Intentionally Anonymous Public Announcements

Thomas Ågotnes[1,2], Rustam Galimullin[1], Ken Satoh[3], and Satoshi Tojo[4]

[1] University of Bergen, Bergen, Norway
[2] Shanxi University, Taiyuan, China
[3] Center for Juris-Informatics, Tokyo, Japan
[4] Asia University, Tokyo, Japan

Abstract. We formalise the notion of an *intentionally anonymous public announcement* in the tradition of public announcement logic. An anonymous announcement can be seen as in-between a public announcement from "the outside" (an announcement of φ) and a public announcement by one of the agents a (an announcement of $K_a\varphi$): we get more information than just φ, but not (necessarily) about exactly who made it. In this paper we assume that it is common knowledge that the announcer *intended* to be anonymous. Like in the Russian Cards puzzle, with that assumption, anonymous announcements in fact reveal more information than without. We introduce an operator for intentionally anonymous announcements, and show that in several ways it all boils down to the notion of a "safe" announcement (again, similarly to Russian Cards). We model safety via a fixed-point operator that is similar to common knowledge. Main formal results include comparisons of expressivity and axiomatic completeness for a language expressing safety.

Keywords: Public announcement logic · Dynamic epistemic logic · Modal logic · Expressive power · Completeness · Anonymity · Privacy · Security

1 Introduction

Taken at face value, the title of this paper seems to be an oxymoron. Indeed, if "public announcement" is taken literally, as in an agent *saying* something in front of everyone else, it will not be anonymous. However, anonymous public communication is almost ubiquitous in our day-to-day lives. Think of posts on social media and message boards done under a username instead of a real name of the poster. Or an anonymous letter to an editor of a news outlet. Other examples include anonymous emails, transactions on a blockchain, whistle blower reports, and even cultural artifacts created anonymously under an alias, like Elena Ferrante and MF DOOM.

The type of anonymity we are interested in here focuses on *action anonymity*, i.e., the inability of an attacker to identify who performed a given action (also sometimes referred to as *unlinkability* in the literature [22]). This is in contrast to *data anonymity*, i.e., the inability of an attacker to know the identity of a subject in an anonymised database, e.g., in medical records (see, e.g., [9]). One

of the standard requirements of both types of anonymity is that they satisfy k-anonymity [25], which intuitively means that a data record or an action cannot be distinguished from at least $k-1$ other records or actions. It is clear that in the case of public communication by someone in a group of agents, we should have at least 3-anonymity. Indeed, if a public announcement is so specific that it could be made only by two agents (and these two agents know that), then the non-announcing agent would be able to deduce the identity of the announcer. In the literature, such a scenario is called "background knowledge attack" [17].

In this paper we formalise anonymous public announcements inspired by *public announcement logic* (PAL) [23], an extension of multi-agent *epistemic logic* with constructs of the form $[\varphi!]\psi$ intuitively meaning that after φ is truthfully announced, ψ is true. In PAL the announced formula φ does not have to actually be known by any agent in the system – the identity of the announcer is left out of the picture. If the announcer indeed is one of the agents a in the system, the announcement in fact contains *more* information: in that case it would be modelled by the announcement $K_a\varphi!$. In this paper we formalise *anonymous* public announcements, conceptually somewhere in-between $\varphi!$ and $K_a\varphi!$ – we get *more* information than just φ but *less* than $K_a\varphi$ for a specific agent a. We make an additional crucial assumption: that it is common knowledge that the anonymous announcer *intended* to be anonymous, i.e., to not reveal her identity. Similarly to the Russian Cards puzzle [7], we shall see that that assumption actually means that the announcement reveals *more* information. To this end, we introduce and study an intentionally anonymous public announcement operator $[\varphi\ddagger]$, such that $[\varphi\ddagger]\psi$ intuitively means that after φ is anonymously announced by some agent, no matter who, and it is common knowledge that the announcement as intended to be anonymous, ψ is true.

Reasoning about anonymity based on (variants of) epistemic logic has been studied in [26, 11], with k-anonymity being discussed in [11]. Building on the runs-and-systems approach of [11], further extensions focus on privacy and onymity [27, 28] and electronic voting [18]. A knowledge-based approach to data anonymity was presented in [15]. Other logical approaches to anonymity and privacy include [24, 14, 32, 6]. The themes of anonymity and privacy are also related to the research on secrets in multi-agent systems (see, e.g., [12, 21, 20, 31]). None of these approaches model anonymous *announcements*. Finally, an approach to anonymity based on *dynamic epistemic logic* (DEL) [8] was presented in [29], where the authors consider scenarios of secret communication between agents in a system. Such secret communication is captured by private announcements, as special type of *action models* [5]. In this setting, verifying whether secret communication remains secret boils down to model checking an epistemic formula in the resulting updated model, i.e. the model that is obtained after an application of a dynamic operator. While our approach is also DEL-based, there are two major differences. First, we model anonymous *public* communication, as opposed to *private* communication. Second, we tackle the crucial issue of *intentional* anonymity, where it is common knowledge that only *safe* announcements are made.

We start out by introducing the machinery of EL and DEL in Section 2. In Section 3, we formalise intentional anonymous announcements as well as introducing a *safety* modality, which is a fixed-point modality somewhat similar to common knowledge. In Section 4 we look at expressivity and show that, in several senses, safety is fundamental, and in particular that the logic with the intentionally anonymous announcement operators can be reduced to epistemic logic with safety (but not to EL, PAL or, in general, DEL *without* the safety modality). In Section 5 we give a sound an complete axiomatisation of the latter. We end with a discussion in Section 6.

2 Background

We give a brief review of relevant background concepts and refer the reader to [8] for further details. Let N be a finite set of *agents*, and P be a countable set of *propositional variables*. All logics in this paper are interpreted on epistemic models.

Definition 1. *An* (epistemic) model *is a triple $M = (S, \sim, V)$, where S is a non-empty set of states, $\sim\colon N \to 2^{S\times S}$ is an equivalence relation for each $i \in N$, and $V : P \to 2^S$ is the valuation function. For $s \in S$, a pair M, s is called a* pointed (epistemic) model.

Definition 2 (Action Models). *Let $\mathcal{L}$ be a language defined over the signature $\langle N, P\rangle$. An* action model *is a triple $\mathsf{M} = (\mathsf{S}, \sim, \mathsf{pre})$, where S is a non-empty set of states, $\sim\colon N \to 2^{\mathsf{S}\times\mathsf{S}}$ is an equivalence relation for each $i \in N$, and $\mathsf{pre} : \mathsf{S} \to \mathcal{L}$ is the precondition function. For $\mathsf{s} \in \mathsf{S}$, we will call a pair M, s a* pointed action model.

Definition 3. *Languages of* epistemic logic *($\mathcal{L}_K$),* public announcement logic *($\mathcal{L}_!$), and* action model logic *($\mathcal{L}_\otimes$) are defined by the following BNFs:*

$$\begin{array}{lll}
\mathcal{L}_K & \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi & \\
\mathcal{L}_! & \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid [\varphi!]\varphi & \\
\mathcal{L}_\otimes & \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid [\pi]\varphi & \pi ::= (\mathsf{M}, \mathsf{s}) \mid \pi \cup \pi
\end{array}$$

where $p \in P$, $i \in N$, and (M, s) is a pointed action model with a finite set of states S, and such that for all $\mathsf{s} \in \mathsf{S}$, the precondition $\mathsf{pre}(\mathsf{s})$ is some $\varphi \in \mathcal{L}_\otimes$ that was constructed in a previous stage of the inductively defined hierarchy. We write $\hat{K}_i\varphi$ for $\neg K_i\neg\varphi$, and when $G \subseteq N$ we write $E_G\varphi$ for $\bigwedge_{i\in G} K_i\varphi$.

The semantics is now defined as follows.

Definition 4. *Let $M, s = (S, \sim, V)$ be a model, $p \in P$, $i \in N$, and (M, s) be an action model.*

$$
\begin{array}{ll}
M, s \models p & \textit{iff } s \in V(p) \\
M, s \models \neg\varphi & \textit{iff } M, s \not\models \varphi \\
M, s \models \varphi \wedge \psi & \textit{iff } M, s \models \varphi \textit{ and } M, s \models \psi \\
M, s \models K_i\varphi & \textit{iff } M, t \models \varphi \textit{ for all } t \in S \textit{ such that } s \sim_i t \\
M, s \models [\psi!]\varphi & \textit{iff } M, s \models \psi \textit{ implies } M, s^{\psi!} \models \varphi \\
M, s \models [\mathsf{M}, \mathsf{s}]\varphi & \textit{iff } M, s \models \mathsf{pre(s)} \textit{ implies } (M \otimes \mathsf{M}, (s, \mathsf{s})) \models \varphi \\
M, s \models [\pi \cup \rho]\varphi & \textit{iff } M, s \models [\pi]\varphi \textit{ and } M, s \models [\rho]\varphi
\end{array}
$$

We write $[\![\varphi]\!]_M$ *for the set* $\{s \in S \mid M, s \models \varphi\}$. *The updated model* $M^{\varphi!}$ *is* $(S^{\varphi!}, \sim^{\varphi!}, V^{\varphi!})$, *where* $S^{\varphi!} = [\![\varphi]\!]_M$, $\sim_i^{\varphi!} = \sim_i \cap (S^{\varphi!} \times S^{\varphi!})$ *for all* $i \in N$, *and* $V^{\varphi!}(p) = V(p) \cap [\![\varphi]\!]_M$ *for all* $p \in P$. *The updated model* $M \otimes \mathsf{M}$ *is* $(S', \sim', V')$, *where* $S' = \{(s, \mathsf{s}) \mid s \in S, \mathsf{s} \in \mathsf{S}, M, s \models \mathsf{pre(s)}\}$, $(s, \mathsf{s}) \sim_i' (t, \mathsf{t})$ *iff* $s \sim_i t$ *and* $\mathsf{s} \sim_i \mathsf{t}$, *and* $(s, \mathsf{s}) \in V'(p)$ *iff* $s \in V(p)$. *We call a formula* φ valid, *or a* validity, *if for all* M, s *it holds that* $M, s \models \varphi$.

Definition 5. *Let* $M^1 = (S^1, \sim^1, V^1)$ *and* $M^2 = (S^2, \sim^2, V^2)$ *be two epistemic models. We say that* M^1 *and* M^2 *are* bisimilar *(denoted* $M^1 \leftrightarroweq M^2$*) if there is a non-empty relation* $Z \subseteq S^1 \times S^2$, *called a* bisimulation, *such that for all* sZt:

Atoms *for all* $p \in P$: $s \in V^1(p)$ *if and only if* $t \in V^2(p)$,
Forth *for all* $i \in N$ *and* $u \in S^1$ *such that* $s \sim_i^1 u$, *there is a* $v \in S^2$ *such that* $t \sim_i^2 v$ *and* uZv,
Back *for all* $i \in N$ *and* $v \in S^2$ *such that* $t \sim_i^2 v$, *there is a* $u \in S^1$ *such that* $s \sim_i^1 u$ *and* uZv.

We say that M^1, s *and* M^2, t *are bisimilar and denote this by* $M^1, s \leftrightarroweq M^2, t$ *if there is a bisimulation linking states* s *and* t.

It is a standard result that $M^1, s \leftrightarroweq M^2, t$ implies $M^1, s \models \varphi$ if and only if $M^2, t \models \varphi$ for $\varphi \in \mathcal{L}_K \cup \mathcal{L}_! \cup \mathcal{L}_\otimes$ (see, e.g, [8, Chapter 5]).

Definition 6. *Let* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *be two languages defined over the same class of models, and let* $\varphi \in \mathcal{L}_1$ *and* $\psi \in \mathcal{L}_2$. *We say that* φ *and* ψ *are* equivalent, *when for all* M, s: $M, s \models \varphi$ *if and only if* $M, s \models \psi$. *If for every* $\varphi \in \mathcal{L}_1$ *there is an equivalent* $\psi \in \mathcal{L}_2$, *we write* $\mathcal{L}_1 \preccurlyeq \mathcal{L}_2$ *and say that* $\mathcal{L}_2$ *is* at least as expressive as $\mathcal{L}_1$. *We write* $\mathcal{L}_1 \prec \mathcal{L}_2$ *iff* $\mathcal{L}_1 \preccurlyeq \mathcal{L}_2$ *and* $\mathcal{L}_2 \not\preccurlyeq \mathcal{L}_1$, *and we say that* $\mathcal{L}_2$ *is* strictly more expressive than $\mathcal{L}_1$. *Finally, if* $\mathcal{L}_1 \preccurlyeq \mathcal{L}_2$ *and* $\mathcal{L}_2 \preccurlyeq \mathcal{L}_1$, *we say that* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *are* equally expressive *and write* $\mathcal{L}_1 \approx \mathcal{L}_2$.

It is well known that $\mathcal{L}_! \approx \mathcal{L}_\otimes \approx \mathcal{L}_K$ [8, Chapter 8].

3 Intentionally Anonymous Public Announcements

An anonymous public announcement of φ (made by an agent in the system) is clearly not identical to a standard public announcement of φ, as in the latter case an announced formula may not even be known to any of the agents. Thus a necessary precondition for an (anonymous or not) announcement made by some

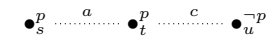

Fig. 1. Three-agent epistemic model with the agent b's relation being the identity.

agent a would be that a knows φ, i.e. that $K_a\varphi$. However, this is not sufficient to guarantee anonymity: for anonymity we also need someone *else* to know φ – in fact, as discussed in the introduction, we need at least *3-anonymity*. We are modeling announcements that are *intentionally* anonymous in the sense that it is common knowledge that an agent will only make the announcement if she knows that it is *safe*. We will discuss in detail what "safe" means, but the intuition is that it is safe for an agent to make an announcement if somebody else could have made it. One subtlety is that the potential announcement by that "somebody else" also should be safe, so the definition has a recursive (or even circular!) flavour. To make our intuitions more precise, consider model M in Figure 1.

- In state s, all three agents a, b and c know p
- In state s, a considers it possible that we are in state t, where it would not be safe for her to announce p since then b would know that it was her. Thus it is not safe for a to announce p in s.
- In state s, c knows that it is not safe for a to announce p in s, and thus it is not safe for c either: if she announces p in s, b would know that it was her since b knows that it is not safe for a to announce p in s.
- It is thus not safe for b to announce p in s either.
- It is not safe for a or b to announce p in t (only two agents know p).
- p cannot be announced by anyone in u.

Thus, no safe announcements can be made in this model. To see when it is actually safe for agents to make anonymous announcements, we need to provide precise definitions of what we mean by safety and intentionally anonymous announcements.

As we have seen, it is not enough that three agents know φ to ensure that φ can be safely announced by any of them; those three agents also need to know that φ is safe, i.e., that three agents know φ and that φ is safe ... and so on. This clearly has the flavour of *common knowledge*. However, the existence of a group of three different agents having common knowledge of φ, $M, s \models \bigvee_{a\neq b\neq c} C_{\{a,b,c\}}\varphi$, is sufficient but not necessary for φ to be safe in s. A weaker condition would also be sufficient: we don't need the three agents to know that *the same* three agents can safely announce φ. It might be, for example, that a considers it possible that a, b and d safely can announce φ.

In order to define that weaker condition, let us recall the fixpoint definition of common knowledge (see, e.g., [10]). It is well known that $C_G\varphi$ is the greatest fixpoint of $E_G(\varphi \wedge x)$ w.r.t. a variable x not occurring in φ, or, more accurately, given an epistemic model M, $[\![C_G\varphi]\!]_M$ is the greatest fixpoint of the function $f(S) = [\![E_G(\varphi \wedge x)]\!]^M_{V[x=S]}$ where $[\![\psi]\!]^M_{V[x=S]}$ is the extension of ψ in the model M where the valuation function has been changed

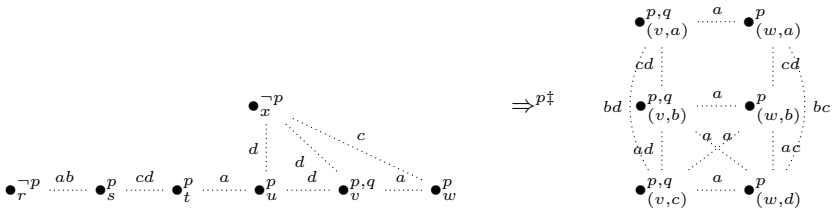

Fig. 2. Four-agent epistemic model M (left), with updated model $M^{p\ddagger}$ (right).

so that the extension of the variable x is S. The greatest fixed-point of monotonic functions is equal to the union of all post-fixed points, so we have that: $[\![C_G\varphi]\!] = \bigcup \left\{ S : S \subseteq [\![E_G(\varphi \wedge x)]\!]^M_{[x=S]} \right\}$.

We now define a similar, weaker, notion of common knowledge in order to capture safety. Let us introduce a safety operator: $\blacktriangle\varphi$ intuitively means that φ is safe in the current state. We let, where $N^3 = \{\{a,b,c\} \mid a,b,c \in N; a \neq b, a \neq c, b \neq c\}$:

$$M, s \models \blacktriangle\varphi \text{ iff } s \in \bigcup \left\{ S : S \subseteq [\![\bigvee_{G \in N^3} E_G(\varphi \wedge x)]\!]^M_{[x=S]} \right\}.$$

In other words, $\blacktriangle\varphi$ is the greatest fixpoint of $\bigvee_{G\in N^3} E_G(\varphi \wedge x)$. Thus, in particular we have that

$$\models \blacktriangle\varphi \leftrightarrow \bigvee_{G \in N^3} E_G(\varphi \wedge \blacktriangle\varphi),$$

and, similarly to the iterative definition of common knowledge, we have the following.

Lemma 1. *$M, s \models \blacktriangle\varphi$ iff $M, s \models \varphi$ and $M, s \models \bigvee_{G_1 \in N^3} E_{G_1}\varphi$ and $M, s \models \bigvee_{G_1 \in N^3} E_{G_1}(\varphi \wedge \bigvee_{G_2 \in N^3} E_{G_2}\varphi)$ and $\cdots$.*

While $\blacktriangle\varphi$ means that φ can safely be announced, $K_a\blacktriangle\varphi$ means that φ can safely be announced *by* a. It is easy to see that: $M, s \models K_a\blacktriangle\varphi$ iff $M, s \models \varphi$ and $M, s \models \bigvee_{G_1 \in N^3, a \in G_1} E_{G_1}\varphi$ and $M, s \models \bigvee_{G_1 \in N^3, a \in G_1} E_{G_1}(\varphi \wedge \bigvee_{G_2 \in N^3} E_{G_2}\varphi)$ and $\cdots$.

Finally we can introduce an operator for intentionally anonymous announcements, $[\varphi\ddagger]$, with the corresponding model update defined as follows.

Definition 7. *The update of epistemic model $M = (S, \sim, V)$ by the intentionally anonymous announcement of φ is the epistemic model $M^{\varphi\ddagger} = (S', \sim', V')$ where:*

- $S' = \{(s, a) : s \in S, a \in N, M, s \models K_a\blacktriangle\varphi\}$
- $(s, a) \sim'_c (t, b)$ *iff* $s \sim_c t$ *and* $a = c$ *iff* $b = c$
- $V'(s, a) = V(s)$

Intuitively, in the updated model it is common knowledge that someone has (truthfully) intentionally anonymously announced φ. A state (s, a) corresponds to that someone being a, in state s of the original model. The precondition is that a knows that φ is safe. An agent can only discern between a situation (s, a) where a made the announcement and a situation (t, b) where a different agent b did, if she could already discern between s and t before the announcement or she is exactly one of a or b. We then let:

$$M, s \models [\varphi\ddagger]\psi \Leftrightarrow \forall a \in N, \big(M, s \models K_a \blacktriangle \varphi \Rightarrow M^{\varphi\ddagger}, (s, a) \models \psi\big)$$

– after the intentionally anonymous announcement of φ, ψ is true no matter who the announcer was.

Thus, we have introduced two new operators: $\blacktriangle$ and $[\varphi\ddagger]$. Do we want both in the formal language? Let's consider all three combinations, which will be useful later[5].

$$\begin{array}{ll} \mathcal{L}_{\ddagger} & \varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [\varphi\ddagger]\varphi \\ \mathcal{L}_{\blacktriangle} & \varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid \blacktriangle\varphi \\ \mathcal{L}_{\ddagger\blacktriangle} & \varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [\varphi\ddagger]\varphi \mid \blacktriangle\varphi \end{array}$$

Let's look at some examples. In the model in Figure 2:

- $M, s \models \neg\blacktriangle p$. p can't be safely announced in s: only two agents know p.
- $M, t \models \neg\blacktriangle p$. While all four agents know p in t, they do not all know that three agents know p: c and d consider it possible that only two agents know p. In other words, even though they know p, they do not know that it can be safely announced.
- $M, u \models \neg\blacktriangle p$. Three agents, a, b and c, know p. They also know that at least three agents know p. But they still do not know that p is safe: they do not know that at least three agents know that at least three agents know p.
- $M, v \models \blacktriangle p$ and $M, w \models \blacktriangle p$. p can be safely announced in v and in w. In fact, $\{w, v\}$ is the greatest fixed-point of $\bigvee_{G \in N^3} E_G(\varphi \wedge x)$.
- $M, v \models K_a \blacktriangle p$. p can be safely announced by a in v.
- Even though a knows *that* p is safe in v, she does not know *why*. She considers it possible that we indeed are in v in which it would be safe for a, b and c to announce p, or in w in which it would be safe for a, b and d to announce p, but she doesn't know which it is. We can say that she knows *de dicto* that p is safe, but not *de re*.
- $M, v \models [p\ddagger]K_c K_d q$. After an intentionally anonymous announcement of p in v, c knows that d has learned that q is true.
- In every state in the updated model, *only the announcer knows who the announcer was*.

The latter point is no coincidence: safe announcements are indeed always safe.

[5] $\mathcal{L}_{\ddagger}$ is well defined without $\blacktriangle$ in the syntax, even though $\blacktriangle$ is used in the semantic condition for $[\varphi\ddagger]$. That semantic condition could of course be written without explicitly mentioning $\blacktriangle$.

Lemma 2. *In any update by an intentionally anonymous announcement, it is common knowledge that no-one except the announcer knows who the announcer is. Formally, let M be a model, let $M^{\varphi\ddagger} = (S', \sim', V')$ be the updated model after the intentionally anonymous announcement of φ, and let $(s, a) \in S'$. Then for any agent $i \neq a$, there is a state $(s', a') \in S'$ such that $(s, a) \sim'_i (s', a')$ and $a \neq a'$.*

Proof. Let M be a model. Let (s, a) be a state in $M^{\varphi\ddagger}$ and let $d \neq a$. By definition, $M, s \models K_a \blacktriangle \varphi$, so there are b, c such that $\{a, b, c\} \in N^3$ and $M, s \models K_b \blacktriangle \varphi \wedge K_c \blacktriangle \varphi$. I.e., (s, b) and (s, c) are states in $M^{\varphi\ddagger}$. Either $b \neq d$ or $c \neq d$ (or both); assume the former. Since all of a, b and d are different, and $\sim_d$ is reflexive, we have that $(s, a) \sim'_d (s, b)$. □

Thus, $K_a \blacktriangle \varphi$ is a *sufficient* condition for a safe anonymous announcement: after it is announced, no one, apart from a herself, will know who the announcer was. It is also *necessary*, in the sense that a must know that at least two other agents know φ, and must know that those two agents know that at least two other agents know φ, and so on, i.e., it is the case that $K_a \blacktriangle \varphi$. That justifies the definition of the model update from an intentionally anonymous announcement: that $K_a \blacktriangle \varphi$ holds for *some* a is the information that is revealed.

The reader might have observed that the updated model $M^{p\ddagger}$ in Figure 2 is bisimilar to the submodel of M it is projected on, i.e., the submodel consisting of states v and w – exactly the submodel where $K_a \blacktriangle p$ holds for some a. This is, in fact, always the case. In the following lemma the model update $M^{\blacktriangle\varphi!}$ is defined like for PAL (Section 2).

Lemma 3. *For any φ, ψ, M, s: $M, s \models [\varphi\ddagger]\psi$ iff $\big(M, s \models \blacktriangle\varphi \Rightarrow M^{\blacktriangle\varphi!}, s \models \psi\big)$.*

Proof. We first show that $M^{\varphi\ddagger}$ and $M^{\blacktriangle\varphi!}$ are bisimilar. Let Z be such that $(s, a) Z s'$ iff $s = s'$, and let $(s, a) Z s$. *Atoms* is straightforward. *Forth* is as well: if $(s, a) \sim_c (t, b)$, then $s \sim_c t$ by definition. For *Back*, let $s \sim_c t$. First, consider that $c \neq a$. Since t is in $M^{\blacktriangle\varphi!}$, $M, t \models \blacktriangle\varphi$. Thus, there are different $b, d, e \in N$ such that $M, t \models K_b \blacktriangle \varphi \wedge K_d \blacktriangle \varphi \wedge K_e \blacktriangle \varphi$. At least one of b, d, e is different from both a and c. Say, b. Thus $(s, a) \sim_c (t, b)$. Second, consider that $c = a$. Since $M, s \models K_c \blacktriangle$ and $s \sim_c t$, we also have that $M, t \models K_c \blacktriangle$. Thus (t, c) is in $M^{\varphi\ddagger}$, and we have that $(s, a) \sim_c (t, c)$.

Note that satisfaction of $\blacktriangle\varphi$ is invariant under bisimulation (this follows, e.g., from Lemma 1). Then we have that $M, s \models [\varphi\ddagger]\psi$ iff for all a, $M, s \models K_a \blacktriangle \varphi$ implies $M^{\varphi\ddagger}, (s, a) \models \psi$ iff for all a, $M, s \models K_a \blacktriangle \varphi$ implies $M^{\blacktriangle\varphi!}, s \models \psi$. We argue that this is equivalent to $M, s \models \blacktriangle\varphi$ implies $M^{\blacktriangle\varphi!}, s \models \psi$. For the implication from left to right, if $M, s \models \blacktriangle\varphi$ then $M, s \models K_a \blacktriangle \varphi$ for some a by the fixed-point property of $\blacktriangle$ (actually that holds for three different $a \in N$). From $M, s \models K_a \blacktriangle \varphi \Rightarrow M^{\blacktriangle\varphi!}, s \models \psi$ for all a, we get that $M^{\blacktriangle\varphi!}, s \models \psi$. For the other direction, let $a \in N$. If $M, s \models K_a \blacktriangle \varphi$, then $M, s \models \blacktriangle\varphi$ by reflexivity, so $M^{\blacktriangle\varphi!}, s \models \psi$. □

As an immediate corollary of Lemma 3 we have that if we allow public announcement operators in the language:

$$\models [\varphi\ddagger]\psi \leftrightarrow [\blacktriangle\varphi!]\psi.$$

So, safe anonymous announcements are exactly public announcements of safety. Thus, we get a "simpler", logically equivalent semantics. The existence of the original semantics and the fact that it corresponds exactly to this alternative semantics is of course still crucial: it is what ensures that safety is safe (Lemma 2)[6]. It also means that PAL extended with $\blacktriangle$ can express intentionally anonymous announcements. However, as we shall see when we now move on to comparing expressive power of the languages introduced above, even less is needed.

4 Expressive power

There is a close relationship between intentionally anonymous announcements and action models. Indeed, consider the language $\mathcal{L}_{\otimes\blacktriangle}$ of action model logic extended with the safety modality. The semantics of $\varphi\ddagger$ given above is equivalent to a certain class of action models of $\mathcal{L}_{\otimes\blacktriangle}$: *anonymous action models*.

Definition 8. *The* anonymous action model *for N and formula φ is the action model* $\mathsf{M}^{\mathsf{N}}_{\varphi} = (\mathsf{S}, \sim, \mathsf{pre})$ *where* $\mathsf{S} = N$, $\mathsf{a} \sim_{\mathsf{c}} \mathsf{b}$ *iff* $(\mathsf{a} = \mathsf{c} \Leftrightarrow \mathsf{b} = \mathsf{c})$, *and* $\mathsf{pre}(\mathsf{a}) = K_a\blacktriangle\varphi$.

An intentionally anonymous announcement now corresponds to the *union* of events $(\mathsf{M}^{\mathsf{N}}_{\varphi}, \mathsf{a})$ for all agents a. The proof of the following is straightforward.

Lemma 4. *For any pointed epistemic model M, s and formulas φ and ψ, $M, s \models [\varphi\ddagger]\psi$ iff $M, s \models \left[\bigcup_{\mathsf{i}\in\mathsf{N}}(\mathsf{M}^{\mathsf{N}}_{\varphi}, \mathsf{i})\right]\psi$.*

Consider now the three languages with safety and/or intentional anonymous announcements: $\mathcal{L}_{\ddagger\blacktriangle}, \mathcal{L}_{\blacktriangle}$ and $\mathcal{L}_{\ddagger}$. The first is similar to that of action model logic with common knowledge $\mathcal{L}_{\otimes C}$. In the same sense, $\mathcal{L}_{\blacktriangle}$ is similar to epistemic logic with common knowledge $\mathcal{L}_{KC}$. $\mathcal{L}_{\ddagger}$ is in-between: it would correspond to $\mathcal{L}_{\otimes C}$ but where common knowledge appears only in the preconditions of action models. It is known that $\mathcal{L}_{\otimes C}$ is strictly more expressive than $\mathcal{L}_{KC}$ [8, Chapter 8]; in particular expressions of the form $[M, s]C_G\varphi$ cannot always be reduced to an $\mathcal{L}_{KC}$ formula. Nevertheless, we shall now see that, perhaps contrary to intuition,

[6] We used invariance under standard bisimulation as a technical tool to establish the correspondence. Of course, when we say that "only the announcer knows who the announcer is" and so on, we implicitly mean that agents can potentially discern between states (s, a) and (s, b), even though, as shown above, they are bisimilar. An extended notion of bisimulation could be introduced to take into account this assumption, but it is in any case not picked up by the logical language. It is of course also possible to think about ways this distinction *could* be picked up by the language, e.g., by special atoms, but that is complicated by the possibility of iterated announcements leading to states like $((s, a), b)$ and so on.

all three languages $\mathcal{L}_{\ddagger\blacktriangle}, \mathcal{L}_{\blacktriangle}$ *and* $\mathcal{L}_{\ddagger}$ *are actually equally expressive.* Let us start with the "intermediate" language $\mathcal{L}_{\ddagger}$. By Lemma 4, every $\varphi\ddagger$ has an equivalent action model with preconditions $K_a\blacktriangle\varphi$. Because preconditions contain the $\blacktriangle$ modality, we cannot translate formulas of $\mathcal{L}_{\otimes\blacktriangle}$ into $\mathcal{L}_K$ since we do not have a reduction axiom for $\blacktriangle\varphi$. However, we *can* reduce formulas of $\mathcal{L}_{\otimes\blacktriangle}$ to formulas of $\mathcal{L}_{\blacktriangle}$.

Lemma 5. $\mathcal{L}_{\ddagger} \preccurlyeq \mathcal{L}_{\blacktriangle}$.

Proof. Consider the language $\mathcal{L}^{-}_{\otimes\blacktriangle}$, which is like $\mathcal{L}_{\otimes\blacktriangle}$ but with the restriction that $\blacktriangle$ is not allowed in the scope of a $[\pi]$ modality. It is easy to see that the reduction axioms for $\mathcal{L}_{\otimes}$ [8, Chapter 6] are still valid for this extended language. Take a $\varphi \in \mathcal{L}_{\ddagger}$, and let ψ be the corresponding $\mathcal{L}_{\otimes\blacktriangle}$ formula by replacing every $[\varphi\ddagger]$ with the corresponding action model expression $[\pi]$ (Lemma 4). Now we recursively reduce the number of occurrences of $[\pi]$ modalities in ψ by starting with an *outermost* one, i.e., one that is not in the scope of any other. Use the corresponding reduction axiom of $\mathcal{L}_{\otimes}$, and repeat until all $[\pi]$ modalities are gone. Every time we use a reduction axiom we might introduce a new subformula $K_a\blacktriangle\psi'$, but never in the scope of a $[\pi]$ modality since we started with the outermost one. Thus, every time we use a reduction axiom the result is in $\mathcal{L}^{-}_{\otimes\blacktriangle}$, and when we are done the result is in pure epistemic logic extended with $\blacktriangle$, i.e. in $\mathcal{L}_{\blacktriangle}$. □

Thus, $\mathcal{L}_{\ddagger}$ can be "reduced" to $\mathcal{L}_{\blacktriangle}$ (even though the latter is not a sublanguage). The second "surprise" is that the $\blacktriangle$ operator can actually be expressed by $[\varphi\ddagger]$.

Lemma 6. *For any* M, s *and any* $\varphi \in \mathcal{L}_{\ddagger\blacktriangle}$, $M, s \models \blacktriangle\varphi$ *iff* $M, s \models \neg[\varphi\ddagger]\bot$.

Proof. $M, s \models \neg[\varphi\ddagger]\bot$ iff there is an $a \in N$ such that $M, s \models K_a\blacktriangle\varphi$ and $M^{\varphi\ddagger}, (s, a) \not\models \bot$ iff there is an $a \in N$ such that $M, s \models K_a\blacktriangle\varphi$ iff (by reflexivity in one direction and the fixed-point definition of $\blacktriangle$ in the other) $M, s \models \blacktriangle\varphi$. □

Corollary 1. $\mathcal{L}_{\ddagger\blacktriangle} \approx \mathcal{L}_{\ddagger} \approx \mathcal{L}_{\blacktriangle}$.

Thus, yet again we see that dynamic announcement operators can actually be expressed in a purely static language: $\mathcal{L}_{\blacktriangle}$. We thus move on to axiomatising safety.

5 Axiomatisation of safety

Observe that $\blacktriangle$ does not distribute over implication: $\not\models \blacktriangle(\varphi \to \psi) \to (\blacktriangle\varphi \to \blacktriangle\psi)$. In particular, $\bigvee_{G\in N^3} E_G(\varphi \to \psi)$ and $\bigvee_{G\in N^3} E_G\varphi$ does not imply $\bigvee_{G\in N^3} E_G\psi$ – it might not be the same G in the first two cases. For similar reasons the conjunctive closure axiom, $(\blacktriangle\varphi \wedge \blacktriangle\psi) \to \blacktriangle(\varphi \wedge \psi)$ does not hold. This is similar to *somebody knows* [2], but unlike most other group knowledge operators which are normal modalities. The other direction of conjunctive closure, *monotonicity*, $\blacktriangle(\varphi \wedge \psi) \to (\blacktriangle\varphi \wedge \blacktriangle\psi)$, does hold.

all instances of propositional tautologies	Prop
From $\varphi \to \psi$ and φ, derive ψ	Modus ponens
$K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	Distribution
$K_a\varphi \to \varphi$	Truth
$\neg K_a\varphi \to K_a\neg K_a\varphi$	Negative introspection
From φ, derive $K_a\varphi$	Necessitation
$\blacktriangle\varphi \to \bigvee_{G\in N^3} E_G(\varphi \wedge \blacktriangle\varphi)$	Mix
From $\varphi \to \bigvee_{G\in N^3} E_G\varphi$, derive $\varphi \to \blacktriangle\varphi$	Induction rule
From $\varphi \to \psi$, derive $\blacktriangle\varphi \to \blacktriangle\psi$	Monotonicity rule

Table 1. The proof system $\mathbf{S}\blacktriangle$.

The axiomatic system $\mathbf{S}\blacktriangle$ for $\mathcal{L}_{\blacktriangle}$ is shown in Table 1. The first two parts form a standard axiomatisation of propositional logic and the individual knowledge operators. The Monotonicity rule combines the monotonicity axiom and the replacement of equivalents rule standard in weak modal logics. We note that necessitation for $\blacktriangle$, from φ derive $\blacktriangle\varphi$, follows (we have $\blacktriangle\top$ from the Induction rule). The Mix axiom (or *fixed-point axiom*) says that $\blacktriangle\varphi$ is indeed a fixed-point of $\bigvee_{G\in N^3} E_G(\varphi \wedge x)$. Finally, the Induction rule give us a way to derive $\blacktriangle\varphi$. Mix and the Induction rule can be seen as adaptions of similar axioms/rules for common knowledge, see, e.g., [10, 8].

We will use the following shorthand:

$$\blacktriangle_n\varphi = \varphi \wedge \bigvee_{G_1\in N^3} E_{G_1}(\varphi \wedge \bigvee_{G_2\in N^3} E_{G_2}(\varphi \wedge \bigvee_{G_3\in N^3} E_{G_3}(\varphi \wedge \cdots \wedge \bigvee_{G_n\in N^3} E_{G_n}\varphi))).$$

Thus, $\blacktriangle_0 = \varphi$, $\blacktriangle_1 = \varphi \wedge \bigvee_{G_1\in N^3} E_{G_1}\varphi$, and so on, and we have that $M, s \models \blacktriangle\varphi$ iff $M, s \models \blacktriangle_n\varphi$ for all n.

It is straightforward to see that the Mix axiom is valid and that the Induction and Monotonicity rules preserve validity, and thus that the following holds.

Lemma 7. $\mathbf{S}\blacktriangle$*is sound.*

We now prove that $\mathbf{S}\blacktriangle$ is also complete. Similarly to epistemic logic with common knowledge, the logic is not *compact*. For example, $\{\blacktriangle_n\varphi : n \geq 0\} \cup \{\neg\blacktriangle\varphi\}$ is not satisfiable, but any finite subset of it is. Thus, we have to settle for *weak* rather than *strong* completeness. To that end, we adapt the standard technique for common knowledge of defining a finite canonical model based on the finite syntactic closure of some formula. This is complicated by the fact that $\blacktriangle$ is not normal: in the axiomatisation we have the (weaker) monotonicity rule instead of the K axiom for $\blacktriangle$. The consequence, aside from the fact that we cannot rely on distribution over implication like in proofs for common knowledge, is that there is no normal relational semantics for $\blacktriangle$ (like "φ is true on all states reachable by a G-path" for $C_G\varphi$). We now define an alternative definition of the semantics and show that it is equivalent, which we will use in the completeness proof.

Given a model M, a *group assignment function* for M is a function $f : S \to \mathcal{P}(N)$ such that $f(s) \in N^3$ or $f(s) = \emptyset$ for each s, i.e., assigning a group of three agents to some of the states, such that for all $s, t \in S$ and $i \in f(s)$, if $s \sim_i t$ then $f(t) \neq \emptyset$. The idea is that f works as an assignment of groups of three agents to certain states, such that if we follow the accessibility relations for those agents, f again assigns a group of three agents to those states, and so on. An *f-consistent path* in M is a finite sequence of states $s_0 s_1 s_2 \cdots s_m$ such that for all $0 \leq i < m$, $s_i \sim_a s_{i+1}$ for some $a \in f(s_i)$. We say that φ is true on a path if it is true in every state on that path.

Lemma 8. *$M, s \models \blacktriangle\varphi$ iff there is a group assignment function f such that $f(s) \neq \emptyset$ and φ is true on every f-consistent path starting in s.*

Proof. Left-to-right: let $M, s \models \blacktriangle\varphi$. Define f as follows: for any state t, $f(t) = G$ for some $G \in N^3$ such that $M, t \models E_G \blacktriangle\varphi$ if such a G exists, $f(t) = \emptyset$ if not. If there are several such G chose any of them. We must show that f is indeed a group assignment function. Let $i \in f(s')$ and $s' \sim_i t$; we must show that $f(t) \neq \emptyset$. Since $i \in f(s')$, there is a $G \in N^3$ such that $M, s' \models E_G \blacktriangle\varphi$ and $i \in G$. Thus $M, s \models K_i \blacktriangle\varphi$ and $M, t \models \blacktriangle\varphi$. By Mix, $M, t \models E_{G'} \blacktriangle\varphi$ for some G', so $f(t) \neq \emptyset$. Finally, $f(s) \neq \emptyset$ from $M, s \models \blacktriangle\varphi$ and Mix.

We proceed by induction on the length of any f-consistent path starting in s. The base case is immediate. Consider an f-consistent path $s_0 \cdots s_n s_{n+1}$ where $s_0 = s$. By the induction hypothesis $s_0, \cdots, s_n$ are φ-states. Since $s_n \sim_i s_{n+1}$ for some $i \in f(s_n)$ and $M, s_n \models E_{f(s_n)} \blacktriangle\varphi$, we thus also have $M, s_{n+1} \models \varphi$.

Right-to-left: by contraposition. Assume that $M, s \not\models \blacktriangle\varphi$, i.e., that $M, s \not\models \blacktriangle_n \varphi$ for some n. Let f be a group assignment function such that $f(s) \neq \emptyset$. We must show that there is a f-consistent path starting in s where φ is not true. We have that $M, s \models \neg\varphi \vee \bigwedge_{G_1 \in N^3} \bigvee_{i_1 \in G_1} \hat{K}_{i_1}(\neg\varphi \vee \bigwedge_{G_2 \in N^3} \bigvee_{i_2 \in G_2} \hat{K}_{i_2}(\neg\varphi \vee \cdots \vee \bigwedge_{G_n \in N^3} \bigvee_{i_n \in G_n} \hat{K}_{i_n} \neg\varphi))$. Thus, there is a $k \leq n$ such that

$$M, s \models \bigwedge_{G_1 \in N^3} \bigvee_{i_1 \in G_1} \hat{K}_{i_1} \bigwedge_{G_2 \in N^3} \bigvee_{i_2 \in G_2} \hat{K}_{i_2} \cdots \bigwedge_{G_k \in N^3} \bigvee_{i_k \in G_n} \hat{K}_{i_k} \neg\varphi.$$

Hence there is a path $s_0 s_1 \cdots s_{k+1}$ where: $s_0 = s$, $s_0 \sim_{i_1} s_1$ for some $i_1 \in f(s_0)$, $s_1 \sim_{i_2} s_2$ for some $i_2 \in f(s_1)$, $\cdots$, $s_k \sim_{i_{k+1}} s_{k+1}$ for some $i_{k+1} \in f(s_k)$, such that $M, s_{k+1} \models \neg\varphi$. This is an f-consistent path. □

We now proceed with defining the finite canonical model, and proving a truth lemma. We adapt the standard proof for common knowledge (see [8, Chapter 7][7]).

Definition 9. *The closure $cl(\gamma)$ of a formula γ is the smallest set that contains all subformulas of γ, is closed under single negations, and that contains $E_G \blacktriangle\varphi$ for all $G \subseteq N$ whenever it contains $\blacktriangle\varphi$.*

[7] There are also other complications, in addition to the lack of normality, including that the correspondent to the *induction axiom* for common knowledge, $\blacktriangle(\varphi \to \bigvee_{G \in N^3} E_G \varphi) \to (\varphi \to \blacktriangle\varphi)$, is not valid. We instead use a variant of the induction *rule* for common knowledge, used in, e.g., [10].

The closure of any formula is finite. As usual we say that a a set of formulas Γ is maximal consistent in $cl(\gamma)$ iff $\Gamma \subseteq cl(\gamma)$, Γ is consistent, and there is no consistent $\Gamma' \subseteq cl(\gamma)$ such that $\Gamma \subset \Gamma'$. When Δ is a finite set of formulas we write $\underline{\Delta}$ for $\bigwedge_{\delta \in \Delta} \delta$.

The canonical model is defined as follows.

Definition 10. *The canonical model for some formula γ is $M^\gamma = (S^\gamma, \sim^\gamma, V^\gamma)$ where: S^γ is the set of all sets of formulas maximal consistent in $cl(\gamma)$, $\Gamma \sim^\gamma \Delta$ iff $\{K_i\varphi : K_i\varphi \in \Gamma\} = \{K_i\varphi : K_i\varphi \in \Delta\}$, and $V^\gamma(p) = \{\Gamma \in S^\gamma : p \in \Gamma\}$.*

This definition of the canonical model is identical to the one used for common knowledge in, e.g., [8], except that the closure is wider. Many of the properties of that model, like deductive closure of maximal consistent sets in $cl(\gamma)$, carry over, and we only need to focus on the following property for the $\blacktriangle$ modality. A φ-path in the canonical model is a path $\Gamma_0\Gamma_1 \cdots$ where $\varphi \in \Gamma_i$ for every i.

Lemma 9. *For any γ and $\Gamma \in S^\gamma$ and $\blacktriangle\varphi \in cl(\gamma)$, $\blacktriangle\varphi \in \Gamma$ iff there is a group assignment function f such that $f(\Gamma) \neq \emptyset$ and every f-consistent path from Γ in M^γ is a φ-path.*

Proof. Left-to-right: let $\blacktriangle\varphi \in \Gamma$. Define f as follows: for any $\Delta \in S^\gamma$, $f(\Delta) = G$ if $E_G\blacktriangle\varphi \in \Delta$ for some $G \in N^3$ and $f(\Delta) = \emptyset$ otherwise. If there are several such G, chose one. Consider a Δ such that $f(\Delta) \neq \emptyset$ and $\Delta \sim_i^\gamma \Delta'$ for some $i \in f(\Delta)$. We must show that $f(\Delta') \neq \emptyset$. Since $K_i\blacktriangle\varphi \in \Delta$, $K_i\blacktriangle\varphi \in \Delta'$ by definition of $\sim_i^\gamma$, and thus $\blacktriangle\varphi \in \Delta'$ by the Truth axiom, and $E_{G'}\blacktriangle\varphi \in \Delta'$ for some G' by Mix and the fact that $E_{G'}\blacktriangle\varphi \in cl(\gamma)$. Finally, $f(\Gamma) \neq \emptyset$ follows from Mix and definition of the closure.

We must show that every f-consistent path from Γ is a φ-path. The proof is by induction on the length of the path. For the base case, we have that $\varphi \in \Gamma$ from the Mix and Truth axioms. For the induction step, consider an f-consistent path $\Gamma_0 \cdots \Gamma_n\Gamma_{n+1}$ where $\varphi \in \Gamma_j$ for all $j \leq n$. From f-consistency we have that $\Gamma_n \sim_i^\gamma \Gamma_{n+1}$ for some $i \in G$ for some $G \in N^3$ such that $E_G\blacktriangle\varphi \in \Gamma_n$. It follows from the definition of $\sim_i^\gamma$ that $K_i\blacktriangle\varphi \in \Gamma_{n+1}$, and thus that $\blacktriangle\varphi \in \Gamma_{n+1}$ from the Truth axiom, and thus that $\varphi \in \Gamma_{n+1}$ from the Mix axiom and Truth axiom again.

Right-to-left: let f be such that $f(\Gamma) \neq \emptyset$ and every f-consistent path is a φ-path. Let $S_{f,\varphi}$ be the set of all sets Δ maximal consistent in $cl(\gamma)$ such that $f(\Delta) \neq \emptyset$ and every f-consistent path from Δ is a φ-path. Let $\chi = \bigvee_{\Delta \in S_{f,\varphi}} \underline{\Delta}$. We first show that $\vdash \chi \rightarrow \bigvee_{G \in N^3} E_G\chi$ (*). Assume, towards a contradiction, that $\chi \wedge \neg \bigvee_{G \in N^3} E_G\chi$ is consistent. Then $\underline{\Delta} \wedge \neg \bigvee_{G \in N^3} E_G\chi$ is consistent for some $\Delta \in S_{f,\varphi}$. Then, for any $G \in N^3$ there must be a $i_G \in G$ such that $\underline{\Delta} \wedge \hat{K}_{i_G} \neg\chi$ is consistent. $\underline{\Delta} \wedge \hat{K}_{i_G} \vee_{\Theta \in S^\gamma \setminus S_{f,\varphi}} \underline{\Theta}$ is consistent for each i_G, and by modal reasoning for individual knowledge, $\underline{\Delta} \wedge \bigvee_{\Theta \in S^\gamma \setminus S_{f,\varphi}} \hat{K}_{i_G}\underline{\Theta}$ is consistent for each i_G. Thus, for every $G \in N^3$ there is an $i_G \in G$ and a $\Theta_{i_G} \in S^\gamma \setminus S_{f,\varphi}$ such that $\underline{\Delta} \wedge \hat{K}_{i_G}\underline{\Theta_{i_G}}$ is consistent. In particular, since $\Delta \in S_{f,\varphi}$ and thus $f(\Delta) \neq \emptyset$, there is an $i_G \in G = f(\Delta)$ and a $\Theta_{i_G} \in S^\gamma \setminus S_{f,\varphi}$ such that $\underline{\Delta} \wedge \hat{K}_{i_G}\underline{\Theta_{i_G}}$ is

consistent. It follows (by a standard property of the canonical model [8, Item 4 of Lemma 7.14]) that $\Delta \sim^{\gamma}_{i_G} \Theta_{i_G}$. Since $i_G \in f(\Delta)$, $f(\Theta_{i_G}) \neq \emptyset$ and since Θ_{i_G} is not in $S_{f,\varphi}$ that means that there is a an f-consistent path from Θ_{i_G} which is not a φ-path. It follows that there is an f-consistent path from Δ that is not a φ-path. But $\Delta \in S_{f,\varphi}$, which leads to a contradiction. Thus, we have shown (*).

From (*) and the induction rule, we get that $\vdash \chi \to \blacktriangle\chi$. Since Γ is one of the disjuncts in χ we have that $\vdash \underline{\Gamma} \to \chi$. Thus, $\vdash \underline{\Gamma} \to \blacktriangle\chi$. Since $\varphi \in \bigcap_{\Delta \in S_{f,\varphi}} \Delta$, we also have that $\vdash \chi \to \varphi$. By Monotonicity, $\vdash \blacktriangle\chi \to \blacktriangle\varphi$, and thus $\vdash \underline{\Gamma} \to \blacktriangle\varphi$. Since $\blacktriangle\varphi \in cl(\gamma)$, we have that $\blacktriangle\varphi \in \Gamma$. □

Lemma 10 (Truth). *For any γ, $\varphi \in cl(\gamma)$ and $\Gamma \in S^{\gamma}$, $\varphi \in \Gamma$ iff $M^{\gamma}, \Gamma \models \varphi$.*

Proof. The proof is by structural induction on φ. We only show the induction step for $\varphi = \blacktriangle\psi$, the other cases are straightforward and/or standard. We have that $M^{\gamma}, \Gamma \models \blacktriangle\psi$ iff (by Lemma 8) there is an f such that $f(\Gamma) \neq \emptyset$ and ψ is true on every f-consistent path starting in Γ iff (by the ind. hyp.) there is an f such that $f(\Gamma) \neq \emptyset$ and every f-consistent path starting in Γ is a ψ-path iff (by Lemma 9) $\blacktriangle\psi \in \Gamma$. □

We immediately get the following.

Theorem 1 (Completeness). S▲ *is complete.*

6 Discussion

Intentionally anonymous announcements $\varphi\ddagger$ capture exactly the announcements that are guaranteed to ensure anonymity. It all boils down to *safety*: intentional anonymous announcements $\varphi\ddagger$ are exactly public announcements of safety $\blacktriangle\varphi!$. Furthermore, the $\varphi\ddagger$ operators can be expressed in epistemic logic extended with only the safety operator $\blacktriangle$. We gave a complete axiomatisation of the latter.

There are still many open problems. While the expressivity results "reduce" the languages with $\varphi\ddagger$ to equivalent languages we have axiomatisations for, they don't directly give us "native" axiomatisations of those languages themselves.

A conceptually related work is van Ditmarsch's analysis of the *Russian Cards Problem* [7] which also models *safe announcements* – albeit with a different notion of safety, namely that a secret "card deal", instead of the identity of the announcer, is not revealed. A safe announcement of φ by a is captured by $[K_a\varphi \wedge [K_a\varphi]C\mathit{cignorant}!]$, where C is common knowledge among all agents and *cignorant* is the safety or secrecy condition – corresponding to only the announcer knowing the identity of the announcer in our case. That is again equivalent to the sequence $[K_a\varphi!][C\mathit{cignorant}!]$, which means that φ is safe to announce for a if $C\mathit{cignorant}$ holds after the announcement of $K_a\varphi$. Our $[\varphi\ddagger]$ operator also satisfies that definition of safety (Lemma 2) (although it cannot be expressed in the logical language).

Safety can be strengthened by using some variant of *distributed knowledge* [13, 30], so that announcements are safe even when non-announcing agents share

their knowledge to deduce the identity of the announcer. Another avenue of further research is to extend the presented formalisms with *quantification over anonymous announcements* like in [3, 1]. This will allow us to express properties like "there is a safe anonymous announcement, such that φ holds afterwards". Also of interest for future work, is to model *self-referential* intentionally anonymous announcements of the form "after this very announcement, no-one will know", as studied recently in [4, 16][8]. Finally, in this paper we have focussed on *knowledge* of agents. In the future, we would also like to explore anonymous announcements also in the setting of non-S5 modal logics that can capture, for example, *beliefs*, potentially false ones, of agents about the identity of the announcer. That would also enable relaxing the strict definition of the safety modality, e.g., similarly to the approximation of common knowledge in [19].

Acknowledgments We thank Hans van Ditmarsch and the LORI reviewers for detailed and constructive comments which has helped us improve the manuscript significantly.

References

1. Ågotnes, T., Balbiani, P., van Ditmarsch, H., Seban, P.: Group announcement logic. Journal of Applied Logic **8**(1), 62–81 (2010)
2. Ågotnes, T., Wáng, Y.N.: Somebody knows. In: KR. pp. 2–11 (2021)
3. Balbiani, P., Baltag, A., Van Ditmarsch, H., Herzig, A., Hoshi, T., De Lima, T.: Knowable as known after an announcement. The Review of Symbolic Logic **1**(3), 305–334 (2008)
4. Baltag, A., Bezhanishvili, N., Fernández-Duque, D.: The topology of surprise. In: KR (2022)
5. Baltag, A., Moss, L.S.: Logics for epistemic programs. Synthese **139**(2), 165–224 (2004)
6. Barthe, G., Gaboardi, M., Arias, E.J.G., Hsu, J., Kunz, C., Strub, P.: Proving differential privacy in hoare logic. In: CSF. pp. 411–424 (2014)
7. van Ditmarsch, H.: The russian cards problem. Stud Logica **75**, 31–62 (2003)
8. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer (2007)
9. Domingo-Ferrer, J., Sánchez, D., Soria-Comas, J.: Database Anonymization: Privacy Models, Data Utility, and Microaggregation-based Inter-model Connections (2016)
10. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: Reasoning About Knowledge. MIT Press (1995)
11. Halpern, J.Y., O'Neill, K.R.: Anonymity and information hiding in multiagent systems. J. Comput. Secur. **13**(3), 483–512 (2005)
12. Halpern, J.Y., O'Neill, K.R.: Secrecy in multiagent systems. TISSEC **12**(1), 5:1–5:47 (2008)
13. van der Hoek, W., van Linder, B., Meyer, J.J.: Group knowledge is not always distributed (neither is it always implicit). Math. Soc. Sci. **38**(2), 215–240 (1999)

[8] This comment also applies to the Russian Cards problem.

14. Hughes, D.J.D., Shmatikov, V.: Information hiding, anonymity and privacy: a modular approach. J. Comput. Secur. **12**(1), 3–36 (2004)
15. Jiang, J., Naumov, P.: De re/de dicto distinction: a logicians' perspective on data anonymity. J. Cybersecur. **11**(1) (2025)
16. Li, Y., Ren, J., Ågotnes, T.: The surprise exam in full modal fixed-point logic. In: Ågotnes, T., Doder, D. (eds.) Logic and Argumentation. pp. 185–200. Springer Nature Singapore, Singapore (2025)
17. Machanavajjhala, A., Kifer, D., Gehrke, J., Venkitasubramaniam, M.: *L*-diversity: Privacy beyond *k*-anonymity. TKDD **1**(1), 3 (2007)
18. Mano, K., Kawabe, Y., Sakurada, H., Tsukada, Y.: Role interchange for anonymity and privacy of voting. J. Log. Comput. **20**(6), 1251–1288 (2010)
19. Monderer, D., Samet, D.: Approximating common knowledge with common beliefs. Games and Economic Behavior **1**(2), 170–190 (1989)
20. More, S.M., Naumov, P.: Hypergraphs of multiparty secrets. Ann. Math. Artif. Intell. **62**(1-2), 79–101 (2011)
21. More, S.M., Naumov, P.: Logic of secrets in collaboration networks. Ann. Pure Appl. Log. **162**(12), 959–969 (2011)
22. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management (2010), http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf
23. Plaza, J.: Logics of public communications. In: ISMIS. pp. 201–216 (1989)
24. Schneider, S.A., Sidiropoulos, A.: CSP and anonymity. In: ESORICS. LNCS, vol. 1146, pp. 198–218 (1996)
25. Sweeney, L.: k-anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **10**(5), 557–570 (2002)
26. Syverson, P.F., Stubblebine, S.G.: Group principals and the formalization of anonymity. In: FM. LNCS, vol. 1708, pp. 814–833 (1999)
27. Tsukada, Y., Mano, K., Sakurada, H., Kawabe, Y.: Anonymity, privacy, onymity, and identity: A modal logic approach. In: CSE. pp. 42–51 (2009)
28. Tsukada, Y., Sakurada, H., Mano, K., Manabe, Y.: On compositional reasoning about anonymity and privacy in epistemic logic. Ann. Math. Artif. Intell. **78**(2), 101–129 (2016)
29. van Eijck, J., Orzan, S.: Epistemic verification of anonymity. ENTCS **168**, 159–174 (2007)
30. Wang, Y., Ågotnes, T.: Public announcement logic with distributed knowledge: expressivity, completeness and complexity. Synthese **190**(1), 135–162 (2013)
31. Xiong, Z., Ågotnes, T.: The logic of secrets and the interpolation rule. Ann. Math. Artif. Intell. **91**(4), 375–407 (2023)
32. Ye, X., Li, Z., Li, Y.: Capture inference attacks for k-anonymity with privacy inference logic. In: DASFAA. LNCS, vol. 4443, pp. 676–687 (2007)