_a \cup \mathsf{Q}_b$, $\mathsf{Q}'_b = \mathsf{Q}_b$, $\mathsf{Q}'_c = \mathsf{Q}_c$, for $c \neq a, b$.

This definition captures the meaning of the direction type: for ab the secrets are shared between the caller and callee , for $a \triangleright b$ they are pushed from the caller to the callee, and for $a \triangleleft b$ they are retrieved by the caller from the callee. Note that $(a \diamond b)(\mathsf{s}) = (b \diamond a)(\mathsf{s})$ and $(a \triangleright b)(\mathsf{s}) = (b \triangleleft a)(\mathsf{s})$, as expected.

In turn, the privacy degree of a call captures what outsiders of the call learn from it and the observance level determines informally what caller and callee can learn about each other's calling history. The meaning of these two parameters will be determined by means of the appropriate equivalence relations between call sequences.

A ***call sequence*** is a *finite* sequence of calls, all of the same call type. The empty sequence is denoted by ϵ. We use $\mathbf{c}$ to denote a call sequence and $\mathbf{C}^\tau$ to denote the set of all call sequences of call type τ. Given the call sequence $\mathbf{c}$ and a call c, $\mathbf{c}.\mathsf{c}$ denotes the sequence obtained by appending $\mathbf{c}$ with c.

The result of applying a call sequence $\mathbf{c}$ to a situation s is defined by induction using Definition 3.5, as follows

[Base] $\epsilon(\mathsf{s}) := \mathsf{s}$,

[Step] $\mathbf{c}.\mathsf{c}(\mathsf{s}) := \mathsf{c}(\mathbf{c}(\mathsf{s}))$.

Note that this definition does not depend on the privacy degree and observance of the calls.

Truth of formulas We illustrated in Examples 3.3 and 3.4 that each call has an effect on the knowledge of the agents. After a sequence of calls took place the agents may be uncertain about the current gossip situation because they do not know which call sequence actually took place. This leads to appropriate indistinguishability relations that allow us to reason about the knowledge of the agents. This is in a nutshell the basis of the approach to epistemic gossip protocols put forth in [1], and upon which we build here.

In general, to determine what agents know after a call sequence we need to consider an appropriate equivalence relation between the call sequences. Let $\mathbf{c}$ and $\mathbf{d}$ be two call sequences of call type τ and a an agent. The statement $\mathbf{c} \sim_a^\tau \mathbf{d}$ informally says that agent a cannot distinguish between $\mathbf{c}$ and $\mathbf{d}$. The definition of $\sim_a^\tau$ crucially depends on the call type τ and is provided in the next subsection. Here we assume that it is given and proceed to define the truth of the formulas of the language $\mathcal{L}$ with respect to a ***gossip model*** (for a given set of agents Ag) $\mathcal{M}^\tau = (\mathbf{C}^\tau, \{\sim_a^\tau\}_{a \in Ag})$ and a call sequence $\mathbf{c}$ as follows:

Definition 3.6. Let $\mathcal{M}^\tau$ be a gossip model for a call type τ and a set of agents Ag, and let $\mathbf{c} \in \mathbf{C}^\tau$. The truth relation for language $\mathcal{L}$ is inductively defined as follows (with Boolean connectives omitted):

$$\begin{array}{lll} (\mathcal{M}^\tau, \mathbf{c}) \models F_a S & \text{iff} & S \in \mathbf{c}(\mathsf{i})_a, \\ (\mathcal{M}^\tau, \mathbf{c}) \models K_a \phi & \text{iff} & \forall \mathbf{d} \in \mathbf{C}^\tau \text{ such that } \mathbf{c} \sim_a^\tau \mathbf{d},\ (\mathcal{M}^\tau, \mathbf{d}) \models \phi. \end{array}$$

Since the gossip model is clear from the context, we will from now on write $\mathbf{c} \models^\tau \phi$ for $(\mathcal{M}^\tau, \mathbf{c}) \models \phi$. We also write $\mathcal{M}^\tau \models \phi$ (ϕ is valid in $\mathcal{M}^\tau$) if for all $\mathbf{c} \in \mathbf{C}^\tau$ we have $\mathcal{M}^\tau, \mathbf{c} \models \phi$.

So the formula $F_a S$ is true after a sequence of calls $\mathbf{c}$ whenever agent a is familiar with the secret S in the gossip situation generated by $\mathbf{c}$ applied to the initial gossip situation i. The knowledge operator K_a is interpreted as is customary in the multimodal $S5_n$ logic (see, e.g., [12]), so using the equivalence relation $\sim_a^\tau$.

It is important to notice that to determine the truth of a propositional formula (so in particular to determine which secrets an agent is familiar with) only the direction parameter of the type of the calls is used. In contrast, to determine the truth of formulas involving the knowledge operator all three parameters of the call type are needed, through the definition of the $\sim_a^\tau$ relations, to which we turn next.

4 Indistinguishability of Call Sequences

Below we say that an agent a is ***involved*** in a call c, and write $a \in \mathsf{c}$, if a is one of the two agents involved in it, i.e., if it is either a caller or a callee in c. So agent a is involved but not affected (a notion introduced in Section 2) by a call c if $\mathsf{c} = a \triangleright b$ or $\mathsf{c} = b \triangleleft a$ for some agent b.

4.1 The $\sim_a^\tau$ relations

For every call type τ and agent a we define the indistinguishability relation $\sim_a^\tau \subseteq \mathbf{C}^\tau \times \mathbf{C}^\tau$ in two steps. First we define the auxiliary relation $\approx_a^\tau$ (Definition 4.1). Intuitively, the expression $\mathbf{c} \approx_a^\tau \mathbf{d}$ can be interpreted as "from the point of view of a, if $\mathbf{c}$ is an (epistemically) possible call sequence, so is $\mathbf{d}$, and vice versa". Then, we define $\sim_a^\tau$ as the least equivalence relation that contains $\approx_a^\tau$.

Definition 4.1. Let $a \in Ag$ and fix a type τ. The relation $\approx_a^\tau$ is the smallest subset of $\mathbf{C}^\tau \times \mathbf{C}^\tau$ satisfying the following conditions:

[Base] $\epsilon \approx_a^\tau \epsilon$.

[Step] Suppose that $\mathbf{c} \approx_a^\tau \mathbf{d}$ and let c and d be calls.

$$\begin{array}{ll} [\text{Step-out}^\tau] & \text{if } \mathit{Out}_a^\tau(\mathsf{c}, \mathsf{d}) \text{ then } \mathit{Concl}_a^\tau(\mathbf{c}, \mathbf{d}, \mathsf{c}, \mathsf{d}), \\ [\text{Step-in}^\tau] & \text{if } \mathit{In}_a^\tau(\mathbf{c}, \mathbf{d}, \mathsf{c}) \text{ then } \mathit{Concl}_a^\tau(\mathbf{c}, \mathbf{d}, \mathsf{c}), \end{array}$$

where the used relations are defined in Table 1. (b is there the partner of a in the call c.)

Agent a is not involved in the last call:

$\tau(\mathsf{p})$	$Out_a^\tau(\mathsf{c},\mathsf{d})$	$Concl_a^\tau(\mathbf{c},\mathbf{d},\mathsf{c},\mathsf{d})$
○	$a \notin \mathsf{c}$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$
◓	$a \notin \mathsf{c}, a \notin \mathsf{d}$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{d}$
●	$a \notin \mathsf{c}$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}$, $\mathbf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$

Agent a is involved in but not affected by the last call:

$In_a^\tau(\mathbf{c},\mathbf{d},\mathsf{c})$	$Concl_a^\tau(\mathbf{c},\mathbf{d},\mathsf{c})$
$\mathsf{c} \in \{a \triangleright b, b \triangleleft a\}$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$

Agent a is involved in and affected by the last call:

$\tau(\mathsf{o})$	$In_a^\tau(\mathbf{c},\mathbf{d},\mathsf{c})$	$Concl_a^\tau(\mathbf{c},\mathbf{d},\mathsf{c})$
α	$\mathsf{c} \in \{a \diamond b, b \diamond a, b \triangleright a, a \triangleleft b\}$, $\mathbf{c}.\mathsf{c}(\mathsf{i})_a = \mathbf{d}.\mathsf{c}(\mathsf{i})_a$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$
β	$\mathsf{c} \in \{a \diamond b, b \diamond a, b \triangleright a, a \triangleleft b\}$, $\mathbf{c}(\mathsf{i})_b = \mathbf{d}(\mathsf{i})_b$	$\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$

Table 1: Defining indistinguishability of call sequences

The definition of $\approx_a^\tau$ captures the complex effect of each of the three parameters of a call type on the knowledge of an agent. Let us discuss it now in detail.

The Base condition is clear. Consider now the Step-out$^\tau$ clause which refers to Table 1, top. Suppose that $\mathbf{c} \approx_a^\tau \mathbf{d}$. Consider first the privacy type ○. According to its informal description the condition $a \notin \mathsf{c}$ means that agent a is not involved in the call c but knows who calls whom. The conclusion $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$ then coincides with this intuition.

Consider now the privacy type ◓. The conditions $a \notin \mathsf{c}$ and $a \notin \mathsf{d}$ mean that agent a is not involved in the calls c and d, thus according to the informal description of ◓ she cannot distinguish between these two calls. This explains the conclusion $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{d}$. Note that this conclusion is not justified for the privacy type ○ because if $\mathsf{c} \neq \mathsf{d}$ then agent a can distinguish between these two calls, so a fortiori between the call sequences $\mathbf{c}.\mathsf{c}$ and $\mathbf{d}.\mathsf{d}$.

Finally, consider the privacy type ●. According to its informal description, the condition $a \notin \mathsf{c}$ means that agent a is not aware of the call c. This justifies the conclusions $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}$ and $\mathbf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$.

Next, consider the Step-in$^\tau$ clause. It spells the conditions that allow one to extend the $\approx_a^\tau$ relation in case agent a is involved in the last call, c. Table 1, middle, formalises the intuition that when agent a is not affected by the call c, then we can conclude that $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$.

Table 1, bottom, focuses on the remaining case. Consider first the observance α. According to its informal description, affected agents incorporate the secrets of their partner with their own secrets and then inspect the result. So we check what secrets agent a is familiar with after the call sequences $\mathbf{c}$ and $\mathbf{d}$ are both extended by c. If these sets are equal, then we can conclude that $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$.

In the case the observance is β, the informal description stipulates that the agent inspects the set of secrets of the call partner before incorporating them with their own secrets. So we

compare these sets of secrets after, respectively, the call sequences $\mathbf{c}$ and $\mathbf{d}$ took place. If these sets are equal, then we conclude that $\mathbf{c}.\mathsf{c} \approx_a^\tau \mathbf{d}.\mathsf{c}$. This explains why in this case a reference to agent b is made in $In_a^\tau(\mathbf{c}, \mathbf{d}, \mathsf{c})$.

4.2 Some observations

The following observations clarify some properties of the indistinguishability relations $\sim_a^\tau$.

Note 4.2. *For all agents a and call types τ*

$$\sim_a^\tau = (\approx_a^\tau)^*,$$

where * is the transitive, reflexive closure operation on binary relations.

Proof. A straightforward proof by induction show that each $\approx_a^\tau$ relation is symmetric. This implies the claim. □

Proposition 4.3. *For all call types τ if $\mathbf{c} \sim_a^\tau \mathbf{d}$, then $\mathbf{c}(\mathsf{i})_a = \mathbf{d}(\mathsf{i})_a$.*

Proof. By Note 4.2 it is sufficient to prove the conclusion under the assumption that $\mathbf{c} \approx_a^\tau \mathbf{d}$.

We proceed by induction on the sum k of the lengths $|\mathbf{c}| + |\mathbf{d}|$ of both sequences. If $k = 0$, then $\mathbf{c} = \mathbf{d} = \epsilon$, so the claim holds. Suppose the claim holds for all pairs of sequences such that the sum of their lengths is $< k$ and that $k > 0$, $|\mathbf{c}| + |\mathbf{d}| = k$ and $\mathbf{c} \approx_a^\tau \mathbf{d}$. By definition $\approx_a^\tau$ is the smallest relation satisfying the Base and Step conditions of Definition 4.1. Let c be the last call of $\mathbf{c}$ or of $\mathbf{d}$ if $\mathbf{c}$ is empty.

If agent a is not involved in c, then four cases arise, depending on the form of $\mathbf{c}$ and $\mathbf{d}$. We consider one representative case, when $\mathbf{c}$ is of the form $\mathbf{c}'.\mathsf{c}$, where $\mathbf{c}' \approx_a^\tau \mathbf{d}$. Then by the assumption about c and the induction hypothesis

$$\mathbf{c}(\mathsf{i})_a = \mathbf{c}'.\mathsf{c}(\mathsf{i})_a = \mathbf{c}'(\mathsf{i})_a = \mathbf{d}(\mathsf{i})_a.$$

If agent a is involved in but not affected by the last call, then $\mathbf{c}$ is of the form $\mathbf{c}'.\mathsf{c}$, $\mathbf{d}$ is of the form $\mathbf{d}'.\mathsf{c}$, $\mathsf{c} \in \{a \triangleright b, b \triangleleft a\}$ and $\mathbf{c}' \approx_a^\tau \mathbf{d}'$. Then by the form of c and the induction hypothesis

$$\mathbf{c}(\mathsf{i})_a = \mathbf{c}'.\mathsf{c}(\mathsf{i})_a = \mathbf{c}'(\mathsf{i})_a = \mathbf{d}'(\mathsf{i})_a = \mathbf{d}'.\mathsf{c}(\mathsf{i})_a = \mathbf{d}(\mathsf{i})_a.$$

Finally, if agent a is involved in and affected by the last call, then $\mathbf{c}$ is of the form $\mathbf{c}'.\mathsf{c}$, $\mathbf{d}$ is of the form $\mathbf{d}'.\mathsf{c}$, $\mathsf{c} \in \{a \diamond b, b \diamond a, b \triangleright a, a \triangleleft b\}$ and $\mathbf{c}' \approx_a^\tau \mathbf{d}'$.

If $\tau(\mathsf{o}) = \alpha$, then by assumption $\mathbf{c}'.\mathsf{c}(\mathsf{i})_a = \mathbf{d}'.\mathsf{c}(\mathsf{i})_a$, i.e., $\mathbf{c}(\mathsf{i})_a = \mathbf{d}(\mathsf{i})_a$. If $\tau(\mathsf{o}) = \beta$, then by assumption $\mathbf{c}'(\mathsf{i})_b = \mathbf{d}'(\mathsf{i})_b$. Also, by the induction hypothesis $\mathbf{c}'(\mathsf{i})_a = \mathbf{d}'(\mathsf{i})_a$, so by the form of c

$$\mathbf{c}(\mathsf{i})_a = \mathbf{c}'.\mathsf{c}(\mathsf{i})_a = \mathbf{c}'(\mathsf{i})_a \cup \mathbf{c}'(\mathsf{i})_b = \mathbf{d}'(\mathsf{i})_a \cup \mathbf{d}'(\mathsf{i})_b = \mathbf{d}'.\mathsf{c}(\mathsf{i})_a = \mathbf{d}(\mathsf{i})_a.$$

□

Corollary 4.4. *For all call types τ, agents a, b and call sequences $\mathbf{c}$*

$$\mathbf{c} \models^\tau K_a F_a B \text{ iff } \mathbf{c} \models^\tau F_a B.$$

Proof. By Proposition 4.3 and the definition of truth of $K_a F_a B$ and $F_a B$. □

5 Classification of the $\sim_a^\tau$ Relations

We introduced in the previous section 18 equivalence relations $\sim_a^\tau$, each parametrised by an agent a. The uniform presentation makes it possible to compare these relations by means of a classification, which we now provide.

Given two call types τ_1 and τ_2 we abbreviate the statement $\forall a \in Ag, \sim_a^{\tau_1} \subset \sim_a^{\tau_2}$ to $\tau_1 \subset \tau_2$ and similarly for $\tau_1 \subseteq \tau_2$ and $\tau_1 = \tau_2$. Such statements presuppose that we systematically change the types of all calls in the considered call sequences.

The following theorem provides the announced classification. It clarifies in total 153 ($= \frac{18 \cdot 17}{2}$) relationships between the equivalence relations.

Theorem 5.1. *The $\sim_a^\tau$ equivalence relations form preorders presented in Figures 1 and 2. An arrow $\rightarrow$ from τ_1 to τ_2 stands here for $\tau_1 \subset \tau_2$, $(\circ, \mathsf{d}, \mathsf{o})$ for the set of six call types with the privacy degree $\circ$ that are all equal, and $(\bullet, \diamond, \mathsf{o})$ for the set $\{(\bullet, \diamond, \alpha), (\bullet, \diamond, \beta)\}$.*

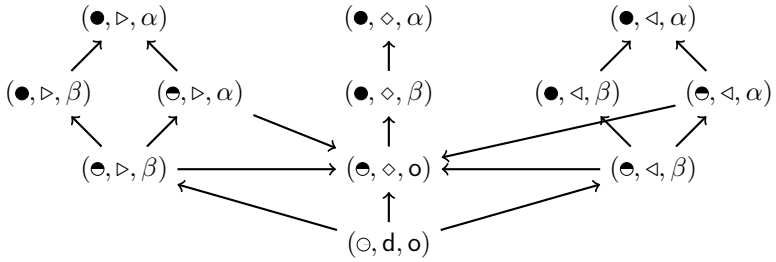

Figure 1: Classification of the $\sim_a^\tau$ relations when $|Ag| = 3$.

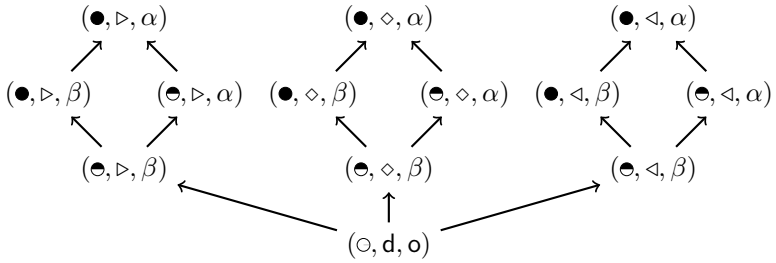

Figure 2: Classification of the $\sim_a^\tau$ relations when $|Ag| > 3$.

The proof of Theorem 5.1 relies on the three lemmas stated below, the proofs of which can be found in the full version of the paper, [2]. We say here that the call types τ_1 and τ_2 are ***incomparable*** when neither $\tau_1 \subseteq \tau_2$ nor $\tau_2 \subseteq \tau_1$ holds. The proofs concerning the incomparability that are established below also hold for a stronger definition, namely that τ_1 and τ_2 are incomparable when for all agents a neither $\sim_a^{\tau_1} \subseteq \sim_a^{\tau_2}$ nor $\sim_a^{\tau_2} \subseteq \sim_a^{\tau_1}$ holds. This way Figures 1 and 2 can be alternatively interpreted as preorders on the $\sim_a^\tau$ equivalence relations, for any agent a, where an arrow $\rightarrow$ from τ_1 to τ_2 stands then for $\sim_a^{\tau_1} \subset \sim_a^{\tau_2}$.

Lemma 5.2.

(i) *Suppose that* $\tau(\mathsf{p}) = \circ$. *Then each* $\sim_a^\tau$ *is the identity relation.*

(ii) *Suppose that* $\tau_1(\mathsf{p}) = \tau_2(\mathsf{p}) = \circ$. *Then* $\tau_1 = \tau_2$.

(iii) *If* $|Ag| = 3$ *then* $(\mathbin{\rotatebox{180}{◓}}, \diamond, \beta) = (\mathbin{\rotatebox{180}{◓}}, \diamond, \alpha)$.

Below the unspecified parameters are implicitly universally qualified. For example, $(\circ, \mathsf{d}, \mathsf{o}) \subset (\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \mathsf{o})$ is an abbreviation for the statement

$$\forall a \in Ag\ \forall \mathsf{d} \in \mathsf{D}\ \forall \mathsf{o} \in \mathsf{O}\ \sim_a^{(\circ,\mathsf{d},\mathsf{o})} \subset \sim_a^{(\mathbin{\rotatebox{180}{◓}},\mathsf{d},\mathsf{o})}.$$

Lemma 5.3.

(i) $(\circ, \mathsf{d}, \mathsf{o}) \subset (\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \mathsf{o})$.

(ii) $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \mathsf{o}) \subset (\bullet, \mathsf{d}, \mathsf{o})$.

(iii) *If* $|Ag| > 3$ *or* $\mathsf{d} \neq \diamond$ *then* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \beta) \subset (\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \alpha)$.

(iv) *If* $|Ag| = 3$, $\mathsf{d} \neq \diamond$ *and* $\mathsf{o}_1, \mathsf{o}_2 \in \mathsf{O}$, *then* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \mathsf{o}_1) \subset (\mathbin{\rotatebox{180}{◓}}, \diamond, \mathsf{o}_2)$.

(v) $(\bullet, \mathsf{d}, \beta) \subset (\bullet, \mathsf{d}, \alpha)$.

Lemma 5.4. *Let* $\mathsf{d}, \mathsf{d}_1, \mathsf{d}_2 \in \mathsf{D}$ *and* $\mathsf{o}_1, \mathsf{o}_2 \in \mathsf{O}$.

(i) *Suppose that* $|Ag| > 3$ *or* $\diamond \notin \{\mathsf{d}_1, \mathsf{d}_2\}$, *and* $\mathsf{d}_1 \neq \mathsf{d}_2$. *Then* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}_1, \mathsf{o}_1)$ *and* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}_2, \mathsf{o}_2)$ *are incomparable.*

(ii) *Suppose that* $\mathsf{d}_1 \neq \mathsf{d}_2$. *Then* $(\bullet, \mathsf{d}_1, \mathsf{o}_1)$ *and* $(\bullet, \mathsf{d}_2, \mathsf{o}_2)$ *are incomparable.*

(iii) *Suppose that* $|Ag| = 3$ *and* $\mathsf{d} \neq \diamond$. *Then* $(\mathbin{\rotatebox{180}{◓}}, \diamond, \alpha)$ *and* $(\bullet, \mathsf{d}, \alpha)$ *are incomparable.*

(iv) *Suppose that* $|Ag| > 3$ *or* $\diamond \notin \{\mathsf{d}_1, \mathsf{d}_2\}$, *and* $\mathsf{d}_1 \neq \mathsf{d}_2$. *Then* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}_1, \beta)$ *and* $(\bullet, \mathsf{d}_2, \alpha)$ *are incomparable.*

(v) *Suppose that* $|Ag| > 3$ *or* $\mathsf{d} \neq \diamond$. *Then* $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}, \alpha)$ *and* $(\bullet, \mathsf{d}, \beta)$ *are incomparable.*

The above Lemmas imply the classification of the $\sim_a^\tau$ relations given in Theorem 5.1 and visualized in Figures 1 and 2. Indeed, the equalities (represented as sets) are established in Lemma 5.2, the strict inclusions (that correspond to the arrows) are established in Lemma 5.3, and Lemma 5.4 implies that no further strict inclusions (i.e., arrows) are present. For example, there is no arrow in Figure 2 between two different diamond shaped subgraphs that correspond to the direction types $\triangleright, \diamond$, and $\triangleleft$ because by Lemma 5.4(*iv*) for $\mathsf{d}_1 \neq \mathsf{d}_2$ the call types $(\mathbin{\rotatebox{180}{◓}}, \mathsf{d}_1, \beta)$ and $(\bullet, \mathsf{d}_2, \alpha)$ are incomparable.

6 Applications of the Classification

The section shows how the above systematisation of $\sim_a^\tau$ relations, through the standard epistemic logic semantics of Definition 3.6, enables general insights into the epistemic effects of call sequences and offers a natural handle on how to model assumptions to the effect that agents have common knowledge of the protocol in use.

6.1 Epistemic effects of communication types

The above classification is useful in order to draw general epistemic consequences in presence of different communication types. Below we will be using two fragments of $\mathcal{L}$:

- $\mathcal{L}_1^+$, consisting of the *literals* F_aS and $\neg F_aS$, $\wedge, \vee$ and K_a,
- $\mathcal{L}_2^+$, consisting of the atomic formulas F_aS, $\wedge, \vee$ and K_a.

Proposition 6.1. *Consider two call types τ_1 and τ_2 such that $\tau_1(\mathsf{d}) = \tau_2(\mathsf{d})$.*

(i) For all literals ψ and all $\mathbf{c}$, $\mathbf{c} \models^{\tau_2} \psi \implies \mathbf{c} \models^{\tau_1} \psi$.

(ii) If $\tau_1 \subseteq \tau_2$ then

$$\text{for all formulas } \phi \in \mathcal{L}_1^+ \text{ and all } \mathbf{c},\ \mathbf{c} \models^{\tau_2} \phi \implies \mathbf{c} \models^{\tau_1} \phi.$$

Proof.
(i) By assumption $\tau_1(\mathsf{d}) = \tau_2(\mathsf{d})$, so both occurrences of $\mathbf{c}$ refer to identical call sequences. Hence for all atomic formulas F_aS and all $\mathbf{c}$, $\mathbf{c} \models^{\tau_2} F_aS$ iff $\mathbf{c} \models^{\tau_1} F_aS$.

(ii) We proceed by induction on the structure of ϕ. The only case that requires explanation is when ϕ is of the form $K_a\psi$. Suppose that $\mathbf{c} \models^{\tau_2} K_a\psi$. To prove $\mathbf{c} \models^{\tau_1} K_a\psi$ take a call sequence $\mathbf{d}$ such that $\mathbf{c} \sim_a^{\tau_1} \mathbf{d}$. By assumption $\tau_1 \subseteq \tau_2$, hence $\mathbf{c} \sim_a^{\tau_2} \mathbf{d}$ and so $\mathbf{d} \models^{\tau_2} \psi$. By the induction hypothesis $\mathbf{d} \models^{\tau_1} \psi$, so by definition $\mathbf{c} \models^{\tau_1} K_a\psi$. □

We finally compare knowledge for call types with different direction types. Then claim (i) in the above Proposition does not hold anymore. Indeed, for τ_1 and τ_2 such that $\tau_1(\mathsf{d}) = \diamond$ and $\tau_2(\mathsf{d}) = \triangleright$ we have $ab \models^{\tau_2} \neg F_aB$ but not $ab \models^{\tau_1} \neg F_aB$. However, the following weaker claim does hold.

Proposition 6.2. *Consider two call types τ_1 and τ_2 such that $\tau_1(\mathsf{d}) = \diamond$.*

(i) For all atomic formulas ψ and all $\mathbf{c}$, $\mathbf{c} \models^{\tau_2} \psi \implies \mathbf{c} \models^{\tau_1} \psi$.

(ii) If $\tau_1 \subseteq \tau_2$ then

$$\text{for all formulas } \phi \in \mathcal{L}_2^+ \text{ and all } \mathbf{c},\ \mathbf{c} \models^{\tau_2} \phi \implies \mathbf{c} \models^{\tau_1} \phi.$$

Proof. By Proposition 6.1 we can assume that $\tau_2(\mathsf{d}) \neq \diamond$.
(i) We use induction on the length $|\mathbf{c}|$ of $\mathbf{c}$. Assume that $\tau_2(\mathsf{d}) = \triangleright$. If $|\mathbf{c}| = 0$ then $\mathbf{c} = \epsilon$ and $\epsilon \models^{\tau_2} F_cD$ iff $D = C$ iff $\epsilon \models^{\tau_1} F_cD$. Now suppose the claim is proven for $\mathbf{c}$ and consider $\mathbf{c}.ab$.

For any agent $c \neq b$, we have by Definition 3.5 $\mathbf{c}.a \triangleright b \models^{\tau_2} F_cD$ iff $\mathbf{c} \models^{\tau_2} F_cD$, which implies by the induction hypothesis $\mathbf{c} \models^{\tau_1} F_cD$, and hence $\mathbf{c}.a\diamond b \models^{\tau_1} F_cD$. For agent b, we have $\mathbf{c}.a \triangleright b \models^{\tau_2} F_bD$ iff ($\mathbf{c} \models^{\tau_2} F_aD$ or $\mathbf{c} \models^{\tau_2} F_bD$) and $\mathbf{c}.a\diamond b \models^{\tau_1} F_bD$ iff ($\mathbf{c} \models^{\tau_1} F_aD$ or $\mathbf{c} \models^{\tau_1} F_bD$), so the claim for b holds by the induction hypothesis, as well.

The proof for $\tau_2(\mathsf{d}) = \triangleleft$ is analogous and omitted.

(ii) The claim follows by (i) and the argument used in the proof of Proposition 6.1. □

Proposition 6.1 holds for example for $\tau_1 = (\circ, \diamond, \beta)$ and $\tau_2 = (\bullet, \diamond, \alpha)$, since by Theorem 5.1

$$(\circ, \diamond, \beta) \subset (\text{◓}, \diamond, \beta) \subset (\bullet, \diamond, \beta) \subset (\bullet, \diamond, \alpha).$$

In turn, Proposition 6.2 holds for example for $\tau_1 = (\circ, \diamond, \beta)$ and $\tau_2 = (\bullet, \triangleright, \alpha)$, since by Theorem 5.1

$$(\circ, \diamond, \beta) = (\circ, \triangleright, \beta) \subset (\text{◓}, \triangleright, \beta) \subset (\bullet, \triangleright, \beta) \subset (\bullet, \triangleright, \alpha).$$

6.2 Common knowledge of protocols

When reasoning about specific protocols it is necessary to limit the set of considered call sequences to those that are 'legal' for it. When the agents form a graph given in advance one can simply limit the set of considered call sequences by allowing only syntactically legal calls. This affects the definition of semantics and can be of importance when reasoning about the correctness of specific protocols.

For example, in [1] a specific protocol for a directed ring is proved correct (Protocol R2 on page 61, for 3 or 4 agents) by allowing for each agent a only the calls between her and her successor $a \oplus 1$, and using the fact that the formula $K_a F_{a\oplus 1} A \ominus 1 \rightarrow F_a A \ominus 1$ is then true. Here $A \ominus 1$ is the secret of the predecessor of agent a, so this formula states that if agent a knows that her successor is familiar with the secret $A \ominus 1$ of her predecessor then agent a is familiar with the secret $A \ominus 1$.

A more challenging task is to incorporate into the framework an assumption that the agents have common knowledge of the underlying protocol.[1]

Example 6.3. Consider Protocol 1 (Hear my Secret) from Section 2 with the direction type $\diamond$. Recall that in this protocol an agent a can call agent b if $\neg K_a F_b a$ is true after the current call sequence. So each pair of agents can communicate at most once.

Assume now four agents a, b, c, d. Then the call sequence ab, bc, bd is compliant with the protocol independently on the assumptions about the privacy degree and observance. Let us analyse the situation after this call sequence took place.

Assume first the privacy degree $\circ$. Then agent c knows which calls took place and hence knows that after the third call agent d is familiar with her secret, C. So after these three calls agent c cannot call agent d anymore.

The situation changes when the privacy degree is $\bullet\!\!\!-$. Through the second call agent c learns the secret A, so she knows that the first call was ab or ba. Agent c is not involved in the third call, but by the assumed privacy degree she still knows that a third call has taken place.

Assume now that the agents have common knowledge of the protocol. So agent c knows that each pair of agents can communicate at most once. Hence she can conclude that d must be involved in the third call and consequently that the third call was between agent d and agent a or b. Agent c therefore now knows that after the third call agent d is familiar with at least 3 secrets: A, B, D if the call was with agent a or A, B, C, D if the call was with agent b. But agent c cannot anymore conclude that agent d is familiar with her secret, C, and consequently can call d.

Finally, consider the privacy degree $\bullet$. Then agent a does not know whether any calls took place after the call ab. In particular she cannot conclude that any of the agents c and d are familiar with her secret and hence can call either c or d. □

To discuss the matters further let us be more precise about the syntax of the protocols. An ***epistemic gossip protocol*** (in short a protocol) consists of the union of $|Ag|$ sets of instructions, one set for each agent. Each instruction is of the form

$$\textit{if } \phi \textit{ then} \text{ execute call } \mathsf{c},$$

in symbols $\phi \rightarrow \mathsf{c}$, where ϕ is a Boolean combination of formulas of the form $K_a\psi$, where a is the caller in the call c. The formula ϕ is referred to as an ***epistemic guard***. Such instructions

[1] This issue was identified as an open problem for epistemic gossip in [1]. The same issue manifests itself in other knowledge-based asynchronous protocols, such as the one investigated recently in [26].

are executed iteratively, where at each time one instruction is selected (at random, or based on some fairness considerations) whose guard is true after the call sequence executed so far.[2]

We therefore view a protocol P as a set of instructions $\phi \to \mathsf{c}$. For example, the instructions composing Protocol 1, are of the form

$$\neg K_a F_b A \to ab$$

for all agents a and b. That is, if a does not know whether b is not familiar with her secret, a calls b.

To justify the restriction on the syntax of the epistemic guards note the following observation.

Note 6.4. *Consider a call type τ such that $\tau(\mathsf{p}) = \bullet$. Then for all agents a, b, c and all call sequences $\mathbf{c}$ and formulas ϕ*

$$\mathbf{c} \models^\tau K_a\phi \text{ iff } \mathbf{c}.bc \models^\tau K_a\phi.$$

Consequently, the same equivalence holds for all formulas that are Boolean combinations of formulas of the form $K_a\phi$, so in particular for all epistemic guards used in the instructions for agent a.

Proof. By Definition 4.1 if the privacy type of τ is $\bullet$ then $\mathbf{c} \sim_a^\tau \mathbf{c}.bc$, which implies the claim. □

This note states that the calls in which agent a is not involved have no effect on the truth of the epistemic guards used in the instructions for agent a. If we allowed in the epistemic guards for agent a as conjuncts formulas not prefixed by K_a, this natural and desired property would not hold anymore.

Indeed, assume the privacy type $\bullet$ and consider the protocol for three agents, a, b, c, in which the only instructions are $\neg F_b A \wedge F_b C \to ab$ for agent a and $\neg F_b C \to bc$ for agent b. Then initially only the call bc can be performed. After it, the call ab can be performed upon which the protocol terminates. In other words, the call bc, of which agent a is not aware, affects the truth of its epistemic guard, which contradicts the idea behind the privacy type $\bullet$.

For the privacy type $\circ$ this restriction on the syntax of the epistemic guards is not needed as then all formulas are equivalent to the propositional ones.

Note 6.5. *Consider a call type τ such that $\tau(\mathsf{p}) = \circ$. Then for all agents a and all formulas ϕ and call sequences $\mathbf{c}$*

$$\mathbf{c} \models^\tau K_a\phi \text{ iff } \mathbf{c} \models^\tau \phi.$$

Proof. This is a direct consequence of the fact that when the privacy type of τ is $\circ$ then by Lemma 5.2(i) each relation $\sim_a^\tau$ is the identity. □

Let us return now to the matter of common knowledge of a protocol. In Definition 4.1 the τ-dependent indistinguishability relations are constructed assuming that any call is possible after any call sequence. This builds in the resulting gossip models $\mathcal{M}^\tau = (\mathbf{C}^\tau, \{\sim_a^\tau\}_{a \in Ag})$ the assumption that agents may consider any call sequence possible in principle, including calls that are not legal if we assume that the agents have common knowledge of the protocol in use.

Specifically, given a gossip model $\mathcal{M}^\tau = (\mathbf{C}^\tau, \{\sim_a^\tau\}_{a \in Ag})$ and a protocol P we define the ***computation tree*** $\mathbf{C}_P^\tau \subseteq \mathbf{C}^\tau$ of P (cf. [1]) as the set of call sequences inductively defined as follows:

[Base] $\epsilon \in \mathbf{C}_P^\tau$,

[2] This simple rendering of protocols suffices for the purposes of this section. More sophisticated formalizations of epistemic gossip protocols have been provided in [5, 1].

[Step] If $\mathbf{c} \in \mathbf{C}^\tau_P$ and $\mathbf{c} \models^\tau \phi$ then $\mathbf{c}.\mathsf{c} \in \mathbf{C}^\tau_P$, where $\phi \to \mathsf{c} \in P$.

So $\mathbf{C}^\tau_P$ is a (possibly infinite) set of finite call sequences that is iteratively obtained by performing a 'legal' call (according to protocol P) from a 'legal' (according to protocol P) call sequence. We refer to such legal call sequences as P-compliant.[3]

Note however, that when building such a computation tree, the epistemic guard ϕ is evaluated with respect to the underlying gossip model $\mathcal{M}^\tau$, which may well include call sequences that are not P-compliant. So in order to restrict the domain of the gossip model to only P-compliant sequences, the epistemic guards of the protocol need to be evaluated, and to do that one needs in turn a gossip model, which contains only P-compliant sequences.

To resolve this circularity we propose a solution showing how under a natural assumption on the syntax of the epistemic guards one can construct a gossip model which consists only of call sequences that are compliant with a given protocol P.

Fix till the end of the section an arbitrary call type τ. First, we introduce the definition of semantics relativised to a set $X \subseteq \mathbf{C}^\tau$ of call sequences. Let $\mathcal{M}^\tau_X = (X, \{\sim^\tau_a\}_{a \in Ag})$, where each $\sim^\tau_a$ relation is restricted to $X \times X$, and let $\mathbf{c} \in X$. Then the definition of semantics is the same as before with the except of the formulas of the form $K_a\phi$:

$$(\mathcal{M}^\tau_X, \mathbf{c}) \models K_a\phi \quad \text{iff} \quad \forall \mathbf{d} \in X \text{ such that } \mathbf{c} \sim^\tau_a \mathbf{d},\ (\mathcal{M}^\tau_X, \mathbf{d}) \models \phi.$$

Fix now a protocol P and a set $X \subseteq \mathbf{C}^\tau$. We define the relativised computation tree of P as the set $\mathbf{C}^\tau_{(P,X)}$ obtained by replacing the above Base and Step conditions by

[Base] $\epsilon \in \mathbf{C}^\tau_{(P,X)}$,

[Step] If $\mathbf{c} \in X \cap \mathbf{C}^\tau_{(P,X)}$ and $(\mathcal{M}^\tau_X, \mathbf{c}) \models \phi$ then $\mathbf{c}.\mathsf{c} \in \mathbf{C}^\tau_{(P,X)}$, where $\phi \to \mathsf{c} \in P$,

and refer to each call sequence from $\mathbf{C}^\tau_{(P,X)}$ as (P, X)-*compliant*.

We now limit the syntax of epistemic guards as follows. A formula $\hat{K}_a\phi$ is an abbreviation for $\neg K_a \neg \phi$ and $\hat{\mathcal{L}}$ denotes the existential fragment of $\mathcal{L}$, consisting of only literals, $\vee$, $\wedge$, and $\hat{K}_a$.

The following lemma clarifies the introduction of the language $\hat{\mathcal{L}}$.

Lemma 6.6. *If $X \subseteq Y \subseteq \mathbf{C}^\tau$ then*

$$\textit{for all formulas } \phi \in \hat{\mathcal{L}} \textit{ and all } \mathbf{c} \in X,\ (\mathcal{M}^\tau_X, \mathbf{c}) \models \phi \implies (\mathcal{M}^\tau_Y, \mathbf{c}) \models \phi.$$

Proof. The only case that requires explanation is when ϕ is of the form $\hat{K}_a\psi$. Suppose that $(\mathcal{M}^\tau_X, \mathbf{c}) \models \phi$. Then for some $\mathbf{d} \in X$ such that $\mathbf{c} \sim^\tau_a \mathbf{d}$, $(\mathcal{M}^\tau_X, \mathbf{d}) \models \psi$. By the induction hypothesis $(\mathcal{M}^\tau_Y, \mathbf{d}) \models \psi$, so by definition $(\mathcal{M}^\tau_Y, \mathbf{c}) \models \phi$. □

Define next an operator $\rho^P : 2^{\mathbf{C}^\tau} \to 2^{\mathbf{C}^\tau}$ by

$$\rho^P(X) = X \cap \mathbf{C}^\tau_{(P,X)}.$$

That is, ρ^P removes from a given set X of call sequences those that are not (P, X)-compliant. What we are after is a set from which no sequences would be removed, so a fixpoint of ρ^P.

Proposition 6.7. *Suppose the epistemic guards of a protocol P are all from $\hat{\mathcal{L}}$. Then there exists an $X \subseteq \mathbf{C}^\tau$ such that $X = \rho^P(X)$.*

[3] We call $\mathbf{C}^\tau_P$ a tree since its elements can be arranged in an obvious way in (a possibly infinite, but finitely branching) tree.

Proof. Suppose that $X \subseteq Y$ and $\mathbf{c} \in X \cap \mathbf{C}^\tau_{(P,X)}$. We prove that $\mathbf{c} \in \mathbf{C}^\tau_{(P,Y)}$ by induction on the length of $\mathbf{c}$. If $\mathbf{c} = \epsilon$, then $\mathbf{c} \in \mathbf{C}^\tau_{(P,Y)}$ by the Base condition.

Otherwise, by the Step condition $\mathbf{c}$ is of the form $\mathbf{c}'.\mathsf{c}$, where $\mathbf{c}' \in X \cap \mathbf{C}^\tau_{(P,X)}$, and for some $\phi \in \hat{\mathcal{L}}$, $(\mathcal{M}^\tau_X, \mathbf{c}') \models \phi$, and $\phi \to \mathsf{c} \in P$. By the induction hypothesis $\mathbf{c}' \in \mathbf{C}^\tau_{(P,Y)}$. Further, $\mathbf{c}' \in Y$ and by Lemma 6.6 $(\mathcal{M}^\tau_Y, \mathbf{c}') \models \phi$, so $\mathbf{c} \in \mathbf{C}^\tau_{(P,Y)}$.

It follows that ρ^P is a monotonic function, that is, $X \subseteq Y$ implies $\rho^P(X) \subseteq \rho^P(Y)$. By the Knaster-Tarski theorem of [32] ρ^P has therefore fixpoints, including a largest and a smallest one. □

Intuitively, when the domain $X \subseteq \mathbf{C}^\tau$ of a gossip model is a fixpoint of ρ^P, then the restriction of the definition of the indistinguishability relations $\sim^\tau_a$ to such a domain has the effect that the call sequences considered possible by the agents coincide with the call sequences generated by the protocol. Such gossip models incorporate then the assumption that there is common knowledge among the agents about the protocol in use.

Furthermore, by the Knaster-Tarski theorem one can construct the largest fixpoint of ρ^P by iteratively applying ρ^P to $\mathbf{C}^\tau$. Such fixpoint $\nu\rho^P$ is the most natural domain for a gossip model that realises the assumption of common knowledge of the protocol, with the $(P, \nu\rho^P)$-compliant call sequences viewed as the P-compliant ones. When the privacy degree is $\circ$ such a gossip model has a very simple structure, namely $(\mathbf{C}^\tau_P, \{\sim^\tau_a\}_{a \in Ag})$.

Corollary 6.8. *Consider a protocol P and a call type τ such that $\tau(\mathsf{p}) = \circ$. Then $\nu\rho^P = \mathbf{C}^\tau_P$.*

Proof. Note that we always have $\rho^P(\mathbf{C}^\tau) = \mathbf{C}^\tau_P$. We now show that $\mathbf{C}^\tau_P \subseteq \mathbf{C}^\tau_{(P,\mathbf{C}^\tau_P)}$ by induction on the length of the call sequences. We only need to consider the induction step. So consider some $\mathbf{c}.\mathsf{c} \in \mathbf{C}^\tau_P$. By definition $\mathbf{c} \in \mathbf{C}^\tau_P$ and $\mathbf{c} \models^\tau \phi$, where $\phi \to \mathsf{c} \in P$, and by the induction hypothesis $\mathbf{c} \in \mathbf{C}^\tau_{(P,\mathbf{C}^\tau_P)}$.

Let ϕ' be obtained from ϕ by removing all occurrences of K_a for all agents a. By Note 6.5 relativised to an arbitrary $X \subseteq \mathbf{C}^\tau$ such that $\mathbf{c} \in X$ we have $\mathbf{c} \models^\tau \phi$ iff $\mathbf{c} \models^\tau \phi'$ iff $(\mathcal{M}^\tau_X, \mathbf{c}) \models \phi'$ iff $(\mathcal{M}^\tau_X, \mathbf{c}) \models \phi$. So in particular $(\mathcal{M}^\tau_{\mathbf{C}^\tau_P}, \mathbf{c}) \models \phi$ and hence by definition $\mathbf{c}.\mathsf{c} \in \mathbf{C}^\tau_{(P,\mathbf{C}^\tau_P)}$.

Consequently $\rho^P(\mathbf{C}^\tau_P) = \mathbf{C}^\tau_P \cap \mathbf{C}^\tau_{(P,\mathbf{C}^\tau_P)} = \mathbf{C}^\tau_P$ and hence $\mathbf{C}^\tau_P$ is the largest fixpoint of ρ^P. □

The syntactic restriction on the epistemic guards used in Proposition 6.7 is clearly satisfied by Protocol 1 as its guards can be rewritten as $\hat{K}_i \neg F_j I$. The same is the case for all protocols studied in [1].

7 Conclusions

We provided an in-depth study of 18 different types of communication relevant for epistemic gossip protocols and modelled their epistemic effects in a uniform way through different indistinguishability relations. This led us to establish a precise map of the relative informativeness of these types of communication (Theorem 5.1). In turn, this result allowed us to prove general results concerning the epistemic effects of call sequences under different communication regimes (Propositions 6.1 and 6.2) and to advance a natural proposal on how to model and analyse agents' common knowledge of gossip protocols (Proposition 6.7), a still under-investigated issue in the literature.

Several natural directions for future research present themselves. We mention three of them. The first question concerns the axiomatisation of the modal language $\mathcal{L}$ introduced in Section 3. This problem is parametrised by the underlying indistinguishability relations introduced in

Section 4. For example, by Note 6.5 the equivalence $\phi \leftrightarrow K_a\phi$ holds for the privacy type $\circ$ but not for the other two.

Actually, even the axiomatization of the F_aS formulas is not straightforward, as it has to take into account the nature of the communication. Indeed, consider the following formula, where $a \neq b$:

$$\Big(F_bA \wedge \bigwedge_{i\neq a,b} \neg F_iA\Big) \rightarrow F_aB.$$

It states that if agent b is the only agent (different from a) familiar with the secret of a, then agent a is familiar with the secret of b. A more general version is:

$$\Big(\bigvee_{i\in X} F_iA \wedge \bigwedge_{i\notin X\cup\{a\}} \neg F_iA\Big) \rightarrow \bigvee_{i\in X} F_aI,$$

where $a \notin X$.

Intuitively it states that if somebody from a group X, to which a does not belong, is familiar with her secret and nobody from outside of the group X (except a) is familiar with this secret, then agent a is familiar with a secret of somebody from the group X. Clearly, both formulas are valid for the $\diamond$ direction type.

In general such an axiomatisation project could be carried out at several levels (cf. [16]): by considering F_iS formulas as primitive, as we did in this paper; or analysing them as "knowing whether" formulas (in epistemic logic notation, $K_iS \vee K_i\neg S$) as in [5]. Whether the latter level of analysis can be easily reconciled with the one proposed in this paper is an interesting open problem.

The second question addresses the problem of decidability of the 18 definitions of truth we introduced. In the terminology of this paper [4] established for the call type $(\bullet, \diamond, \alpha)$ that the semantics and the definition of truth are both decidable for the formulas without nested modalities. It would be interesting to establish analogous results for the remaining call types.

The final question concerns the robustness of our analysis, and specifically of the relationships identified in Theorem 5.1, with respect to modes of gossip that involve the transfer of higher-order epistemic information as introduced and studied in [22, 23]. Intuitively, we would expect this type of higher-order epistemic communication to have an impact on the effects of the asymmetric communication types $\triangleright$ and $\triangleleft$ and for the full privacy $\bullet$ degree.

Finally, one could envisage other aspects of a call not considered in this framework. For example in [5] yet another notion of privacy was considered, according to which given a call ab every agent $c \neq a, b$ noted that at most one call took place. Then for agent c the call sequences ϵ and ab are equivalent but ϵ and ab, ab are not. Another possibility could be to consider a notion of privacy that is intermediate between $\circ$ and $\bullet$, according to which the caller is anonymous but the callee not. Then for agent c the call sequences ab and ad are equivalent but ab and bd are not.

Acknowledgments

We thank Hans van Ditmarsch for most useful and extensive discussions on the subject of this paper. We are also grateful to anonymous referees of this and earlier versions of this paper for helpful comments. The first author was partially supported by the NCN grant nr 2014/13/B/ST6/01807.

References

[1] K. R. Apt, D. Grossi, and W. van der Hoek. Epistemic protocols for distributed gossiping. In *Proceedings Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2015)*, volume 215, pages 51–56. EPTCS, 2016.

[2] K. R. Apt, D. Grossi, and W. van der Hoek. When are two gossips the same? Types of communication in epistemic gossip protocols. Computing Research Repository (CoRR), available at https://arxiv.org/abs/1807.05283, 2018.

[3] K. R. Apt and D. Wojtczak. On the computational complexity of gossip protocols. In *Proceedings of IJCAI 2017*, pages 765–771, 2017.

[4] K. R. Apt and D. Wojtczak. Verification of distributed epistemic gossip protocols. *Journal of Artificial Intelligence Research*, 62:101–132, 2018.

[5] M. Attamah, H. van Ditmarsch, D. Grossi, and W. van der Hoek. Knowledge and gossip. In *Proceedings of ECAI'14*, pages 21–26. IOS Press, 2014.

[6] N. Bailey. *The Mathematical Theory of Epidemics*. Griffen Press, 1957.

[7] B. Baker and R. Shostak. Gossips and telephones. *Discrete Mathematics*, 2:197–193, 1972.

[8] O. Bataineh and R. van der Meyden. Abstraction for epistemic model checking of dining-cryptographers based protocols. In *Proceedings of TARK'11*, 2011.

[9] R. Bumby. A problem with telephones. *SIAM Journal of Algorithms and Discrete Methods*, 2:13–18, 1981.

[10] B. Chlebus and D. Kowalski. Robust gossiping with an application to consensus. *Journal of Computer and System Sciences*, 72:1262–1281, 2006.

[11] M. Cooper, A. Herzig, F. Maffre, F. Maris, and P. Régnier. Simple epistemic planning: Generalised gossiping. In *Proceedings of ECAI 2016*, pages 1563–1564, 2016.

[12] H.P. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*, volume 337 of *Synthese Library*. Springer, 2007.

[13] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. Knowledge-based programs. *Distributed Computing*, 10:199–225, 1997.

[14] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about knowledge*. The MIT Press, Cambridge, 1995.

[15] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Applied Mathematics*, 53:79–133, 1994.

[16] M. Gattinger. *New Directions in Model Checking Dynamic Epistemic Logic*. PhD thesis, ILLC, 2018.

[17] van Ditmarsch H., Grossi D., Herzig A., van der Hoek W., and Kuijer L. Parameters for epistemic gossip problems. In *Proceedings of LOFT'16*, 2016.

[18] A. Hajnal, E. C. Milner, and E. Szemeredi. A cure for the telephone disease. *Canadian Mathematical Bulletin*, 15:447–450, 1972.

[19] J. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[20] J. Halpern and L. Zuck. A little knowledge goes a long way: Knowledge-based derivations and correctness proofs for a family of protocols. *Journal of the ACM*, 39(3):449–478, 1992.

[21] S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18(4):319–349, 1988.

[22] A. Herzig and F. Maffre. How to share knowledge by gossiping. In *Proceedings of EUMAS/AT*, pages 249–263, 2015.

[23] A. Herzig and F. Maffre. How to share knowledge by gossiping. *AI Communications*, 30(1):1–17, 2017.

[24] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. Dissemination of information in intercon-

nection networks (broadcasting and gossiping). In *Combinatorial Network Theory*, pages 125–212. Kluwer, 1996.

[25] J. Hromkovic, R. Klasing, A. Pelc, P. Ruzicka, and W. Unger. *Dissemination of Information in Communication Networks: Broadcasting, Gossiping, Leader Election, and Fault-Tolerance*. Springer, 2005.

[26] S. Knight, B. Maubert, and F. Scharzentruber. Reasoning about knowledge and messages in asynchronous multi-agent systems. *Mathematical Structures in Computer Science*, pages 1–42, 2017.

[27] R. Kurki-Suonio. Towards programming with knowledge expressions. In *Proceedings of POPL'86*, pages 140–149, 1986.

[28] J.-J. Ch. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*, volume 41 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1995.

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

[30] A. Procaccia, Y. Bachrach, and J. Rosenschein. Gossip-based aggregation of trust in decentralized reputation systems. In *Proceedings of IJCAI'07*, pages 1470–1475, 2007.

[31] Á. Seress. Quick gossiping without duplicate transmissions. *Graphs and Combinatorics*, 2:363–383, 1986.

[32] Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5:285–309, 1955.

[33] R. Tijdeman. On a telephone problem. *Nieuw Archief voor Wiskunde*, 3(XIX):188–192, 1971.

[34] H. van Ditmarsch, J. van Eijck, P. Pardo, R. Ramezanian, and F. Scharzentruber. Epistemic protocols for dynamic gossip. *Journal of Applied Logic*, 20:1–31, 2017.