When Are Two Gossips the Same?

Types of Communication in Epistemic Gossip Protocols

Paper #XXX

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 different indistinguishability relations. We then study these relations establishing results concerning the relative informativeness of the different types of communication identified.

KEYWORDS

Epistemic logic; gossip protocols

1 INTRODUCTION

In the gossip problem [5, 24] a number n 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 [5, 6, 13, 23, 24], 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 in 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 [7] and probabilistic information transmission, where the spreading of gossips is used to model the spread of an epidemics [4, 22]. Surveys are [11, 15, 17, 18].

Scientific context: Epistemic Gossip Protocols. The present paper proposes 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 [1, 3]. '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 distributed systems [9, 14, 19, 21].

Besides the aforementioned [1, 3], a number of papers have recently focused on epistemic gossip protocols. In [16] 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 [8] gossip protocols were studied as an instance of multi-agent epistemic planning that is subsequently translated into the classical planning language PDDL. More recently, [2] studied the computational complexity of epistemic gossip protocols and [25] 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 links (that is, which other agents can be called).

Paper Contribution. All these works often make different assumptions on the nature of communication upon which the protocol is based. Little work aiming at a systematic analysis has been attempted, with the notable exception of [12], 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 would 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. The comparison of the resulting definitions of knowledge is also of obvious interest from the general point of view of the study of epistemic aspects of communication. From an epistemic logic point of view, each call type induces a specific notion of knowledge.

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 whether the agents exchange all or some information they hold. 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. The number of possibilities is surprisingly large: 18. The paper provides a unified framework in which we model, systematically analyze and compare these possibilities.

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. Andre, S. Koenig (eds.), July 2018, Stockholm, Sweden

https://doi.org/doi

Paper outline. Section 2 introduces gossip protocols by example. Section 3 identifies the features of calls we will focus on, introduces a simple epistemic language to study them, and provides some motivating examples. Section 4 provides a uniform formal semantics of the epistemic language parametrised by the indistinguishability relations which for each type identify the call sequences that the agents cannot distinguish. These relations are systematically introduced and explained. They are studied and compared in terms of their relative informativeness in Section 5. Section 6 sketches how such relations can also be used to capture the notion of common knowledge of a gossip protocol. Section 7 concludes.

2 GOSSIP PROTOCOLS

Gossip protocols aim at sharing knowledge between agents in a pre-described way. This is the paradigmatic setup:

> *Six friends each know a secret. They can call each other by phone. In each call they exchange all the secrets they know. How many calls are needed for everyone to know all secrets?*

Let us generalise this to the case of $n \geq 2$ agents and focus on protocols that are *sufficient* (in the sense that they spread all secrets). If $n = 2$, the two agents a and b need to make only one phone call, which we denote by ab ('a calls b'). For $n = 3$, the call sequence ab, bc, ca will do. Let us look at a protocol for $n \geq 4$ agents.

Protocol 1. *Choose four from the set of agents Ag, say a, b, c, d, and one of those four, say a. First, a makes $n-4$ calls to each agent in $Ag \setminus \{a, b, c, d\}$. Then, the calls ab, cd, ac, bd are made. Finally a makes another call to each agent from $Ag \setminus \{a, b, c, d\}$.*

This adds up to $(n-4)+4+(n-4) = 2n-4$ calls. For $n = 6$ we get a call sequence $ae, af, ab, cd, ac, bd, ae, af$ of 8 calls. All agents are then familiar with all secrets. Less than $2n-4$ calls is insufficient to distribute all secrets [24].

The above protocol assumes that the agents can coordinate their actions before making the calls. But often such coordination is not possible. Suppose all students of the cohort the above friends are part of, receive an unexpected invitation for a party. The friends may be curious to find out about each other whether they will accept, in which case they will have to make phone calls based on the knowledge, or better, ignorance, they have about the secrets of others. Since in such a distributed protocol several agents may decide to initiate a call at the same time, we assume the presence of an *arbiter* who breaks the ties in such cases. Let us now consider such an epistemic protocol: an agent calls another agent depending on her knowledge (or ignorance) only, and choices are random.

Protocol 2 (Hear my secret). *Any agent a calls agent b if a does not know whether b is familiar with a's secret.*

This protocol has been proven to terminate and be correct—in the sense that upon termination everybody is familiar with all secrets—in [1], under specific assumptions on the type of communication taking place during each call. In this paper we aim at providing a systematic presentation of what such assumptions can reasonably be, and at an analysis of their logical interdependencies.

3 KNOWLEDGE IN THE GOSSIP PROBLEM

Throughout the paper we assume a fixed finite set Ag of at least three ***agents***. We 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.

3.1 Calls

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 aspects 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:

- ○: every agent $c \neq a, b$ notes that a calls b,
- ◓: every agent $c \neq a, b$ notes that some call takes place, though not between whom,
- ●: no agent $c \neq a, b$ notes that a call is taking place.

Intuitively, these degrees can be ordered as $\circ <_p \text{◓} <_p \bullet$, with ○ meaning no privacy at all, ◓ ensuring anonymity of the caller and callee, and ● denoting full privacy. Conversely, from the perspective of the agents not involved in the call, a call with the privacy level ○ is the most informative, while a call with the privacy level ● is the most opaque. We will elaborate on this notion in Examples 3.2–3.3.

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

- ***push***, written as ▷. As a result of the call the callee learns all the secrets held by the caller.
- ***pull***, written as ◁. As a result of the call the caller learns all the secrets held by the callee.
- ***push-pull***, written as ◇. 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. For a call of direction ◇ these agents are a and b, for a call of direction ▷ this is b, and for a call of direction ◁ this is a. This distinction allows us to consider two possible levels of ***observance*** of a call:

- α: During the call the affected agent(s) incorporate the secrets of their partner with their own secrets, and only *a*fter that, inspect the result.[1]
- β: During the call the affected agent(s) inspect the secrets of their partner *b*efore 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 affected agent before adding them to the secrets she is familiar with. Let

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

Each call is of the shape ab^{τ}, where $\tau = (\mathrm{p}, \mathrm{d}, \mathrm{o}) \in \mathrm{P} \times \mathrm{D} \times \mathrm{O}$ is called its ***type***. So we defined in total 18 call types.

The types $(\circ, \diamond, \beta)$ and $(\ominus, \diamond, \beta)$ were studied in [3] 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 $\tau(\mathrm{p}) = \circ, \tau(\mathrm{d}) = \diamond$ and $\tau(\mathrm{o}) = \beta$.

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 or (a, b) for the direction type $\diamond$, $a \triangleright b$ for the direction type $\triangleright$ and $a \triangleleft b$ for the direction type $\triangleleft$.

3.2 Modal language

We are interested in determining agents' knowledge after a sequence of calls took place. To this end we use the standard modal language $\mathcal{L}$ for epistemic logic (see [1]):

$$\phi \quad ::= \quad F_a S \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid K_a\phi \mid \hat{K}_a\phi,$$

where $a \in Ag$ and $S \in \mathrm{S}$. We will be using also fragments of $\mathcal{L}$: the *universal fragment* $\mathcal{L}^+$, consisting of the *literals* $F_a S$ and $\neg F_a S$, $\wedge$ and K_a; and the *existential fragment* $\hat{\mathcal{L}}$, consisting of only literals, $\vee$ and $\hat{K}_a$.

In what follows we refer to the elements ϕ of $\mathcal{L}$ as ***epistemic formulas***, or in short, just formulas. We read $F_a B$ as 'agent a is familiar with the secret B' (or 'B belongs to the set of secrets a has learned') and $K_a\phi$ as 'agent a knows that formula ϕ is true'. Formula $\hat{K}_a\phi$ states that agent a considers possible that ϕ is true. So $\mathcal{L}$ is an epistemic language in which atomic formulas consist of 'being familiar with' statements about secrets.

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").

Consider now the statement that initially each agent is familiar only with her own secret. This can be expressed as the formula

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

[1]This mode is akin to the caller and callee interacting through a third party, who first collects the caller's and callee's secrets separately, and then shares their union with both. We are indebted to [omitted for blind reviewing] for this observation.

Finally, consider the statement that it is not the case that some agent, say a, is familiar with all the secrets and each other agent is familiar with at most her own secret and that of a. This can be expressed as the formula

$$\bigwedge_{a \in Ag} \neg(Exp_a \wedge \bigwedge_{b \in Ag, b \neq a, C \in \mathrm{S}, C \neq A, B} \neg F_b C).$$

□

We now 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 4.

Example 3.2. Suppose there are four agents, a, b, c and d. Consider the call type is $(\circ, \diamond, \alpha)$. Assume the call sequence ab, bc.

Let us reason from the perspective of agent d. Because of the assumed privacy level after the first call, ab, agent d knows that both agents a and b are familiar with A and B. This can be expressed as the formula

$$K_d(F_a A \wedge F_a B \wedge F_b A \wedge F_b B).$$

This then implies that after the second call, bc, agent d also knows that both b and c are familiar with A, B, and C. Her factual knowledge can be expressed as the formula $K_d\phi$, where

$$\phi = F_a A \wedge F_a B \wedge F_b A \wedge F_b B \wedge F_b C \wedge F_c A \wedge F_c B \wedge F_c C.$$

In fact, because of the assumed privacy level $\circ$, how the knowledge evolves during communication is completely transparent to all agents. Hence after both calls everybody knows ϕ, i.e.,

$$\bigwedge_{a \in Ag} K_a\phi.$$

An analogous argument applies for the call type $(\circ, \diamond, \beta)$.

Suppose now that the privacy level is $\ominus$. Then we cannot conclude the formula $K_d\phi$, since agent d only knows then that two calls took place, but not between which pairs of agents. In fact, in this case we can only conclude (note that the same call can be made twice):

$$K_d(\bigvee_{a, b \in Ag \setminus \{d\}, a \neq b} F_a B).$$

Finally, if the privacy level is $\bullet$, then d is not aware of calls ab and bc. She considers it possible that a, b, c are already familiar with all secrets except her own, but also considers it possible that all other agents only know their own secret. As she has not yet been involved in a call, she knows that they are not familiar with D. □

Example 3.3. Suppose there are three agents, a, b and c. Consider the two call types $(\bullet, \diamond, \mathrm{o})$, where $\mathrm{o} \in \mathrm{O}$, 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.

In both cases agent c (who also is an expert) does not know that agents a and b communicated, so she does not know that they are experts. In other words, the formula $K_c\phi$ is not true. This changes when the privacy degree is ○, i.e., in that case the formula $K_c\phi$ is true. Moreover, because there are three agents, the same conclusion holds when the privacy degree is ◓. However, the last conclusion does not hold anymore when there are more than three agents. □

Example 3.4. Assume the same call sequence as in the previous case but suppose that the call parameters are now (○, ◃, o), where o ∈ O. So we consider now the call sequence $\mathbf{c} = a \triangleleft c, b \triangleleft c, a \triangleleft b$.

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 $\mathbf{c}$

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

If the privacy degree is ◓ or ● we only have $K_a F_b C$ if the observance is β: agent a then cannot distinguish $\mathbf{c}$ from $a \triangleleft c, c \triangleleft b, a \triangleleft b$. Clearly, $K_c F_a B$ does not hold then after $\mathbf{c}$. □

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.

4 SEMANTICS

We provide now a formal semantics for the modal language $\mathcal{L}$.

4.1 Gossip situations and calls

First we recall the following crucial notions introduced in [1]. A ***gossip situation*** is a sequence $\mathrm{s} = (\mathrm{Q}_a)_{a\in Ag}$, where $\mathrm{Q}_a \subseteq \mathrm{S}$ 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 $\mathrm{s} = (\mathrm{Q}_a)_{a\in Ag}$, 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 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 4.1. The application of a call c to a gossip situation s is defined as follows, where $\mathrm{s} := (\mathrm{Q}_a)_{a\in Ag}$:

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

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

$\boxed{\mathrm{c} = a \triangleleft b}$ $\mathrm{c}(\mathrm{s}) = (Q'_a)_{a\in Ag}$, where $\mathrm{Q}'_a = \mathrm{Q}_a \cup \mathrm{Q}_b$, $\mathrm{Q}'_b = \mathrm{Q}_b$, $\mathrm{Q}'_c = \mathrm{Q}_c$, 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 $(a \diamond b)(\mathrm{s}) = (b \diamond a)(\mathrm{s})$ and $(a \triangleright b)(\mathrm{s}) = (b \triangleleft a)(\mathrm{s})$, as expected.

In turn, the privacy degree of a call captures what outsiders of the call learn from it and the observance level informally determines 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 the call sequences.

A ***call sequence*** is a *finite* sequence of calls, all of the same 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 type τ. Given the call sequence $\mathbf{c}$ and a call c, $\mathbf{c}$.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 4.1, as follows:

$$[\text{Base}]\ \epsilon(\mathrm{s}) := \mathrm{s};\ [\text{Step}]\ \mathbf{c}.\mathrm{c}(\mathrm{s}) := \mathrm{c}(\mathbf{c}(\mathrm{s})).$$

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

Example 4.2. Let Ag be $\{a, b, c\}$. We use the following concise notation for gossip situations. Sets of secrets will be written down as lists. E.g., the set $\{A, B, C\}$ will be written as ABC. Gossip situations will be written down as lists of lists of secrets separated by dots. E.g., $\mathrm{i} = A.B.C$ and the gossip situation $(\{A, B\}, \{A, B\}, \{C\})$ will be written as $AB.AB.C$. So, $(ab, ca, ab)(A.B.C) = ABC.ABC.ABC$. □

4.2 Truth of formulas

We illustrated in Examples 3.2–3.4 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 a basis of the approach to epistemic gossip protocols put forth in [1], and upon which we build here.

Consider for instance the situation analyzed in Example 3.3. We noticed there that depending on the assumed observance level the knowledge of agent a differs. This has to do with the call sequences the agent considers possible. If the call type is $(\bullet, \diamond, \alpha)$ agent a cannot distinguish between the call sequences ac, ab and ac, bc, ab. Indeed, after both sequences she is familiar with all the secrets but she cannot determine whether agents b and c communicated. From his perspective both call sequences are possible, that is, he cannot distinguish between them. In contrast, if the call type is $(\bullet, \diamond, \beta)$ agent a can distinguish between these two call sequences, which has in turn an effect on her knowledge.

So 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 relation $\mathbf{c} \sim_a^\tau \mathbf{d}$ informally means 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 Ag) $\mathcal{M}^\tau = (\mathbf{C}^\tau, \{\sim_a^\tau\}_{a\in Ag})$ and a call sequence $\mathbf{c}$ as follows:

Definition 4.3. Let $\mathcal{M}^\tau$ be a gossip model for call type τ and 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 and the case of $\hat{K}_a\phi$, which is the dual of the case for $K_a\phi$, omitted):

$$\mathcal{M}^\tau, \mathbf{c} \models F_aS \quad \text{iff} \quad S \in \mathbf{c}(\mathrm{i})_a,$$
$$\mathcal{M}^\tau, \mathbf{c} \models K_a\phi \quad \text{iff} \quad \forall \mathbf{d} \text{ such that } \mathbf{c} \sim_a^\tau \mathbf{d},\ \mathcal{M}^\tau, \mathbf{d} \models \phi.$$

When the gossip model is clear from the context, we will sometimes 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_aS 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 customary in epistemic logic [10, 20] using the equivalence relations $\sim_a^\tau$.

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 type of the calls are needed, through the definition of the $\sim_a^\tau$ relations.

4.3 Indistinguishability of call sequences

Below we use two intuitive notions. We say that an agent a is ***involved*** in a call c, and write $a \in$ 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. We say that an agent a is ***affected*** by a call c if c is one of the following forms:

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

Intuitively, a is affected by the call c if it can affect the set of secrets a is familiar with. So agent a is involved but not affected by the call c if $\mathrm{c} = a \triangleright b$ or $\mathrm{c} = b \triangleleft a$.

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 $\precsim_a^\tau$ (Definition 4.4). Intuitively, the expression $\mathbf{c} \precsim_a^\tau \mathbf{d}$ can be interpreted as "from the point of view of a, if $\mathbf{c}$ is an (epistemically) possible call sequence, so is $\mathbf{d}$". Then, we define $\sim_a^\tau$ as the least equivalence relation that contains $\precsim_a^\tau$, that is:

$$\sim_a^\tau = (\precsim_a^\tau \cup (\precsim_a^\tau)^{-1})^*,$$

where $^{-1}$ is the inverse operation on binary relations and * is the transitive, reflexive closure operation on binary relations.

Definition 4.4. Let $a \in Ag$ and fix a type τ. Relation $\precsim_a^\tau$ is the smallest sub-relation of $\mathbf{C}^\tau \times \mathbf{C}^\tau$ satisfying the following conditions, where $b, c \in Ag$:

[Base] $\epsilon \precsim_a^\tau \epsilon$.
[Closure] If $\tau(\mathrm{d}) = \diamond$ and $\mathbf{c}.b\diamond c \precsim_a^\tau \mathbf{d}$, then $\mathbf{c}.c\diamond b \precsim_a^\tau \mathbf{d}$.
[Step] Suppose that $\mathbf{c} \precsim_a^\tau \mathbf{d}$ and let c and d be calls.

Step-out$^\tau$ if $Out_a^\tau(\mathrm{c}, \mathrm{d})$ then $Concl_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c}, \mathrm{d})$,
Step-in$^\tau$ if $In_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$ then $Concl_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$,

where the used relations are defined in Table 1.

To reflect on the definition of $\sim_a^\tau$, it 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. The Closure condition captures the fact that from the point of view of agent a the calls $b\diamond c$ and $c\diamond b$ are the same (with $a = b$ or $a = c$ allowed). By symmetry we can also conclude that $\mathbf{c} \sim_a^\tau \mathbf{d}.b\diamond c$ implies $\mathbf{c} \sim_a^\tau \mathbf{d}.c\diamond b$.

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

Consider now the privacy type ◓. The conditions $a \notin$ c and $a \notin$ 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 $\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{d}$. Note that this conclusion is not justified for the privacy type ○ because if $\mathrm{c} \neq \mathrm{d}$ then agent a can distinguish between these two calls, so a fortiori between the call sequences $\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{d}$.

Finally, consider the privacy type ●. According to its informal description, the condition $a \notin$ c means that agent a is not aware of the call c. This justifies the conclusion $\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}$. (The conclusion $\mathbf{c} \sim_a^\tau \mathbf{d}.\mathrm{c}$ follows by symmetry.)

Next, consider the Step-in$^\tau$ clause. It spells the conditions that allow one to extend the indistinguishability 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 two

Agent a is not involved in the last call:

τ	$Out_a^\tau(\mathrm{c}, \mathrm{d})$	$Concl_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c}, \mathrm{d})$
$\tau(\mathrm{p}) = ○$	$a \notin \mathrm{c}$	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{c}$
$\tau(\mathrm{p}) = ◓$	$a \notin \mathrm{c}, a \notin \mathrm{d}$	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{d}$
$\tau(\mathrm{p}) = ●$	$a \notin \mathrm{c}$	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}$

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

$In_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$	$Concl_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$
$a \in \mathrm{c}$, a not affected by c	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{c}$

Agent a is involved in and affected by the last call, and agent b is the partner of a in the call:

τ	$In_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$	$Concl_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$
$\tau(\mathrm{o}) = \alpha$	$a \in \mathrm{c}$, a affected by c, $\mathbf{c}.\mathrm{c}(\mathrm{i})_a = \mathbf{d}.\mathrm{c}(\mathrm{i})_a$	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{c}$
$\tau(\mathrm{o}) = \beta$	$a \in \mathrm{c}$, a affected by c, $\mathbf{c}(\mathrm{i})_b = \mathbf{d}(\mathrm{i})_b$	$\mathbf{c}.\mathrm{c} \precsim_a^\tau \mathbf{d}.\mathrm{c}$

Table 1: Defining indistinguishability of call sequences

call sequences that are indistinguishable for a can be both extended by 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 agent a cannot distinguish between the resulting call sequences $\mathbf{c}$.c and $\mathbf{d}$.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 check these sets of secrets after, respectively, the call sequences $\mathbf{c}$ and $\mathbf{d}$ took place. If these sets are equal, then agent a cannot distinguish between the resulting call sequences $\mathbf{c}$.c and $\mathbf{d}$.c. This explains why in this case a reference to agent b is made in $In_a^\tau(\mathbf{c}, \mathbf{d}, \mathrm{c})$.

The following simple observation is useful.

Observation 4.5. *For all types τ if $\mathbf{c} \sim_a^\tau \mathbf{d}$, then $\mathbf{c}(\mathrm{i})_a = \mathbf{d}(\mathrm{i})_a$.*

Example 4.6. We first illustrate Table 1, top, by analyzing situations in which the considered agent is not involved in the last call. Assume four agents, a, b, c and d.

Suppose that the privacy of τ is $\circ$. We have $ab, bc \nsim_a^\tau ab, cd$, because $ab, bc \not\preceq_a^\tau ab, cd$ as $bc \neq cd$ and $bc \neq dc$. So we fail to apply Table 1, top, first row and the symmetric closure does not give us that either.

On the other hand, if the privacy of τ is $\bullet\!\!\!-$, we have $ab, bc \sim_a^\tau ab, cd$, because $ab, bc \preceq_a^\tau ab, cd$, as $a \notin bc$, $a \notin cd$ and $ab \sim_a^\tau ab$ (Table 1, top, second row), and $\sim_a^\tau$ is a reflexive closure. On the other hand, $ab, bc \nsim_a^\tau ab, cd, bc$ as now the clause in the second row fails to apply, as the lengths of the compared sequences are different.

Finally, if the privacy of τ is $\bullet$, we of course also have $ab, bc \sim_a^\tau ab, cd$ for the same reason as in the previous paragraph, but we now also have $ab, bc \sim_a^\tau ab, cd, bc$, because $ab, bc \preceq_a^\tau ab, cd, bc$. Indeed, we have $ab \sim_a^\tau ab$ and hence by Table 1, top, third row, applied three times (with symmetric closure), first $ab \sim_a^\tau ab, cd$, then $ab, bc \sim_a^\tau ab, cd$, and finally $ab, bc \sim_a^\tau ab, cd, bc$. □

Example 4.7. To illustrate Table 1, middle, consider the same four agents and sequence $d \triangleright c, b \triangleright c$, and $\bullet\!\!\!-$. Then $d \triangleright c, b \triangleright c \sim_b^\tau c \triangleright d, b \triangleright c$, because (skipping closure details of $\preceq_b$) agent b is involved in the second call but not affected (Table 1, middle), and $d \triangleright c \sim_b^\tau c \triangleright d$, because $b \notin d \triangleright c$ and $b \notin c \triangleright d$ (Table 1, top, second row). □

Example 4.8. Now consider Table 1, bottom. The difference between observancies α and β is seen in Example 3.4. For observancy α ('after') we have that $a \triangleleft c, b \triangleleft c, a \triangleleft b \sim_a^\tau a \triangleleft c, c \triangleleft b, a \triangleleft b$, because agent a is afterwards familiar with the same set of secrets on the left and on the right, namely $\{A, B, C\}$ (Table 1, bottom, first row). On the other hand, for observancy β ('before') we get $a \triangleleft c, b \triangleleft c, a \triangleleft b \nsim_a^\tau a \triangleleft c, c \triangleleft b, a \triangleleft b$, because $a \triangleleft c, b \triangleleft c \nsim_b^\tau a \triangleleft c, c \triangleleft b$ (note that this concerns indistinguishability for agent b, not a); here the second row of Table 1, bottom, applies.

As a final example, we have that $d \triangleright c, b \triangleright c \nsim_c^\tau c \triangleright d, b \triangleright c$, because unlike agent b who is involved and not affected, agent c is involved and affected in the second call $b \triangleright c$. Observe that after $d \triangleright c, b \triangleright c$ agent c is familiar with secrets B, C, D, whereas after $c \triangleright d, b \triangleright c$ agent c is only familar with B, C. Now, apply Observation 4.5. □

5 COMPARISON OF THE $\sim_a^\tau$ RELATIONS

We defined in total 18 $\sim_a^\tau$ equivalence relations. We now study how they compare to one another.

5.1 No privacy

We start with the simplest case concerning the privacy degree $\circ$, and begin with the following observation, where we use the relation on call sequences $\approx$ defined as:

> $\mathbf{c} \approx \mathbf{d}$ if for some k, $\mathbf{c} = \mathrm{c}_1, \ldots, \mathrm{c}_k$, $\mathbf{d} = \mathrm{d}_1, \ldots, \mathrm{d}_k$ and for each $i \in \{1, \ldots, k\}$ either $\mathrm{c}_i = \mathrm{d}_i$ or for some agents a, b, $\mathrm{c}_i = a \diamond b$ and $\mathrm{d}_i = b \diamond a$.

Intuitively $\approx$ weakens the equality relation between call sequences by coalescing the two possible instances of push-pull calls.

Theorem 5.1. *Suppose that $\tau(\mathrm{p}) = \circ$. Fix an agent a.*

(i) If $\tau(\mathrm{d}) \neq \diamond$ then $\sim_a^\tau$ is the identity relation.
(ii) If $\tau(\mathrm{d}) = \diamond$ then $\sim_a^\tau$ is the $\approx$ relation.

Proof. (*i*) Each $\sim_a^\tau$ is an equivalence relation, so by its definition it suffices to prove that $\mathbf{c} \preceq_a^\tau \mathbf{d}$ implies $\mathbf{c} = \mathbf{d}$. 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 $|\mathbf{c}'.\mathrm{c}| + |\mathbf{d}| > k$ and $\mathbf{c}'.\mathrm{c} \preceq_a^\tau \mathbf{d}$ (induction hypothesis). By definition $\preceq_a^\tau$ is the smallest relation satisfying the Base, Closure and Step conditions of Definition 4.4. So, by the Step condition, $\mathbf{d}$ is of the form $\mathbf{d}'.\mathrm{d}$ where $\mathbf{c}' \preceq_a^\tau \mathbf{d}'$ and $\mathrm{c} = \mathrm{d}$. By induction hypothesis $\mathbf{c}' = \mathbf{d}'$, so $\mathbf{c} = \mathbf{d}$.

(*ii*) We apply the same inductive argument, and use in addition the Closure condition of Definition 4.4 to conclude that $\mathbf{c} \preceq_a^\tau \mathbf{d}$ implies $\mathbf{c} \approx \mathbf{d}$. □

Before moving further, let us introduce some auxiliary notation. Given two call types τ_1 and τ_2 we abbreviate the statement $\forall a \in Ag, {\sim_a^{\tau_1}} \subseteq {\sim_a^{\tau_2}}$ to $\tau_1 \subseteq \tau_2$ and similarly for the equality (written as $\equiv$) and the strict inclusion (written as $\subset$) between these equivalence relations. Such statements presuppose that we systematically change the types of all calls in a given call sequence. Further, the unspecified parameters are implicitly universally qualified. For example, $(\circ, \mathrm{d}, \mathrm{o}) \equiv (\bullet\!\!\!-, \mathrm{d}, \mathrm{o})$ is an abbreviation for the statement

$$\forall a \in Ag\ \forall \mathrm{d} \in \mathrm{D}\ \forall \mathrm{o} \in \mathrm{O}\ {\sim_a^{(\circ, \mathrm{d}, \mathrm{o})}} = {\sim_a^{(\bullet\!\!\!-, \mathrm{d}, \mathrm{o})}}\ .$$

This notation and the way of modifying call sequences is used in the following consequence of Theorem 5.1.

Theorem 5.2. *Suppose that $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p}) = \circ$.*

(i) If $\tau_1(\mathrm{d}) \neq \diamond$ and $\tau_2(\mathrm{d}) \neq \diamond$ then $\tau_1 \equiv \tau_2$.
(ii) If $\tau_1(\mathrm{d}) = \tau_2(\mathrm{d}) = \diamond$ then $\tau_1 \equiv \tau_2$.
(iii) If $\tau_1(\mathrm{d}) \neq \diamond$ and $\tau_2(\mathrm{d}) = \diamond$ then $\tau_1 \subset \tau_2$.

Proof. (*i*) and (*ii*) are direct consequences of Theorem 5.1(*i*).

(*iii*) follows from the fact that $= \subseteq \approx$. To appreciate that the inclusion is strict, note that $a \diamond b \sim_a^{\tau_2} b \diamond a$, while $a \triangleright b \nsim_a^{\tau_1} b \triangleright a$ and $a \triangleleft b \nsim_a^{\tau_1} b \triangleleft a$. □

5.2 Other privacy degrees

The above situation changes for the other privacy degrees.

Theorem 5.3. *Assume $|Ag| \geq 4$. Suppose that $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p}) = \textcircled{\scriptsize}$ and $\tau_1(\mathrm{d}) \neq \tau_2(\mathrm{d})$. Then the relations $\sim_a^{\tau_1}$ and $\sim_a^{\tau_2}$ are incomparable.*

Proof. For each pair of distinct direction types we exhibit appropriate call sequences. In each case the conclusions do not depend on the observance level.

(1) Suppose that $\tau_1(\mathrm{d}) = \triangleleft$ and $\tau_1(\mathrm{d}) = \diamond$.
Then $b \triangleleft c, c \triangleleft a \sim_a^{\tau_1} b \triangleleft d, c \triangleleft a$, while $b \diamond c, c \diamond a \not\sim_a^{\tau_2} b \diamond d, c \diamond a$.
Further, $b \diamond c, a \diamond c \sim_a^{\tau_2} c \diamond b, a \diamond c$, while $b \triangleleft c, a \triangleleft c \not\sim_a^{\tau_1} c \triangleleft b, a \triangleleft c$.
(2) Suppose that $\tau_1(\mathrm{d}) = \triangleright$ and $\tau_1(\mathrm{d}) = \diamond$.
Then $c \triangleright b, a \triangleright c \sim_a^{\tau_1} d \triangleright b, a \triangleright c$, while $c \diamond b, a \diamond c \not\sim_a^{\tau_2} d \diamond b, a \diamond c$.
Further, $c \diamond b, c \diamond a \sim_a^{\tau_2} b \diamond c, c \diamond a$, while $c \triangleright b, c \triangleright a \not\sim_a^{\tau_1} b \triangleright c, c \triangleright a$.
(3) Suppose that $\tau_1(\mathrm{d}) = \triangleright$ and $\tau_1(\mathrm{d}) = \triangleleft$.
Then $b \triangleright c, a \triangleright c \sim_a^{\tau_1} c \triangleright b, a \triangleright c$, while $b \triangleleft c, a \triangleleft c \not\sim_a^{\tau_2} c \triangleleft b, a \triangleleft c$.
Further, $c \triangleleft b, c \triangleleft a \sim_a^{\tau_2} b \triangleleft c, c \triangleleft a$, while $c \triangleright b, c \triangleright a \not\sim_a^{\tau_1} b \triangleright c, c \triangleright a$.

The assumption $|Ag| \geq 4$ is necessary as, for $|Ag| \leq 3$ and $\mathrm{p} = \textcircled{\scriptsize}$, even though an agent cannot observe a call, she knows between which two agents it took place. □

In view of the above results we will focus on comparing the $\sim_a^{\tau}$ relations with the same direction type. We shall discuss the case of three agents below. For the privacy degree ● the assumption about the number of agents can be dropped.

Theorem 5.4. *Assume $|Ag| \geq 3$. Suppose that $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p}) = \bullet$ and $\tau_1(\mathrm{d}) \neq \tau_2(\mathrm{d})$. Then the relations $\sim_a^{\tau_1}$ and $\sim_a^{\tau_2}$ are incomparable.*

Proof. The proof of Theorem 5.3 remains valid and the assumption that there are at least four agents is now not needed. In (1) in the proof of Theorem 5.3 one can now remove the calls $b \triangleleft d$ and $b \diamond d$. □

Finally, we have the following result.

Theorem 5.5.

(i) *If $|Ag| = 3$ then $(\circ, \diamond, \mathrm{o}) \equiv (\textcircled{\scriptsize}, \diamond, \mathrm{o})$.*
(ii) *If $|Ag| \geq 4$ or $\mathrm{d} \neq \diamond$ then $(\circ, \mathrm{d}, \mathrm{o}) \subset (\textcircled{\scriptsize}, \mathrm{d}, \mathrm{o})$.*
(iii) $(\textcircled{\scriptsize}, \mathrm{d}, \mathrm{o}) \subset (\bullet, \mathrm{d}, \mathrm{o})$.
(iv) $(\circ, \mathrm{d}, \beta) \equiv (\circ, \mathrm{d}, \alpha)$.
(v) $(\textcircled{\scriptsize}, \mathrm{d}, \beta) \subset (\textcircled{\scriptsize}, \mathrm{d}, \alpha)$.
(vi) $(\bullet, \mathrm{d}, \beta) \subset (\bullet, \mathrm{d}, \alpha)$.

Sketch of proof. All the inclusions $\subseteq$ are direct consequences of Definition 4.4.
(i) To prove the inclusion $\supseteq$ suppose $Ag = \{a, b, c\}$. Then, if a is not involved in a call, the call has to be between b and c. So the claim follows from Definition 4.4.
(ii) To show that the inclusion is strict assume first that $Ag = \{a, b, c\}$. Suppose that for some $\mathrm{o} \in \mathrm{O}$, $\tau_1 = (\circ, \triangleright, \mathrm{o})$ and $\tau_2 = (\textcircled{\scriptsize}, \triangleright, \mathrm{o})$. Then $b \triangleright c \sim_a^{\tau_2} c \triangleright b$ while $b \triangleright c \not\sim_a^{\tau_1} c \triangleright b$. A similar argument holds for the $\triangleleft$ direction. Next, assume four agents a, b, c, d. Suppose that for some $\mathrm{o} \in \mathrm{O}$, $\tau_1 = (\circ, \diamond, \mathrm{o})$ and $\tau_2 = (\textcircled{\scriptsize}, \diamond, \mathrm{o})$. Then $bc \sim_a^{\tau_2} bd$, while $bc \not\sim_a^{\tau_1} bd$. A similar argument holds for the other direction type.
(iii) To see that the inclusion is strict notice that, by Definition 4.4, for the privacy degree ⓞ only call sequences of the same length can be indistinguishable, while this fails to be the case for the privacy degree ●.
(iv) This is a direct consequence of Theorem 5.1.
(v) To see that the inclusion is strict consider four agents a, b, c, d. Suppose that for some $\mathrm{d} \in \{\diamond, \triangleleft\}$ we have $\tau_1 = (\textcircled{\scriptsize}, \mathrm{d}, \alpha)$ and $\tau_2 = (\textcircled{\scriptsize}, \mathrm{d}, \beta)$. Then $ab, ac, bc, ab \sim_a^{\tau_1} ab, ac, cd, ab$, while $ab, ac, bc, ab \not\sim_a^{\tau_2} ab, ac, cd, ab$. For $\mathrm{d} = \triangleright$, a similar example can be constructed.
(vi) To see that the inclusion is strict consider three agents a, b, c. Suppose that for some $\mathrm{o} \in \mathrm{O}$, $\tau_1 = (\bullet, \diamond, \alpha)$ and $\tau_2 = (\bullet, \diamond, \beta)$. Then $ac, ab \sim_a^{\tau_1} ac, bc, ab$, while $ac, ab \not\sim_a^{\tau_2} ac, bc, ab$. A similar argument holds for the other two direction types. □

Note that in *(i)* the restriction to $\diamond$ is necessary. Indeed, for $\tau_1 = (\textcircled{\scriptsize}, \triangleright, \mathrm{o})$ and $\tau_2 = (\circ, \triangleright, \mathrm{o})$ we have $a \triangleright b \sim_c^{\tau_1} b \triangleright a$, while $a \triangleright b \not\sim_c^{\tau_2} b \triangleright a$, and similarly for $\triangleleft$.

5.3 Epistemic effects of communication types

The above systematization is useful to draw general epistemic consequences from different communication types. The following Proposition explains why such results are of interest.

Proposition 5.6. *Consider two call types τ_1 and τ_2 such that $\tau_1(\mathrm{d}) = \tau_2(\mathrm{d})$ and $\tau_1 \subseteq \tau_2$. Then*

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

Proof. We proceed by induction on the structure of ϕ. If ϕ is atomic, say $F_c D$ the conclusion directly follows from the assumption that $\tau_1(\mathrm{d}) = \tau_2(\mathrm{d})$.

For the induction step only the case when ϕ is of the form $K_a\psi$ requires explanation. 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 $\mathbf{c} \models^{\tau_1} K_a\psi$. □

By Proposition 5.6 if $\tau_1(\mathrm{d}) = \tau_2(\mathrm{d})$ and $\tau_1 \subseteq \tau_2$ then for all call sequences $\mathbf{c}$, $\mathbf{c} \models^{\tau_2} K_a F_b C$ implies $\mathbf{c} \models^{\tau_1} K_a F_b C$. Informally, under τ_1 the agents are then more informed than under τ_2. This is for example the case for $\tau_1 = (\circ, \diamond, \beta)$ and $\tau_2 = (\bullet, \diamond, \alpha)$, since by Theorem 5.5

$$(\circ, \diamond, \beta) \equiv (\circ, \diamond, \alpha) \subseteq (\textcircled{\scriptsize}, \diamond, \alpha) \subset (\bullet, \diamond, \alpha).$$

We finally compare knowledge for call types that have a different direction type.

Proposition 5.7. *Consider two call types τ_1 and τ_2 such that $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p})$, $\tau_1(\mathrm{o}) = \tau_2(\mathrm{o})$ but $\tau_1(\mathrm{d}) = \diamond \neq \tau_2(\mathrm{d})$. Then*

$$\text{for all atoms } F_c D \text{ and all } \mathbf{c}, \quad \mathbf{c} \models^{\tau_2} F_c D \implies \mathbf{c} \models^{\tau_1} F_c D. \quad \text{(At)}$$

Moreover, equation (For) *holds iff $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p}) = \circ$.*

Proof. We use induction on the length $|\mathbf{c}|$ of $\mathbf{c}$. Assume that $\tau_1 = (\mathrm{p}, \diamond, \mathrm{o})$ and $\tau_2 = (\mathrm{p}, \triangleright, \mathrm{o})$. (The proof for $\tau_2 = (\mathrm{p}, \triangleleft, \mathrm{o})$ is analogous.) If $|\mathbf{c}| = 0$ then $\mathbf{c} = \epsilon$ and $\epsilon \models^{(\mathrm{p}, \triangleright, \mathrm{o})} F_c D$ iff $D = C$ iff $\epsilon \models^{(\mathrm{p}, \diamond, \mathrm{o})} F_c D$. Now suppose the claim is proven for $\mathbf{c}$ and consider $\mathbf{c}.ab$. For any agent $c \neq b$, we have by Definition 4.1 $\mathbf{c}.a \triangleright b \models^{\tau_2} F_c D$ iff $\mathbf{c} \models^{\tau_2} F_c D$, which implies by the induction hypothesis $\mathbf{c} \models^{\tau_1} F_c D$, and hence $\mathbf{c}.a \diamond b \models^{\tau_1} F_c D$. For agent b, we have $\mathbf{c}.a \triangleright b \models^{\tau_2} F_b D$ iff ($\mathbf{c} \models^{\tau_2} F_a D$ or $\mathbf{c} \models^{\tau_2} F_b D$) and $\mathbf{c}.a \diamond b \models^{\tau_1} F_b D$ iff ($\mathbf{c} \models^{\tau_1} F_a D$ or

$\mathbf{c} \models^{\tau_1} F_b D$), so the claim for b holds by the induction hypothesis, as well.

For the second part of the proposition, if $\tau_1(\mathrm{p}) = \circ$ then (For) holds because of (At), together with Theorem 5.1. If $\tau_1(\mathrm{p}) = \tau_2(\mathrm{p}) \neq \circ$, we provide a counterexample to (For) for the case $\tau_2(\mathrm{d}) = \triangleright$ and $\tau_1(\mathrm{o}) = \tau_2(\mathrm{o}) = \alpha$: $ac, cb, ba \models^{\tau_2} K_a K_c F_b C$, but $ac, cb, ba \not\models^{\tau_1} K_a K_c F_b C$ since $ac, cb, ba \sim_a^{\tau_1} ac, de, ba$. □

6 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, for example to call sequences that respect Protocol 2 (Hear my Secret). This is relevant in particular when one tries to incorporate into the framework an assumption that the agents have common knowledge of the underlying protocol.[2]

Example 6.1. Consider again Protocol 2 (Hear my Secret), four agents a, b, c, d, privacy degree ◓, and direction $\diamond$. Call sequence ab, bc, ad is compliant with the protocol. Because the privacy degree is ◓, agent c, who is not involved in the third call, still knows that a third call has taken place. Now if c knows that other agents also respect Protocol 2, she can conclude that d must be involved in the third call: given that she was not involved in the third call, c considers that the third call could have been ad, da, bd, or db. She therefore now knows that agent d is familiar with at least 3 secrets. Without that assumption, she cannot rule out that the third call is again ab. She then does not have this knowledge about d: she still considers it possible that d is only familiar with her own secret. □

First of all notice that Definition 4.4 constructs τ-dependent indistinguishability relations 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 with respect to the underlying protocol.

An ***epistemic gossip protocol*** (in short a protocol) consists of the union of $|Ag|$ sets of instructions, one for each agent, each instruction in the form

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

in symbols $\phi \to \mathrm{c}$, where $\phi \in \mathcal{L}$ is of the form $K_a\psi$ or $\hat{K}_a\psi$, and is referred to as an ***epistemic guard***. Such instructions 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.[3] We therefore think of a protocol P as a set of instructions $\phi \to \mathrm{c}$. For example, the instructions composing Protocol 2, are of the form

$$\hat{K}_a \neg F_b A \to ab$$

for all agents a and b. That is, if a considers it possible that b is not familiar with her secret, a calls b. Formally, given a gossip model $\mathcal{M}^\tau = (\mathbf{C}^\tau, \{\sim_a^\tau\}_{a \in Ag})$ we define the ***computation tree*** $\mathbf{C}_P^\tau \subseteq \mathbf{C}^\tau$ of a given protocol P (cf. [1]) as the set of call sequences inductively defined as follows:

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

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

So $\mathbf{C}_P^\tau$ 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 sequences as P-compliant.

Notice however, that in 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 should contain only P-compliant sequences. This circularity is not problematic for call types involving privacy degrees ○ and ◓, as the epistemic relation links only sequences of equal length, allowing therefore for call sequences and equivalence relations to be inductively constructed in parallel. That is however not the case for call types involving privacy degree ●, where call sequences of any length may be indistinguishable from the actual call sequence.

We sketch here a partial solution to the above issue, showing how one can construct a gossip model also for privacy degree ●, which consist only of sequences compliant with a given protocol P, under some assumptions on the logical form of epistemic guards. Define the following operation $\rho^P : 2^{\mathbf{C}^\tau} \to 2^{\mathbf{C}^\tau}$ such that $\rho(X) = \{\mathrm{c} \in X \mid \mathrm{c} \text{ is } P\text{-compliant}\}$, that is, ρ^P removes from a given set of call sequences those that are not P-compliant. What we are after is a set from which no sequences would be removed, a fixpoint of ρ^P:

Proposition 6.2. *Let the epistemic guards ϕ of P be in $\hat{\mathcal{L}}$. There exists an $X \subseteq \mathbf{C}^\tau$, such that $X = \rho^P(X)$.*

Proof. Observe that, under the assumption that $\phi \in \hat{\mathcal{L}}$, a call sequence is not P-compliant only if a call is executed while such a ϕ of type $\hat{K}_a\psi$ was not satisfied, which means that it lacked a sequence witnessing ψ. From this it follows that ρ^P is a monotonic function on sets, that is, $X \subseteq Y$ implies $\rho^P(X) \subseteq \rho^P(Y)$. By the Knaster-Tarski theorem, ρ^P has therefore fixpoints, including a largest and smallest fixpoint. □

Furthermore, by the Knaster-Tarski theorem one can construct the largest fixpoint of ρ^P by iteratively applying ρ^P to $\mathbf{C}^\tau$. Fixpoints of ρ^P—and most naturally the largest such fixpoint—can be used as domains for gossip models, restricting the definition of the indistinguishability relations $\sim_a^\tau$ to such domains. Gossip models constructed in this way incorporate the assumption that there is common knowledge among the agents about the protocol in use.

7 CONCLUSION

We provided an in-depth study of 18 different types of communication in gossip protocols and modelled their epistemic effects in a uniform way through different indistinguishability relations. This allowed us to establish various results concerning the relative informativeness of these types of communication. Such an analysis is highly relevant for the formal investigation of epistemic gossip protocols. Independently, it provides insights into various ways knowledge can spread in distributed systems and paves the way to an axiomatisation and proof theoretic comparison of the considered forms of communication.

[2] This issue was identified as an open problem in [1].

[3] This simple rendering of protocols suffices for the purposes of this section. More sophisticated formalizations of epistemic gossip protocols have been provided in [1, 3].

REFERENCES

[1] K. R. Apt, D. Grossi, and W. Van der Hoek. 2016. Epistemic Protocols for Distributed Gossiping. In *Proceedings Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2015)*. EPTCS, 51–56.

[2] K. R. Apt and D. Wojtczak. 2017. On the Computational Complexity of Gossip Protocols. In *Proceedings of IJCAI 2017*. 765–771.

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

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

[5] B. Baker and R. Shostak. 1972. Gossips and Telephones. *Discrete Mathematics* 2 (1972), 197–193.

[6] R. Bumby. 1981. A Problem with Telephones. *SIAM Journal of Algorithms and Discrete Methods* 2 (1981), 13–18.

[7] B. Chlebus and D. Kowalski. 2006. Robust Gossiping with an Application to Consensus. *J. Comput. System Sci.* 72 (2006), 1262–1281.

[8] M. Cooper, A. Herzig, F. Maffre, F. Maris, and P. Régnier. 2016. Simple Epistemic Planning: Generalised Gossiping. In *Proceedings of ECAI 2016*. 1563–1564.

[9] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. 1997. Knowledge-Based Programs. *Distributed Computing* 10 (1997), 199–225.

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

[11] P. Fraigniaud and E. Lazard. 1994. Methods and Problems of Communication in usual Networks. *Discrete Applied Mathematics* 53 (1994), 79–133.

[12] van Ditmarsch H., Grossi D., Herzig A., van der Hoek W., and Kuijer L. 2016. Parameters for Epistemic Gossip Problems. In *Proceedings of LOFT'16*.

[13] A. Hajnal, E. C. Milner, and E. Szemeredi. 1972. A Cure for the Telephone Disease. *Canad. Math. Bull.* 15 (1972), 447–450.

[14] J. Halpern and L. Zuck. 1992. A Little Knowledge Goes a Long Way: Knowledge-Based Derivations and Correctness Proofs for a Family of Protocols. *J. ACM* 39, 3 (1992), 449–478.

[15] S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman. 1988. A survey of Gossiping and Broadcasting in Communication Networks. *Networks* 18, 4 (1988), 319–349.

[16] A. Herzig and F. Maffre. 2017. How to Share Knowledge by Gossiping. *AI Communications* 30, 1 (2017), 1–17.

[17] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. 1996. Dissemination of Information in Interconnection Networks (Broadcasting and Gossiping). In *Combinatorial Network Theory*. Kluwer, 125–212.

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

[19] R. Kurki-Suonio. 1986. Towards Programming with Knowledge Expressions. In *Proceedings of POPL'86*. 140–149.

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

[21] R. Parikh and R. Ramanujam. 1985. Distributed Processing and the Logic of Knowledge. In *Logic of Programs (LNCS 193)*. Springer, 256–268. Similar to *JoLLI* 12: 453–467, 2003.

[22] A. Procaccia, Y. Bachrach, and J. Rosenschein. 2007. Gossip-Based Aggregation of Trust in Decentralized Reputation Systems. In *Proceedings of IJCAI'07*. 1470–1475.

[23] Á. Seress. 1986. Quick Gossiping without Duplicate Transmissions. *Graphs and Combinatorics* 2 (1986), 363–383.

[24] R. Tijdeman. 1971. On a Telephone Problem. *Nieuw Archief voor Wiskunde* 3(XIX) (1971), 188–192.

[25] H. van Ditmarsch, J. van Eijck, P. Pardo, R. Ramezanian, and F. Scharzentruber. 2017. Epistemic Protocols for Dynamic Gossip. *Journal of Applied Logic* 20 (2017), 1–31.