Review of Analytic Philosophy
Vol. 5 No. 1 (2025)

Research article

Logics for Personalized Announcements and Attention Dynamics

Gaia Belardinelli, Lei Li, Sonja Smets, and Anthia Solaki

Abstract:
Access to information online usually undergoes double filtering: first, by the built-in algorithms of the platform in use, which select and propose information that is tailored to each user's preferences, and second, by the cognitive capacity and attention resources of the users themselves, which only allow the agent to receive a *portion* of the already personalized incoming information. Here, we introduce a framework based on dynamic epistemic logic, where announcements are personalized by semantically specified filtering conditions and are only received by the agents if their attention resources allow it. To achieve this, we first introduce opinion models to represent agents' opinions on topics and special edge-conditioned action models to more compactly represent the access of agents to personalized announcements. Then, we extend this framework to encompass attentive limitations as well. It is shown that the proposed logics are sound and complete with respect to the underlying classes of models.

Keywords:
Dynamic epistemic logic, Recommender systems, Attention dynamics, Opinion models, Personalized announcements

1. Introduction

A vast amount of information is created in social media platforms such as Facebook every day. To keep the platform experience engaging for the users, social media companies must find ways to curate such information and present users with content that captures their interest. One way to do this is by tailoring the information that users receive to the users' personal preferences,

Received November 15, 2023; accepted June 18, 2025; online published August 28, 2025; https://doi.org/10.18494/SAM.RAP.2025.0020

for example, by presenting content that matches their expressed opinions or their content-interaction history (Zhou *et al.*, 2012). While this personalized design achieves the goal of retaining users on the platform, it also has the effect of impacting users' beliefs and reasoning in ways that are not always desirable. In fact, social media companies have been accused of inducing reasoning biases such as confirmation bias, and of creating virtual environments that have distorting and isolating epistemic effects, such as filter bubbles, echo chambers, and more (Flaxman *et al.*, 2016). Aside from typical users, agents who use online platforms for profit are also clearly impacted by their design. Influencers on Instagram, for instance, had reason to worry when the platform tested hiding the number of 'likes' a post receives. Their popularity not only depends on users' pre-existing attitudes towards the content of a post, but also on parameters such as likes, views and shares, which affect the higher-order reasoning of users (Paul, 2019). The design choices underlying social media platforms thus play a central role in the way agents access the vast amount of information that can be found on such platforms, which in turn substantially impacts the beliefs they form.

However, even if the platform makes some content accessible to an agent, in reality, there is still a chance that the agent will not actually process it. Indeed, real agents may simply not have enough *attention* to read and keep track of all the information that is made available to them, as, even if filtered and curated, such information typically still comes in large amounts and keeps evolving dynamically (likes, visualizations, and resharings continue to increase, new posts are shared, and so on). The relevance of attention limitations in an information-rich world was first emphasized by Simon (1971). The feeds of online social networks nowadays exemplify this information-rich world and the overabundance of information we have to deal with. As attention is an intrinsically bounded resource, one cannot possibly process all the information popping up in one's feed, such as all the posts about other agents' opinions and preferences, and will inevitably pay attention to a mere subset. Attention thus acts as a second kind of filter impacting the reception and processing of information in virtual environments, and therefore, also the beliefs of agents.

Logical approaches have been extensively applied to the study of social networks, for example, to analyze the propagation of opinions or the dynamics of the network's structure (Baltag *et al.*, 2019; Smets and Velázquez-Quesada, 2017, 2018, 2019a,b, 2020; Solaki *et al.*, 2016; Young Pedersen *et al.*, 2021). While such approaches allow for a precise, fine grained analysis of information dynamics in highly complex environments such as social media platforms, they do not yet account for the *constraints* that interfere with agents' *access* to information, such as those imposed by the limitations of human cognitive resources or by the personalization mechanisms of social media platforms. Several of the existing models instead assume that agents have access to a far greater amount of information than real people actually do. For example, agents are considered to know all the preferences or opinions of all their social media friends. This is clearly not the case in reality, where a user may only have partial or even incorrect information about other users' features.

In our setting, we focus on doxastic logics and on two types of limitations: (i) structural filtering of information, imposed by social media design, and (ii) inherent cognitive bounds of human reasoners, mainly driven by their limited and selective attention.

Structural filtering. The necessity of maintaining browsing interesting and engaging for the users, given the large amount of user-generated content, has fostered the introduction of *recommender systems* in sorting which posts appear to the users. Recommender systems rely heavily on *personalization*, that is, the subset of posts that will end up appearing on an agent's feed are the ones deemed sufficiently compatible with the preference profile built for this agent (Jannach *et al.*, 2010). For example, *user-based collaborative filtering* recommends posts by finding other users with similar profiles, whereas *item-based collaborative filtering* suggests posts based on similarities among the posts themselves. *Content-based filtering*, in turn, recommends posts with features the user has interacted favorably with in the past. Even in very compact networks, each of these personalization mechanisms can result in agents becoming exposed to radically different feeds, a practice that has been blamed for the increasing polarization in social media (Vīķe-Freiberga *et al.*, 2013).

Attention-driven filtering. While the first type of filtering determines what information is, in principle, accessible to an agent, this second type pertains to the information that an agent *actually* accesses. As attention is limited, not all information on the news feed will actually be read and processed by an agent. For that to happen, the agent must be paying attention to it, that is, she must have enough cognitive resources (attention span) to actually process it. Each post in social media may be seen as having a cognitive cost attached to it, which an agent must pay if she wants to learn from it (Kahneman, 1973). If this requirement fails, the post is not processed and the agent's beliefs remains unchanged.

These observations have implications for the standard way of modelling epistemic phenomena based on the use of S5-modal logic. Such logics have been criticized for their limited potential to bridge formal modeling and the cognitive lives of human agents (Stalnaker, 1991; Solaki, 2021; van Benthem, 2008; Verbrugge, 2009). In line with this, we make only minimal assumptions on the information states of agents in this paper and model the information flow as it is, i.e., as the outcome of dynamic message-posting actions that are sensitive to the aforementioned types of limitations. We model such scenarios in terms of update methods coming from Dynamic Epistemic Logic (DEL) (Baltag *et al.*, 1998; Baltag and Renne, 2016; van Benthem, 2011; van Ditmarsch *et al.*, 2008), which can deal with knowledge and belief dynamics triggered by new incoming public or private information. To do so, our approach relies on (a) the format of the so-called edge-conditioned event models in Bolander (2018), an extension of DEL where filtering limitations are treated as edge conditions[1] on events that are partially observed by agents, and (b) an attention-budget function, suitably updated in the aftermath of informational actions (similar to the work using such functions in the context of epistemic planning (Belardinelli and Rendsvig, 2021) and in the context of modeling the cognitive cost for memory use (Solaki, 2021). This approach can potentially yield further insights into the emergence of various phenomena, such as polarization in online networks, thus providing valuable guidance to the design of social networks and our adopted practices therein.

While logic-based methods (e.g., fuzzy logics) have been applied in the context of recommender systems (Jain and Gupta, 2018), the use of modal logics to represent the underlying filtering conditions as well as their effect, is, as far as we know, a new development. The work by Liu *et al.* (2024) is an example of recent work in this direction that focuses particularly on content-based filtering and adopts a mixture of models with a Kripke semantics and neighborhood semantics. While we similarly use modal logic, we specifically adopt the tools from Dynamic Epistemic Logics. In this context, we work with edge-conditioned event models that are particularly well fitted to model filtering constraints on events. Formally, we apply this setting to model a new type of scenario involving the kind of filtering used by recommender systems, as well as the filtering that may occur as a result of an agent's own limited attention span. We also study the effect that this filtering can have on the beliefs of agents. Note that in this paper only recommended items that are considered to be 'acceptable' (i.e., understood as being 'coherent with an agent's belief system') can trigger a belief change when the recommender system actually sends out the message. Recommended items that do not fit well with an agent's belief system will not lead to a belief change.

An unpublished previous version of this joint work appeared in the PhD thesis Li (2023).

1 Work in Li (2023); Belardinelli *et al.* (2025) shows that edge-conditioned event (or action) models are as expressive as standard event models in DEL but offer a more succinct representation of specific event types. This is done by explicitly imposing propositions that act as conditions directly on the accessibility relations between events.

Outline. The paper proceeds as follows. We first focus on a formal framework centered around the structural filtering limitations (Section 2), explaining how we embed the dynamics of personalized announcements in doxastic-logical models. Next, we provide a sound and complete axiomatization for this framework (Section 3). In Section 4, we enrich the framework with additional machinery to represent also the attentive limitations of agents, and provide a sound and complete axiomatization. We conclude with a discussion and directions for future work in Section 5.

2. Logics for Personalization in Social Platforms

In what follows, we build a doxastic logical framework suitable to represent the dynamics of personalization in an online social platform (i.e., the structural filtering discussed above). The goal is to model the effects of these personalization mechanisms on the information flow and thus ultimately on agents' beliefs. We first present the static basis of our account, which we call *Static Logics of Personalized Announcements* ($SLPA_i$, where the index i denotes the specific personalization (or structural filtering) method adopted, as introduced below). Formally, a filtering method is going to be treated as a parameter of our class of models, and used to decide which announcements (i.e., which social media posts) an agent receives.[2] Subsequently, we proceed with the development of a fully dynamic logic framework that captures belief updates due to the structural filtering methods of social platforms.

2.1 Static logics of personalized announcements

We begin by introducing the models of the static logics for personalized announcements SLPA_i, and later provide the syntax. Let $Ag \neq \emptyset$ be a finite set of agents and let $T \neq \emptyset$ be a finite set of topics about which the agents can adopt a position, i.e., they can express an opinion such as a posted 'like' on their social media platform. For this part of our logic, we follow (Smets and Velázquez-Quesada, 2019a) in denoting agents' positions on topics (i.e., expressed opinions such as posted 'likes') using $\{R_t\}_{t\in T}$ as a pairwise disjoint collection providing a finite non-empty set R_t of positions for each $t \in T$. We denote by R the set $\cup_{t\in T} R_t$.

2 Note that, since we introduce more than one filtering method, each of which can be used as a parameter of our models, there will be several corresponding static logics SLPA_i, one for each method i, all based on the same language.

Definition 1 (Opinion Model). *Given Ag, T, and $\{R_t\}_{t\in T}$, an opinion model is a tuple $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j\in Ag}, \mathbf{V}\rangle$ where*

- $\mathbf{W}$ *is a non-empty set of possible worlds;*
- $\rightarrow_j$ *is the doxastic accessibility relation for agent $j \in Ag$. This relation is taken to be serial, i.e., for any $w \in \mathbf{W}$, there exists $u \in \mathbf{W}$ such that $w \rightarrow_j u$ (Consistency of Belief);*
- $\mathbf{V}$: $W \times Ag \times T \rightarrow \mathcal{P}(R)$ *is the valuation function, where $\mathbf{V}(w, i, t) \subseteq R_t$ indicates the positions (i.e., expressed opinions) of agent i on topic $t \in T$ at world w.*

We additionally require that an agent's valuation about her own expressed opinions does not change across the doxastically accessible worlds in opinion models. This is captured via the following Opinion-Belief Condition, which is imposed in this semantics to model agents that have correct and true beliefs about their own posted 'likes' and expressed opinions on their social media platform (i.e., beliefs about propositions that are true at the world of evaluation). Intuitively, this translates in the assumption that agents cannot be mistaken about their own posts, or, in other words, that if an agent has expressed an opinion, the agent believes she has done so. Note that no such constraint is imposed in these models on the beliefs of agents about the posted 'likes' of others. It is well possible in these models for agents to have false beliefs or to suspend their beliefs about the expressed beliefs of others.

- If $w \rightarrow_i u$, then $\mathbf{V}(w, i, t) = \mathbf{V}(u, i, t)$, for any $t \in T$, $i \in Ag$ (*Opinion-Belief Condition*).

The two model conditions, *Consistency of Belief* and *Opinion-Belief*, represent the two minimal assumptions we make concerning the agents' reasoning about opinions. We refrain from imposing additional constraints which are typical in epistemic-doxastic logics, such as introspection properties, given that the introspective capacity of an agent may actually be rather limited in scenarios of imperfect agents. For example, agents with bounded resources, such as memory or time, might not be *fully* introspective and their introspective abilities might be limited to a certain nesting level (van Ditmarsch and Labuschagne, 2007), while studies on implicit cognition reveal even more fundamental issues with such assumptions (Litman and Reber, 2005; Schwitzgebel, 2010).[3]

[3] In other words, we drop full introspective abilities because the infinite arrays of positive and negative introspection presuppose infinite reasoning capacity, which is clearly unattainable for human agents with finite cognitive resources. This is consistent with our focus on attentional limitations (Section 4) and the aforementioned empirical evidence. Consistency of belief has been criticized, too. Inspecting one's belief set for consistency can be a difficult, and potentially intractable task, that also depends on how many beliefs one holds (Cherniak, 1986), which, for the specific context of an online social

As mentioned above, we note that the class of opinion models will come to be parametrized by a filtering method, which we introduce below.

Example 1. *Consider the following positions on three topics about which Alice (A), John (J), and Helen (H) may express a specific opinion (i.e. a 'like,' an interest or preference) on a social media platform.*

- *Cuisine (C): Chinese (c), French (f), Italian (i), Korean (k), Mexican (m);*
- *Dance (D): Ballet (b), Jazz (j), Hip-hop (h), Tango (t);*
- *Game (G): Action (a), Puzzle (p), Racing (r), Strategy (s).*

Throughout the paper, we represent a position on a topic (i.e., elements of R_t for $t \in T$) using the lowercase initial of the selected option and upper case initial of the considered topic. It is assumed that Alice, John, and Helen form a group of users, and they do not necessarily know each others' opinions on these topics, but they form beliefs about them based on what others have posted in the past. In Fig. 1, there are three worlds, w, u, and v. The relation $w \xrightarrow{A} u$ denotes that world u is doxastically accessible from world w for Alice, i.e., Alice entertains the distribution of opinions depicted in world u, and the same goes for the other labelled arrows.

The model $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j \in Ag}, \mathbf{V} \rangle$ in Fig. 1 is an opinion model, where $Ag = \{A, J, H\}$, $\mathbf{W} = \{w, u, v\}$, and where the accessibility relations $\{\rightarrow_j\}_{j \in Ag}$ and the valuation function $\mathbf{V}$ are as shown in Fig. 1.

Definition 2. *Let Φ be the set of atomic formulas i_t^r. The language $\mathcal{L}$ of the class of static logics of personalized announcements is given by:*

$$\phi ::= i_t^r \mid \neg\phi \mid \phi \wedge \phi \mid \mathsf{F}(j, i_t^r) \mid B_j \phi$$

where $i, j \in Ag$ and $r \in R_t$, for $t \in T$.

The Boolean operators $\rightarrow$, $\vee$ and $\leftrightarrow$ are defined as usual. For $j \in Ag$, the dual operator $\hat{B}_j$ of the belief operator B_j is defined as follows $\hat{B}_j \phi := \neg B_j \neg \phi$. Intuitively, the atomic formula i_t^r can be read as follows: "agent i expressed opinion r on topic t." We introduce the operator

network could vary. We acknowledge that the current modal-logic framework does not entirely do justice to boundedly rational agents, as it does not take into account the cognitive effort for computing all the consequences of one's own beliefs nor for the effort of keeping one's beliefs consistent. Still, we accept consistency of belief as a minimal assumption to proceed with personalization and attention dynamics and because lifting all idealizing assumptions is independently treated by different logical systems addressing the problem of logical omniscience (see e.g., Halpern and Pucella (2011)).

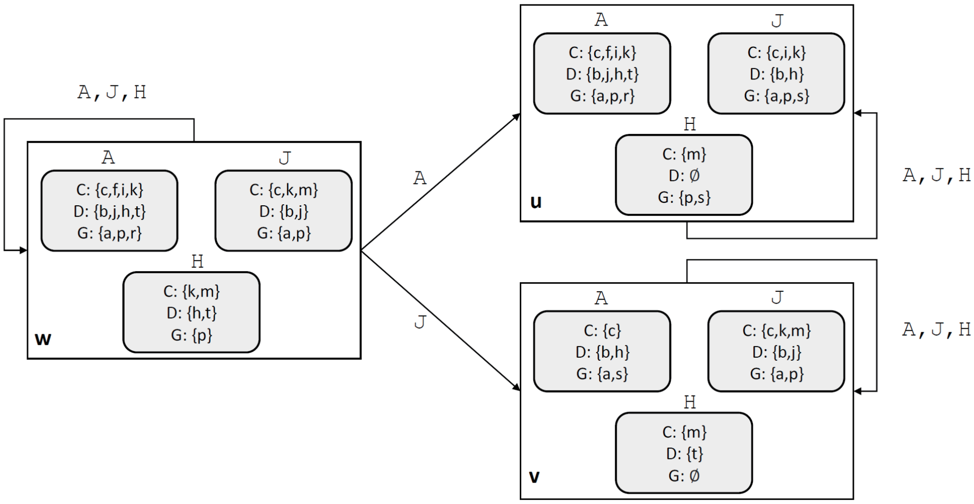

Fig. 1. Opinion model $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j \in Ag}, \mathbf{V} \rangle$. Worlds are represented by rectangles with sharp corners. The opinions expressed (or 'likes' posted) at each world by each agent *A*, *J*, and *H* on the topics *C*, *D*, and *G* are represented inside rectangles with rounded corners inside each world. For example, at *w*, Alice *A* has expressed opinions *c*, *f*, *i*, *k* about topic *C*, opinions *b*, *j*, *h*, *t* about topic *D*, and opinions *a*, *p*, *r* about topic *G*. The arrows indicate doxastic accessibility relations. For example, if *w* is taken to be the actual world, Helen's beliefs about the other agents' opinions happen to be correct.

$\mathsf{F}(j, i_t^r)$ to express that the post i_t^r by the recommender system *fits well* with the prior beliefs of agent j, i.e., agent j considers the recommended post i_t^r to be 'acceptable' (or 'credible'). This idea of expressing that a proposition 'fits well' with an agent's belief system can be viewed in terms of how well a proposition 'coheres' with an agent's set of beliefs.[4] To evaluate the 'acceptability' of the recommended post i_t^r for agent j (with $j \neq i$), i.e., $\mathsf{F}(j, i_t^r)$, we will analyze both whether j's beliefs and i_t^r fit together and check if i_t^r can pass the filtering constraints imposed by the recommender system. We here assume that a recommender system works on the basis of an algorithm that has access to all expressed opinions of all the users of the platform. So to recap, the guiding principle is that the truth value of $\mathsf{F}(j, i_t^r)$ rests upon whether agent j's beliefs are deemed sufficiently "compatible" with the post i_t^r that the recommender system of the platform in question can offer, i.e., whether it passes the filtering process and j would also accept the post i_t^r if it were to appear in her feed.

4 While the notion of 'coherence of a set of propositions' is here only used to stress its property of being logically consistent, in formal epistemology, different additional requirements are typically added to it (see Olsson (2023)).

Definition 3 (Semantic clauses). *Given an opinion model* $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j \in Ag}, \mathbf{V} \rangle$ *and world* $w \in \mathbf{W}$

- $\mathbf{M}, w \vDash i^r_t$ *iff* $r \in \mathbf{V}(w, i, t)$*;*
- $\mathbf{M}, w \vDash \neg\phi$ *iff* $\mathbf{M}, w \nvDash \phi$*;*
- $\mathbf{M}, w \vDash \phi \wedge \psi$ *iff* $\mathbf{M}, w \vDash \phi$ *and* $\mathbf{M}, w \vDash \psi$*;*
- $\mathbf{M}, w \vDash B_i\phi$ *iff* $\mathbf{M}, u \vDash \phi$ *for all* $u \in \mathbf{W}$ *such that* $w \rightarrow_i u$*;*
- $\mathbf{M}, w \vDash \mathsf{F}(j, i^r_t)$ *iff (a)* $\mathbf{M}, w \vDash i^r_t$*, (b) there is a* $v \in \mathbf{W}$ *with* $w \rightarrow_j v$ *such that* $\mathbf{M}, v \vDash i^r_t$*, and (c) if* $j \neq i$ *then the* ***filtering condition*** *(as specified below) holds at w.*

The evaluation of an agent's expressed opinion i^r_t is treated as the evaluation of an atomic sentence, and the evaluation of the Boolean operators is standard in modal logic. Beliefs are here modelled as Kripke modalities, one for each agent, and the truth-value assignments use the doxastic accessibility relations. The evaluation of $\mathsf{F}(j, i^r_t)$ is based on three ingredients. First, we require that the post i^r_t is true at the world of evaluation. Intuitively, we think about recommender systems that can push an already expressed item by agent i onto the feed of another agent j.[5] Second, the post i^r_t needs to be acceptable to j. This means that the post has to be compatible with j's prior beliefs, in the sense of $\hat{B}_j i^r_t$.[6] Note that by the Opinion-Belief condition, we have that an agent's own posts or expressed opinions are always acceptable to her, implying that $\mathsf{F}(i, i^r_t)$ is always true for such opinions. Third, for $i \neq j$, the evaluation depends on the filtering condition that the online social platform adopted. We discuss next four variants of filtering conditions, namely the *Radical Item-based Push Condition*, the *Conservative Item-based Push Condition*, the *User-Based Push Condition*, and the *Feature Push Condition*. In reality, personalized information distribution is often determined by machine learning methods. Our filtering conditions are abstractions that will

5 Note however that this does not mean that the posts themselves reveal true factual information about a topic, as i^r_t only captures an expressed opinion by an agent. Including a fact-checking constraint that refers to the truth-value of a post would require a further constraint that can be added to the filtering conditions in this paper.

6 The notion of acceptability in this paper ties in with a conservative belief-expansion strategy for an agent. What is acceptable to agents in this paper cannot contradict their prior beliefs. Different belief-revision strategies will lead to different notions of acceptability and those can alternatively be considered. We comment more on this point in the conclusion.

The filtering conditions are inspired by methods used in recommender systems: itembased, user-based, and content-based filtering (Aggarwal, 2016; Jannach *et al.*, 2010); the underlying ideas behind those are summarized below. The target agent is depicted in red, while thick lines denote "liking" and dashed lines denote "recommending." Shapes are used as placeholders of the "items" that can be liked by the agents. The shaded areas indicate the type of similarity used to generate the recommendation, as far as collaborative methods are concerned.

allow us to embed the underlying idea behind popular filtering methods in a doxastic-logical framework and thus offer a meta-analytical study of how such methods shape the informational dynamics and affect diffusion phenomena.

The filtering conditions. We now elucidate the filtering condition for agent j concerning agent i's opinion r on topic t (i.e., i^r_t), where $j \neq i$.

- **Radical Item-based Push Condition.** This filtering condition is inspired by the so-called *item-based collaborative filtering* method of recommender systems (Fig. 2, (Jannach *et al*., 2010, Ch. 2)). It relies on a similarity metric between *items* (which, in this paper, are agents' posts), calculated in terms of how they were interacted with by all agents. The method then predicts that the posts that should be recommended to the target agent are the ones that are similar to those the agent has interacted positively with. One everyday example of an application of this, encountered in many social media, can be summarized in "many users who liked that item, like you, have also liked this other item." Given $\theta \in \mathbb{N}$ as a threshold, it represents the minimum level of consensus required to trigger the recommendation of a post. Specifically, if the number of individuals within a group who hold a certain opinion and like a specific item exceeds the given threshold, then that item is expected to be acceptable to the entire group.
  In this case, the filtering condition between agent j and the target formula i^r_t is satisfied at the world of evaluation w, whenever for some opinion r' about topic t held by agent j, there are at least θ agents agreeing on their positions on both r and r'. Therefore, since agent j expressed the opinion r' about t, j should also tend to hold opinion r, and so a message about position r is "pushed" to agent j accordingly. In formal terms, there exists $r' \in \mathbf{V}(w, j, t)$ such that

$$|\{k \in Ag \mid r' \in \mathbf{V}(w, k, t) \leftrightarrow r \in \mathbf{V}(w, k, t)\}| \geq \theta$$

- **Conservative Item-based Push Condition.** This filtering condition follows the same principle as the previous one but it encodes a more stringent interpretation of it. It requires that, for *all* opinions r' of agent j about topic t, there is a sufficient number of agents (above some threshold θ) that have opinion r if and only if they have opinion r'. Under this requirement, there should be more reasons to believe that opinion r could also be a viewpoint that agent j is inclined to hold and would like to be pushed to them. Stated formally, for all $r' \in \mathbf{V}(w, j, t)$ such that

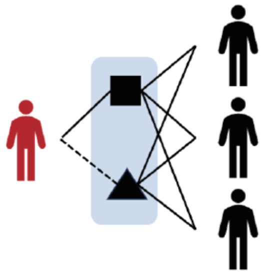

Fig. 2 The underlying idea behind item-based (collaborative) filtering: when a sufficient number of agents like two items and one item is liked by the target agent, the other item will also be recommended to her. That is, filtering is based on the similarity of items.

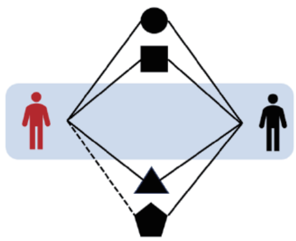

Fig. 3 The underlying idea behind user-based (collaborative) filtering: when agents share the liking of a sufficient number of items and one item is liked by one agent, it will be recommended to the other. That is, filtering is based on the similarity of ***users***.

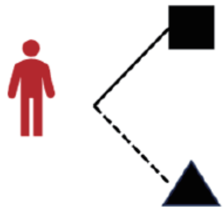

Fig. 4. The underlying idea behind content-based filtering: when an agent likes a certain item and another item has similar ***content***, i.e., shares a feature with it (here depicted as having the same color), it will be recommended to her.

$$|\{k \in Ag \mid r' \in \mathbf{V}(w, k, t) \leftrightarrow r \in \mathbf{V}(w, k, t)\}| \geq \theta$$

- **User-based Push Condition.** This filtering condition is inspired by the so-called *user-based collaborative filtering* method of recommender systems (Fig. 3, (Jannach *et al*., 2010, Ch. 2)). It relies on similarity metrics between *agents*, calculated in terms of their stance towards several "items." This method recommends a post to a target user if agents similar to her have interacted positively with it. One of the most common applications of this method can be found in online shopping outlets, where we often encounter the remark "other customers like you have purchased... ." To specify a semantic condition inspired by this, we will say that the filtering condition between agent j and the target i_t^r is satisfied at the world of evaluation w, whenever i and j have a sufficient number of opinions regarding topic t in common, i.e., for a given threshold $\theta \in \mathbb{N}$

$$|\mathbf{V}(w, i, t) \cap \mathbf{V}(w, j, t)| \geq \theta$$

- **Feature Push Condition.** This filtering condition is inspired by the so-called *content-based filtering* method of recommender systems (Fig. 4, (Jannach *et al*., 2010, Ch. 3)). It relies on the features of the items the target user has interacted favourably with. It then simply recommends posts that share these features.

In this case, the filtering condition between agent j and the target i_t^r is satisfied at the world of evaluation w, whenever agent j does have opinions on topic t. In formal terms,

$$\mathbf{V}(w, j, t) \neq \emptyset$$

Example 2. *In Example 1, fix a threshold $\theta = 2$.*

- *Suppose we are interested in whether John's expressed opinion c on topic C would be acceptable for Helen if it were recommended to her, i.e., does it pass the filter for Helen and is it compatible with her beliefs at world w? To answer this question, we need to check the truth value of the formula $\mathsf{F}(H, J_C^c)$ at w. Using the Radical Item-based Push Condition: given that at w, Helen holds opinion k on topic C, and Alice and John hold opinions c and k, the condition is satisfied at w. Moreover, $\mathbf{M}, w \vDash \hat{B}_H J_C^c$. Overall, this means that $\mathbf{M}, w \vDash \mathsf{F}(H, J_C^c)$ and therefore, it can be expected that the post from John that he holds opinion c on C would be accepted by Helen if it appeared in her feed at world w.*
- *Suppose we are interested in the same question but now use the Conservative Item-based Push Condition: given that Helen expressed opinions k and m on topic C and that both John and Alice did not express opinions m and k together, the condition is not satisfied at w. Hence, $\mathbf{M}, w \nvDash \mathsf{F}(H, J_C^c)$, because even if John's post J_C^c holds in w and is also compatible with Helen's beliefs, it is not a post that will appear in her feed as it does not pass the filtering condition of the recommender system at world w.*
- *Suppose we are interested in the same question, but now use the User-based Push Condition: given that Helen shares opinions k and m with John on topic C, the two agents are deemed sufficiently similar. Hence together with the fact that the post is compatible with Helen's beliefs, we have $\mathbf{M}, w \vDash \mathsf{F}(H, J_C^c)$.*
- *Suppose we are interested in the same question, but now use the Feature Push Condition. Since Helen has an interest in topic C at w ($\mathbf{V}(w, H, C) \neq \emptyset$), and the post is compatible with her beliefs, the condition is satisfied at w. As a result, $\mathbf{M}, w \vDash \mathsf{F}(H, J_C^c)$.*

  On the other hand, if the actual world was u, Helen would have no interest in topic D, formally $V(u, H, D) = \emptyset$. Thus the Feature Push Condition for any post i^r coming from another agent is not satisfied at u for Helen, no matter which other agent $i \neq H \in Ag$ posts on topic D we consider. As a result, $\mathbf{M}, u \nvDash \mathsf{F}(H, i_D^r)$, for $i \in \{A, J\}$, $r \in R_D$.

The recommendation rules are indeed critical, as they shape the agents' access to information. In the above example, Helen did not adopt position c on topic C at w, but if the social media platform had used the *Radical Item-based Push Condition* instead and Helen received the post in her feed, the post could have had an impact on her beliefs. If the social media platform had used the *Conservative* variant instead, then Helen would not have received this post from the recommender system. Hence, depending on which filtering condition is used, an agent may or may not be given access to a post that another agent shared. For instance, the conservative version of the Item-Based Push Condition, which is much more restrictive than its radical counterpart, will allow fewer posts to circulate in the platform and may thus cause agents to miss posts that are relevant or interesting to them. It is therefore clear that the information dynamics is highly sensitive to the structural choices underlying the design of a social network. Note that many other filtering conditions can be envisioned. However, for the purpose of this study, we here only consider an illustrative selection of such conditions and their implications for the belief dynamics (below).[7]

2.2 Dynamic logic of personalized announcements

In what follows, we extend the basic static setting with the dynamics of information flow in online social platforms and propose the Dynamic Logic of Personalized Announcements (DLPA_i, where i is the personalization or filtering condition adopted).[8] To present this logic we first look at the semantics.

First, we propose *access action models* to describe events where agents learn new posts from their social media feed. More specifically, we utilize an action model to represent the act of learning personalized announcements of the type i_t^r, which is subject to the agent's acceptability of the recommended post. Moreover, the belief states of agents are updated through these personalized announcements, which is reflected in the dynamics of our models defined later. In order to refine standard action models for public announcements so that they represent personalization phenomena, we define access action models inspired by the edge-conditioned models in Bolander (2018), albeit adapted to our setting. It is worth mentioning that, instead of using edge-conditioned action models, we could have used other

7 Besides other filtering conditions, we can also envision variations of the ones proposed in this paper. For instance, as the above 'item-based' conditions make essential use of a specific similarity measure between agents, one could use different measures and/or add constraints in terms of which 'expressed opinions' should hold for similar agents.

8 Similarly to SLPA_i, there will be several dynamic logics DLPA_i, one for each corresponding filtering condition i, all based on the same dynamic language (introduced below).

frameworks such as Baltag and Moss (2004); van Benthem (2011); Bjorndahl and Nalls (2021); Engesser *et al*. (2021).

Definition 4 (Access action model). *An access action model is a tuple* $\mathbf{C} := \langle \mathbf{E}, \{\mathbf{f_j}\}_{j \in Ag}, \mathbf{pre} \rangle$*, where*

- $\mathbf{E}$ *is a non-empty set of events;*
- $\mathbf{f_j} : \mathbf{E} \times \mathbf{E} \rightarrow \mathcal{L}$ *is the 'acceptability and filtering' function, assigning a condition to each pair of events for each agent j;*
- $\mathbf{pre} : \mathbf{E} \rightarrow \mathcal{L}$ *is the precondition function.*

Next, we propose specialized access models, i.e, the *models for personalized announcements*, for scenarios in which agents learn the recommended post i_t^r that passes the 'acceptability and filtering' condition. Otherwise, they will not gain any information about it.

Definition 5 (Action model for the personalized announcement of i_t^r). *The model for the personalized announcement of* i_t^r *is a tuple* $\mathbf{C} := \langle \mathbf{E}, \{\mathbf{f_j}\}_{j \in Ag}, \mathbf{pre} \rangle$*, where*

- $\mathbf{E} = \{e_0, e_1\}$ *is a set of events where event* e_1 *indicates the act of learning a recommended post* i_t^r*, while event* e_0 *indicates that no learning happens and hence no new information is being processed.*
- *The 'acceptability and filtering' function* $\mathbf{f_j}$: $E \times E \rightarrow \mathcal{L}$ *is defined as follows. It ensures that agent i always accepts her own post* i_t^r*, while others learn the announcement if it is acceptable to them and the platform does not filter out the post for them.*

$$-for\ j = i,\ \ \mathbf{f_j}(e,e') = \begin{cases} \top, & if\ \ e = e' = e_1 \\ \top, & if\ \ e = e' = e_0 \\ \bot, & otherwise \end{cases}$$

$$-for\ j \neq i,\ \ \mathbf{f_j}(e,e') = \begin{cases} \mathsf{F}(j, i_t^r), & if\ \ e = e' = e_1 \\ \neg\mathsf{F}(j, i_t^r), & if\ \ e = e_1\ \ and\ \ e' = e_0 \\ \bot, & if\ \ e = e_0\ \ and\ \ e' = e_1 \\ \top, & if\ \ e = e' = e_0 \end{cases}$$

- $\mathbf{pre}$ *is the precondition function,* $\mathbf{pre}(e) = \begin{cases} i_t^r, & if\ \ e = e_1 \\ \top, & if\ \ e = e_0 \end{cases}$

The events that are part of these event models are only of two possible types: either an agent learns the recommended post i_t^r (event e_1) or they learn nothing (event e_0). While event e_0 can always occur, event e_1 can only occur if i had previously expressed the position i_t^r (as its pre-condition contains i_t^r). These event models also specify which agents consider which event is possible; this is done in the above definition by making use of the function f_j. These 'acceptability and filtering' functions capture the fact that agent i can fully distinguish between the event e_1 in which she processes her own post i_t^r (which is always acceptable to her) and the event e_0 in which she does not learn anything. Besides agents who are recommended their own post, also other receiving agents ($j \neq i$) who pass the 'acceptability and filtering' condition for i's post - i.e., agents j for which $\mathsf{F}(j, i_t^r)$ holds – have access to event e_1 and can distinguish it from e_0. However, all other agents who do not learn i's post - i.e., the agents who do not pass the 'acceptability and filtering' condition namely the agents for which $\neg\mathsf{F}(j, i_t^r)$ holds – will only entertain event e_0.

As such, these event models come with a number of doxastic assumptions, similar to those of (misleading) private announcements in DEL (Baltag and Renne (2016)). Let us unpack these assumptions to understand how our framework impacts different types of agents' access to post contents, as well as their access to higher-order information about each other's access to these contents. Agents who do not receive a post because it *does not pass the filtering condition* and so do not learn anything about the post are treated as agents who completely ignore its contents. The agents who receive the post (because it passes the filtering condition) and accept it (because it is compatible with their beliefs) are aware that agents who cannot accept the post will ignore it (i.e., receiving a non-acceptable message is here equivalent to not receiving it at all). Instead, the agents who do not learn anything do not consider it any *more* likely that others could be receiving and accepting the post. These assumptions can be seen as simplifying assumptions to model a particular type of personalized announcements, e.g., scenarios where the agents belong to a specific (sub-)network, such as a music or poetry club, where they express positions on specific topics and receive exclusive posts that are targeted to club members only and non club members cannot process the posts. In such scenarios, members will know that non-members cannot process these posts.

Although this two-event construction in which non-receivers solely entertain event e_0 (*where nothing new is learned*) may suggest that non-receivers are entirely blocked from considering the possibility that others (come to) believe the content of a post, this is not the case, as we shall see. Our event model construction is motivated as follows. In an online social platform, posts are constantly exchanged, which is the primary function of the platform. In other words, the 'default' in our framework always is that *some* post may be made, but any more enhanced information about it (over its contents and acceptability) is mediated by the recommended system, which remains opaque to the ordinary user. In principle, in an event, a user may implicitly entertain *all* combinations of agent-opinions and higher-order beliefs thereof. However, there is no *special* learning (i.e., beyond the trivial fact that *some* post may have been made – which we consider covered by event e_0) conveyed to non-receivers by means of a personalized announcement. In essence, it is compatible with our setting that, whenever something is posted, a higher-order belief about it is formed. However, it comes as a feature of the initial doxastic situation of agents and not as something that is 'learned' by non-receivers via a personalized announcement.[9]

This can be evinced by our example. Initially, in Example 1, agent J considers possible that H_D^h (i.e., $\hat{B}_{\mathsf{J}} H_D^h$ holds), as well as that agent A considers that possible (i.e., $\hat{B}_{\mathsf{J}} \hat{B}_{\mathsf{A}} H_D^h$ holds). As we shall see, these carry over to the updated model (Example 3). That is, agent J initially entertains these possibilities and *does not cease to entertain them*, following the personalized announcement. So overall, this is not precluded from our framework, but it comes as a feature of the doxastic situation of the agents and does not have to be 'forced' by an explicit representation in the event model of how non-receivers may entertain possibilities on the receivers' situation.[10]

9 This design choice can be aligned with studies in the social sciences showing that users of social media can show general unawareness about the technical facets of their social media platforms (for an overview, we refer to the background section in Proferes (2017)). Translated into our models, such agents can be unaware of the real event that happens, be it the fact that a message was posted or that a post was filtered out of their personal feed by the recommender system. More specifically, an initial study in Eslami *et al*. (2015) on user's algorithmic awareness regarding Facebook revealed the surprising result that the majority of the participants in the conducted study were not aware of the algorithm's existence. When the unaware participants were confronted with an unfiltered alternative of their incoming feeds, 'many of the Unaware participants (n = 15) were initially very surprised by how long the 'All Stories' column was in comparison to the 'Shown Stories' column [. . .].' Moreover, Eslami *et al*. (2015) reports that 'Observing the algorithm outputs in FeedVis surprised some Unaware participants (n = 11) by revealing misperceptions about their friends whose stories were not shown in the participants' News Feed at all.' These users' initial false beliefs were attributed to them because some believed that 'those friends simply did not post on Facebook,' some 'falsely believed that those friends had left Facebook' and some participants 'mistakenly believed that their friends intentionally chose not to show them stories because they were not interpersonally close enough' (Eslami *et al*. (2015)).

10 Note how our syntax is built on atomic propositions that give information about specific topics and agent preferences. Then, granting access to events with non-trivial preconditions to agents who do not pass the needed conditions risks endowing them with too much information, in any case more than they would be justified to have (i.e., more than 'some post may have been made'). The matching event models for these explicit representations are variations of semi-private announcements (see Baltag and Renne (2016)) such that the corresponding edge-conditioned event models contain an 'acceptability and filtering' function that ensures that for $j \neq i$, $\neg\mathsf{F}(j, i_t^r)$ applies when the non-receivers cannot distinguish between the event types e_1 and e_0 for each possible post on a topic. The alternative for such an

On the semantic level, we have proposed action models for the learning of personalized announcements above. On the syntactic level, in order to incorporate the learning events of personalized announcements, we add formulas of the form $[\mathbf{C}, e]\phi$ for a given action model $\mathbf{C}$, where e is an event in this model. The full dynamic language is obtained as follows,

Definition 6. *The language $\mathcal{L}^+$ of the dynamic logics of personalized announcements is given by*

$$\phi ::= i_t^r \mid \neg\phi \mid \phi \wedge \phi \mid \mathsf{F}(j, i_t^r) \mid B_i\phi \mid [\mathbf{C}, e]\phi$$

where $i_t^r \in \Phi$, $i, j \in Ag$, and $r \in R_t$ for $t \in T$. The symbol $\mathbf{C}$ stands for an action model for personalized announcements and e is an event of it.

The satisfaction of dynamic formulas such as $[\mathbf{C}, e]\phi$ depends on the update of the original opinion model, which is formally defined as follows,

Definition 7 (Updated model). *Given an opinion model $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}\rangle$ and an action model $\mathbf{C} = \langle \mathbf{E}, \{\mathbf{f}_{\mathbf{j}}\}_{j\in Ag}, \mathbf{pre}\rangle$ for the personalized announcement of i_t^r, the updated model is given by $\mathbf{M} \otimes \mathbf{C} := \langle \mathbf{W}', \{\rightarrow'_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}'\rangle$ where*

- $\mathbf{W}' = \{(w, e) \mid \mathbf{M}, w \vDash pre(e)\}$;
- $(w, e) \rightarrow'_{\mathbf{j}} (w', e')$ *iff* $w \rightarrow_j w'$, $\mathbf{M}, w \vDash \mathbf{f}_{\mathbf{j}}(e, e')$, *where* $\mathbf{f}_{\mathbf{j}}(e, e')$ *denotes the filtering function;*
- $\mathbf{V}'((w, e), j, t) = \mathbf{V}(w, j, t)$, *for* $w \in \mathbf{W}$, $j \in Ag$, $t \in T$.

The intuition underlying this definition is that when an agent posts an announcement on the social platform, the other agents in the community may update their beliefs with respect to it, depending on whether the information reaches them via the recommender system and if they consider the information to be acceptable.

It is easy to check that the *Opinion-Belief Condition*, imposed on opinion models, is preserved by the product update operation as specified in the above definition. Moreover, *Consistency of Belief* is also preserved by the update operation, i.e., the new doxastic accessibility relations are serial.

Lemma 1. *The doxastic accessibility relations in the updated model (Definition 7) are serial.*

explicit representation would be a more complex event model, introducing a separate event standing for 'some post is made' (and a dedicated syntactic item as its precondition). However, as the added value in terms of learning is not significant at this stage, we opt, in this paper, to use a more simple event model with only two event types.

Proof. We aim to show that for any (w, e) in $\mathbf{W}'$, there exists (u, e') in $\mathbf{W}'$ such that $(w, e) \rightarrow_j (u, e')$, for any $j \in Ag$. For any (w, e) in $\mathbf{W}'$, since $\mathbf{M}$ is serial, then for any agent k, there exists $u \in \mathbf{W}$ such that $w \rightarrow_k u$.

- For $j = i$, we have $(w, e) \rightarrow'_i (u, e)$, since $\mathbf{f_i}(e, e) = \top$ for any $e \in \mathbf{E}$.
- For $j \in Ag - \{i\}$, if e is e_0, then $(w, e_0) \rightarrow'_j (u, e_0)$, since $\mathbf{f_j}(e_0, e_0) = \top$.
  If e is e_1, there are two cases. Suppose $\boldsymbol{M}, w \vDash \mathsf{F}(j, i^r_t)$, then it follows that $\mathbf{M}, v \vDash i^r_t$ for some $v \in \mathbf{W}$ with $w \rightarrow_j v$, thus $(v, e_1) \in \mathbf{W}\times\mathbf{E}$. Since $\mathbf{f}_j(e_1, e_1) = \mathsf{F}(j, i^r_t)$, then we have that $(w, e_1) \rightarrow'_j (v, e_1)$. Else, suppose $\mathbf{M}, w \vDash \neg\mathsf{F}(j, i^r_t)$, since $\mathbf{f}_j(e_1, e_0) = \neg\mathsf{F}(j, i^r_t)$, then $(w, e_1) \rightarrow'_j (u, e_0)$. □

The semantics of dynamic formulas is given as follows.

Definition 8 (Dynamic semantics). *Consider an opinion model* $\mathbf{M}$, *a world* w *in* $\mathbf{M}$, *and an action model* $\mathbf{C}$ *for personalized announcement. We define the truth of* $\phi \in \mathcal{L}^+$ *at* w *in* $\mathbf{M}$ *inductively as in Definition 3 with the additional clause:*

$$\mathbf{M}, w \vDash [\mathbf{C}, e]\phi \textit{ iff } \mathbf{M}, w \vDash pre(e) \textit{ implies } \mathbf{M} \otimes \mathbf{C}, (w, e) \vDash \phi.$$

We present an example of an updated model below.

Example 3. *Consider the opinion model in Example 1 and the model for the personalized announcement of Helen's opinion h on topic D in Fig. 5, where preconditions are specified on the left, and the label* $\mathsf{A} \Leftarrow \neg\mathsf{F}(\mathsf{A}, \mathsf{H}^h_D)$ *on the edge from* e_1 *to* e_0 *expresses that* $\mathbf{f}_\mathsf{A}(e_1, e_0) = \neg\mathsf{F}(\mathsf{A}, \mathsf{H}^h_D)$*, and similarly for the rest. Suppose that the filtering condition employed by the social platform is the User-Based Push Condition with threshold* $\theta = 2$. *The updated model* $\mathbf{M}' = \langle \mathbf{W}', \{\rightarrow'_\mathbf{j}\}_{j\in Ag}, \mathbf{V}' \rangle$ *given these events is depicted in Fig. 6.*

In the example, since $\mathbf{M}, w \nvDash \mathsf{F}(A, \mathsf{H}^h_D)$ *and* $\mathbf{M}, w \nvDash \mathsf{F}(\mathsf{J}, \mathsf{H}^h_D)$*, we have, respectively,* $(w, e_1) \rightarrow'_\mathsf{A} (w, e_1)$ *while* $(w, e_1) \rightarrow'_\mathsf{J} (w, e_0)$ *and* $(w, e_1) \rightarrow'_\mathsf{J} (v, e_0)$. *As a result,* $\mathbf{M}', (w, e_1) \vDash B_\mathsf{A}\mathsf{H}^h_D$ *while* $\mathbf{M}', (w, e_1) \nvDash B_\mathsf{J}\mathsf{H}^h_D$. *That is, due to personalization, the very same announcement appeared in the feed of Alice, thereby changing her beliefs, but it did not appear in the feed of John, hence leaving his beliefs about Helen intact.*

Note that in the example above, when the event e_1 happens, Helen learns that Alice comes to believe her post H^h_D after it has been recommended to her. We also have $\mathbf{M}, w \vDash B_\mathsf{H}\neg B_\mathsf{J}\mathsf{H}^h_D$

and $\mathbf{M}', (w, e_1) \vDash B_H \neg B_J H_D^h$, that is, Helen H previously held the opinion that John J did not believe H_D^h, and this has not been changed by event e_1.

This shows that, in this particular example, the update dynamics requires some transparency with respect to the personalization mechanisms, where agents may gain information about what information other agents have learned. While this situation covers a subset of the social media phenomena (e.g., phenomena where some agents have special access to other agent's information, for instance, because they have administrative status and thus access to information about who receives what), our model can be easily generalized to capture other cases, such as private or semi-private announcements, by using standard action model types (Baltag and Moss (2004); Baltag and Renne (2016)). Moreover, note that our general class of access action models can be utilized to model situations where information does not only selectively reach members of the intended audience, but additionally no member of the network is aware of the filtering outcome of the recommender system, arguably due to the lack of transparency of how the chosen recommender system works. The possibility of utilizing such

$$pre(e) = \begin{cases} \mathtt{H}_D^h & \text{if } e = e_1 \\ \top, & \text{if } e = e_0 \end{cases}$$

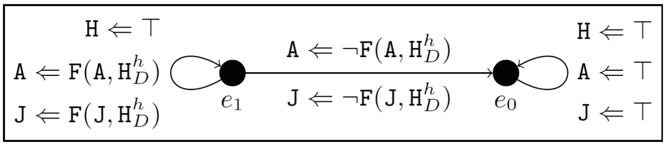

Fig. 5 The model for the personalized announcement of H_D^h.

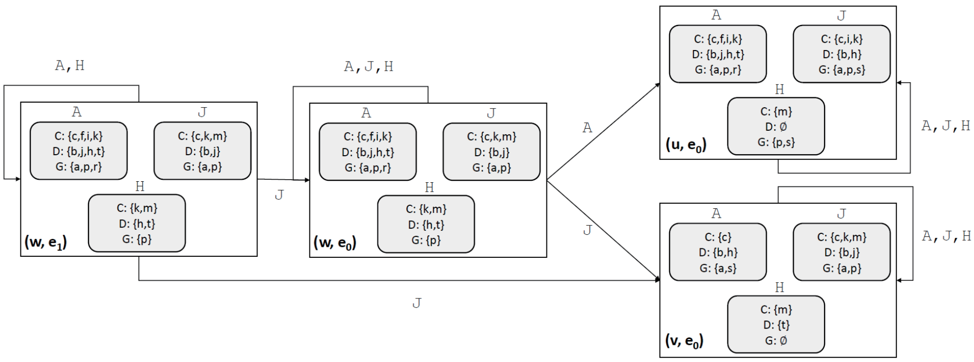

Fig. 6 The updated model $\mathbf{M}'$, after the personalized announcement of H_D^h.

generalizations of the standard action models to cover subtle issues of higher-order uncertainty is also evinced by the work of Baltag and Moss (2004); Bjorndahl and Nalls (2021); Bolander (2018); Engesser *et al*. (2021).

Finally, and related to our previous remark on page 16, the doxastic possibilities the non-receiver agent *J* entertained carry over to the updated model, namely, $[\mathbf{C}, e_1]\hat{B}_{\mathsf{J}}H^h_D$ and $[\mathbf{C}, e_1]\hat{B}_{\mathsf{J}}\hat{B}_{\mathsf{A}}H^h_D$ also hold at $\mathbf{M}, w$. In fact, more complex statements hold at that world, such as $\hat{B}_{\mathsf{J}}[\mathbf{C}, e_1]B_{\mathsf{A}}H^h_D$. Clearly, agent J entertained these possibilities initially and *does not cease to entertain them*, following the personalized announcement. This is due to the doxastic situation of the agents in the initial model, which the event model and the corresponding update do not undo; it is not that the event model itself represents the possibility that non-receiver J will come to entertain some possibility over receiver A and the post H^h_D, even though J does not in any case learn it.

3. Axiomatization

In this section, we propose sound and complete Hilbert-style proof systems for the logics described in Section 2. The axiom system of the each static logic SLPA_i is based on the minimal modal logic *K* with the axiom *D* for serial accessibility relations (known as *KD*), equipped with an axiom for the corresponding filtering condition *i*. The axiomatization of each dynamic logic DLPA_i expands that of the corresponding static logic with a set of reduction axioms.

3.1 Complete axiomatization of SLPA_i

We note that, as mentioned earlier, since in Section 2.1 we introduced four filtering condtions, we will now also introduce different axiom systems, each corresponding to a different condition. In this section then, we provide the "core" axioms of our logics, which always apply, and also list the different axioms corresponding to the four filtering conditions. Combining the core axioms with an axiom for the chosen filtering condition of a social network platform results in the full static axiomatization for the network in question. We use the following abbreviations, where $j \neq i$:

$$RIP(j, i^r_t) := \bigvee_{r' \in R_t} (j^{r'}_t \wedge \bigvee_{\{i_1,\ldots,i_\theta\} \subseteq Ag} \bigwedge_{1 \leq n \leq \theta} ({i_n}^{r'}_t \leftrightarrow {i_n}^{r}_t))$$

$$CIP(j, i_t^r) := \bigwedge_{r' \in R_t} (j_t^{r'} \to \bigvee_{\{i_1,\dots,i_\theta\} \subseteq Ag} \bigwedge_{1 \leq n \leq \theta} (i_{n_t}^{r'} \leftrightarrow i_{n_t}^{r}))$$

$$UP(j, i_t^r) := \bigvee_{\{r_1,\dots,r_\theta\} \subseteq R_t} \bigwedge_{1 \leq n \leq \theta} (i_t^{r_n} \wedge j_t^{r_n})$$

$$FP(j, i_t^r) := \bigvee_{r' \in R_t} j_t^{r'}$$

Let us first discuss the axioms and rules in Table 1. The main part of the axioms and rules stem from the axiom system of the modal logic *KD*, in which beliefs are required to be consistent (i.e., agents do not believe contradictions). This is supplemented in Table 1 with the positive and negative Opinion-Belief axioms, which together capture the condition of *Opinion-Belief* imposed on opinion models. The axiom for 'own opinions' expresses that no filtering condition imposed by the recommender system applies to one's own expressed opinions. In Table 2, we list the axioms for the different filtering conditions (and conditions of acceptability).

Definition 9. *The axiom system of $SLPA_i$ is given by the axioms and inference rules of Table 1, extended with one axiom from Table 2 corresponding to the selected filtering condition i.*[11]

Theorem 1. *The axiom system of $SLPA_i$ is sound and complete with respect to the class of opinion models with filtering condition i.*

Proof. For soundness, we can follow the standard way to show that all axioms and rules in Table 1 are valid. Note that the set of Ag and R_t are all finite when we prove the soundness of each axiom of Table 2.

For completeness, we can follow the canonical method (Blackburn *et al*., 2001). In the following, we give the definition of our canonical model.

Definition 10. *The canonical model* $\mathbf{M}^c$ *is a tuple* $\langle S^c, \{\to_i^c\}_{i \in Ag}, V^c \rangle$ *defined as follows*

- $S^c = \{w \mid w \text{ is a maximal } SLPA_i\text{-consistent set}\}$;
- $w \to_i^c u$ *iff for all formulas* ϕ, $\phi \in u$ *implies* $\hat{B}_i \phi \in w$;
- $V^c(w, i, t) = \{r \in R_t \mid i_t^r \in w\}$.

11 As an example, if the filtering condition is Radical Item-based Push, the corresponding axiom system is $SLPA_{\text{Radical Item-based Push}}$, namely the axiom system in Table 1 extended with the axiom Radical Item-based Push axiom from Table 2.

Table 1. The core Hilbert-style proof system of SLPA_i.

PC All instances of classical propositional tautologies
K $B_i(\phi \to \psi) \to (B_i\phi \to B_i\psi)$
Consistency $\mathrm{B}_i\phi \to \neg B_i\neg\phi$
Positive Opinion-Belief $i_t^r \to B_i i_t^r$
Negative Opinion-Belief $\neg i_t^r \to B_i \neg i_t^r$
Own opinion $\mathsf{F}(j,\ j_t^r) \leftrightarrow \hat{B}_j\, j_t^r$
Modus Ponens From ϕ and $\phi \to \psi$, infer ψ
Necessitation of B_i From ϕ infer $B_i\phi$

We now have to show that the canonical model is an opinion model. Since the Consistency axiom ensures the seriality of the doxastic accessibility relation, we only need to show that the canonical model satisfies the condition *Opinion-Belief*.

Suppose that $w \to_i^c u$ and $\mathbf{V}(w,i,t) \neq \mathbf{V}(u,i,t)$, then there are two cases.

- There exists $r \in \mathbf{V}(w, i, t)$ and $r \notin \mathbf{V}(u, i, t)$. It follows that $i_t^r \in w$ [1] and $i_t^r \notin u$, by the definition of V^c. From the definition of S^c, this means that $i_t^r \in w$ and $\neg i_t^r \in u$. From the definition of $\to^c$, $\hat{B}_i \neg i_t^r \in w$. However, from the Positive Opinion-Belief axiom and [1] we have that $B_i i_t^r \in w$, a contradiction.
- There exists $r \notin \mathbf{V}(w,i,t)$ and $r \in \mathbf{V}(u,i,t)$. It follows that $i_t^r \notin w$ and $i_t^r \in u$ by the definition of V^c. From the definition of S^c, this means that $\neg i_t^r \in w$ [1] and $i_t^r \in u$. From the definition of $\to^c$, $\hat{B}_i i_t^r \in w$. However, from the Negative Opinion-Belief axiom and [1] we have that $B_i \neg i_t^r \in w$, a contradiction.

Next, we can obtain the Existence Lemma following the standard method as well as the Truth Lemma, making use of induction on the complexity of ϕ and the properties of maximal SLPA_i-consistent sets. □

In this static system SLPA_i, we higlight the following interesting validities. For instance, when an agent j has expressed an opinion i_t^r, it will always be acceptable to her. This is stated in the form of the following validity in opinion models:

$$j_t^r \to \mathsf{F}(j,\ j_t^r)$$

Table 2. The different axioms for the acceptability and filtering conditions, where $j \neq i$.

Radical Item-based Push axiom
$\mathsf{F}(j, i_t^r) \leftrightarrow i_t^r \wedge \hat{B}_j i_t^r \wedge RIP(j, i_t^r)$
Conservative Item-based Push axiom
$\mathsf{F}(j, i_t^r) \leftrightarrow i_t^r \wedge \hat{B}_j i_t^r \wedge CIP(j, i_t^r)$
User-based Push axiom
$\mathsf{F}(j, i_t^r) \leftrightarrow i_t^r \wedge \hat{B}_j i_t^r \wedge UP(j, i_t^r)$
Feature Push axiom
$\mathsf{F}(j, i_t^r) \leftrightarrow i_t^r \wedge \hat{B}_j i_t^r \wedge FP(j, i_t^r)$

which follows from Theorem 1, using the combination of 'positive opinion-belief,' the 'consistency' of beliefs, and the 'own opinion' axiom. Observe also that agents hold only true beliefs about their own expressed opinions in this logic, i.e., the following implication is valid in opinion models:

$$B_j j_t^r \rightarrow j_t^r$$

which follows from Theorem 1, using the 'consistency' of belief together with the contraposition of 'negative opinion-belief.'

3.2 Complete axiomatization of DLPA$_i$

We now proceed to provide a complete axiomatization for the full dynamic logic. The full dynamic logic for a given social network platform is obtained from the reduction axioms and rules in Table 3, the core axioms and rules of Table 1 and an axiom from Table 2 corresponding to the filtering condition of interest.

The reduction axioms play a crucial role in the completeness proof, as they allow us to reduce the dynamic properties of the models to their static counterparts. All reduction axioms are valid – in the following two lemmas, we show that this holds for the last two reduction axioms in Table 3.

Lemma 2. $[\mathbf{C}, e]\mathsf{F}(j, i_t^r) \leftrightarrow (pre(e) \rightarrow \mathsf{F}(j, i_t^r))$ *is valid.*

Proof. Note that the satisfiablility of $RIP(j, i_t^r)$, $CIP(j, i_t^r)$, $UP(j, i_t^r)$, and $FP(j, i_t^r)$ at w are determined by $V(w, i, t)$ for $i \in Ag$, $r \in R_t$, where $t \in T$, thus $RIP(j, i_t^r)$, $CIP(j, i_t^r)$, $UP(j, i_t^r)$, and $FP(j, i_t^r)$ are satisfiable at w iff $RIP(j, i_t^r)$, $CIP(j, i_t^r)$, $UP(j, i_t^r)$ and $FP(j, i_t^r)$ are satisfiable at (w, e) respectively, where $pre(e)$ holds at w. We prove the two directions of the Lemma distinctly.

From left to right, given $\mathbf{M}, w \vDash [\mathbf{C}, e]\mathsf{F}(j, i_t^r)$, suppose $\mathbf{M}, w \vDash pre(e)$, then $\mathbf{M} \otimes \mathrm{C}, (w, e) \vDash \mathsf{F}(j, i_t^r)$. According to the truth condition of $\mathsf{F}(j, i_t^r)$, it follows that $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash i_t^r \wedge \hat{B}_j i_t^r$ and the filtering condition is satisfiable at (w, e). Since $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash i_t^r \wedge \hat{B}_j i_t^r$, then $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash i_t^r$ and there exists (v, e') in $\mathbf{M} \otimes \mathbf{C}$ such that $(w, e) \to'_j (v, e')$ and $\mathbf{M} \otimes \mathbf{C}, (v, e') \vDash i_t^r$. Thus we have that $\mathrm{M}, w \vDash i_t^r$, $w \to_j v$ and $\mathbf{M}, v \vDash i_t^r$, which implies $\mathbf{M}, w \vDash i_t^r \wedge \hat{B}_j i_t^r$ Thus, $\mathbf{M}, w \vDash \mathsf{F}(j, i_t^r)$.

From right to left, given that $\mathbf{M}, w \vDash pre(e) \to \mathsf{F}(j, i_t^r)$, suppose $\mathbf{M}, w \vDash pre(e)$, then $\mathbf{M}, w \vDash \mathsf{F}(j, i_t^r)$. Thus we have $\mathbf{M}, w \vDash i_t^r \wedge \hat{B}_j i_t^r$, then $\mathbf{M}, w \vDash i_t^r$ and $\mathbf{M}, u \vDash i_t^r$ for some $u \in \mathbf{W}$ with $w \to_j u$. It follows that $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash i_t^r$ and there exists (u, e) in $\mathbf{M} \otimes \mathbf{C}$, such that $(w, e) \to'_j (u, e)$ and $\mathbf{M} \otimes \mathbf{C}, (u, e) \vDash i_t^r$, which implies $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash i_t^r \wedge \hat{B}_j i_t^r$. Thus, $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash \mathsf{F}(j, i_t^r)$. □

Lemma 3. $[\mathbf{C}, e]B_i\phi \leftrightarrow (pre(e) \to \bigwedge_{e' \in E}(\mathbf{f}_i(e, e') \to B_i[\mathbf{C}, e']\phi))$ *is valid.*

Proof. Form left to right, given $\mathbf{M}, w \vDash [\mathbf{C}, e]B_i\phi$ [1], suppose $\mathbf{M}, w \vDash pre(e)$ [2]. We want to show that $\mathbf{M}, w \vDash \bigwedge_{e' \in E}(\mathbf{f}_i(e, e') \to B_i[\mathbf{C}, e']\phi)$. Take any $e' \in E$ such that $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$. It suffices to show that $\mathbf{M}, w \vDash B_i[\mathbf{C}, e']\phi$, i.e., $\mathbf{M}, w' \vDash [\mathbf{C}, e']\phi$, for all $w' \in \mathbf{W}$ such that $w \to_i w'$. That is, $\mathbf{M}, w' \vDash pre(e')$ implies $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$, for all $w' \in \mathbf{W}$ such that $w \to_i w'$.

From [1] we obtain: $\mathbf{M}, w \vDash pre(e)$ implies $\mathbf{M} \otimes \mathbf{C}, (w, e) \vDash B_i\phi$, i.e., $\mathbf{M}, w \vDash pre(e)$ implies $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$, for all $(w', e') \in \mathbf{W}'$ such that $(w, e) \to'_i (w', e')$. That is, $\mathbf{M}, w \vDash pre(e)$ implies $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$, for all $(w', e') \in \mathbf{W}'$ such that $w \to_i w'$ and $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$. From [1],[2]: $\mathbf{M} \otimes \mathrm{C}, (w', e') \vDash \phi$, for all $(w', e') \in \mathbf{W}'$ such that $w \to_i w'$ and $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$ [3].

Now take arbitrary $e' \in E$ such that $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$ and arbitrary $w' \in \mathbf{W}$ such that $w \to_i w'$. Suppose that $\mathbf{M}, w' \vDash pre(e')$, i.e., $(w', e') \in \mathrm{W}'$. From [3], we obtain $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$, which is precisely what we wanted to show.

Table 3. The additional reduction axioms and rules of DLPA_i.

Reduction axioms
$[\mathbf{C}, e]i_t^r \leftrightarrow pre(e) \to i_t^r$
$[\mathbf{C}, e](\phi_1 \wedge \phi_2) \leftrightarrow [\mathbf{C}, e]\phi_1 \wedge [\mathbf{C}, \mathrm{e}]\phi_2$
$[\mathbf{C}, e]\neg\phi \leftrightarrow pre(e) \to \neg[\mathbf{C}, e]\phi$
$[\mathbf{C}, e]\mathsf{F}(j, i_t^r) \leftrightarrow pre(e) \to \mathsf{F}(j, i_t^r)$
$[\mathbf{C}, e]B_i\phi \leftrightarrow pre(e) \to \bigwedge_{e' \in E}(\mathbf{f}_i(e, e') \to B_i[\mathbf{C}, e']\phi)$
Necessitation rule of $(\mathbf{C}, e)$
From ϕ infer $[\mathbf{C}, e]\phi$

From right to left, given $\mathbf{M}, w \vDash pre(e) \rightarrow \bigwedge_{e' \in E} (\mathbf{f}_i(e, e') \rightarrow B_i[\mathbf{C}, e']\phi)$ [1], we want to show that $\mathbf{M}, w \vDash [\mathbf{C}, e]B_i\phi$, i.e., $\mathbf{M}, w \vDash pre(e)$ implies $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$, for all $(w', e') \in \mathbf{W}'$ such that $w \rightarrow_i w'$ and $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$.

Suppose $\mathbf{M}, w \vDash pre(e)$. Take arbitrary $w' \in \mathbf{W}$ and $e' \in \mathbf{E}$ such that $\mathbf{M}, w' \vDash pre(e')$, $w \rightarrow_i w'$ and $\mathbf{M}, w \vDash \mathbf{f}_i(e, e')$. From [1] we obtain that $\mathbf{M}, w \vDash B_i[\mathbf{C}, e']\phi$. As a result, $\mathbf{M}, w' \vDash [\mathbf{C}, e']\phi$, i.e., $\mathbf{M}, w' \vDash pre(e')$ implies $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$. Because of our assumption, we obtain $\mathbf{M} \otimes \mathbf{C}, (w', e') \vDash \phi$ as desired. □

Definition 11. *The axiom system of DLPA$_i$ is given by the reduction axioms and rules of Tables 1 and 3, extended with one axiom from Table 2 corresponding to the selected filtering condition i.*

Theorem 2. *The axiom system of DLPA$_i$ is sound and complete with respect to the class of opinion models with filtering condition i and action models for personalized announcements.*

Proof. Since all reduction axioms in Table 3 are valid, and the inference rule in Table 3 preserves the validity, soundness follows.

For completeness, since the reduction axioms define a validity-preserving translation from $\mathcal{L}^+$ to $\mathcal{L}$, for any formula in DLPA$_i$, there is an equivalent formula in SLPA$_i$. According to Theorem 1, completeness follows. □

A number of interesting validities can be stated by making use of these reduction axioms of DLPA$_i$. The following three validities make essential use of the last two reduction axioms in Table 3 together with the axioms in the static system for the regulation of $\mathsf{F}(j, i_t^r)$ in Table 2.[12] Consider, for instance, a true post i_t^r that is acceptable for an agent j. If agent j also believes it to be acceptable, she will actually come to believe the post when the recommender system shows it in her feed, (i.e., when event type e_1 happens and j learns about the post i_t^r for any of the filtering conditions in Table 2 that successfully forward the post to j):

$$(i_t^r \wedge \mathsf{F}(j, i_t^r) \wedge B_j \mathsf{F}(j, i_t^r)) \rightarrow [\mathbf{C}, e_1]B_j i_t^r$$

Agents j who consider an expressed opinion i_t^r acceptable (i.e., $\mathsf{F}(j, i_t^r)$), when i_t^r was actually expressed but not disbelieved (i.e., $\neg B_j \neg i_t^r$), will still not disbelieve the post when the recommender system shows it in their feed (i.e., after event type e_1 happens and j learns about the post i_t^r):

12 The antecedents contain redundant conjuncts, which are nonetheless included to facilitate the reading of the formulas.

$$(\neg B_j \neg i_t^r \wedge i_t^r \wedge \mathsf{F}(j, i_t^r)) \rightarrow [\mathbf{C}, e_1]\neg B_j \neg i_t^r$$

Agents j, for whom a true post i_t^r is not acceptable because it is incompatible with their prior beliefs, will not come to believe it after event type e_1 happens:

$$(\neg B_j \neg i_t^r \wedge i_t^r \wedge \neg\mathsf{F}(j, i_t^r)) \rightarrow [\mathbf{C}, e_1]\neg B_j i_t^r$$

Consistent with our earlier remarks on pages 16 and 20, an agent z for whom the true post i_t^r is unacceptable and who recognizes that another agent j, for whom the post is acceptable, views i_t^r as possible, will continue to entertain that possibility even if she herself does not come to believe the post.

$$(i_t^r \wedge \mathsf{F}(j, i_t^r) \wedge \neg\mathsf{F}(z, i_t^r) \wedge \hat{B}_z \hat{B}_j i_t^r) \rightarrow [\mathbf{C}, e_l](B_j i_t^r \wedge \hat{B}_z \hat{B}_j i_t^r)$$

4. Logics for Attention in Personalized Social Platforms

In this section, we extend the previous setting, considering the second type of limitation, i.e., attention-driven filtering. As above, we sequentially introduce logics within both static and dynamic contexts to account for the fact that, given the limited attention available, agents only learn a subset of the information that they in principle have access to in a social platform.

4.1 Static logics of personalized announcements with attention

At first, we propose the Static Logics of Personalized Announcements with Attention (SLPAA_i), which makes use of the above introduced setting in addition to ideas coming from logics for bounded-rational agents in e.g., Alechina *et al*. (2011); Solaki (2021) and logics to model attention as a bounded resource in epistemic planning in Belardinelli and Rendsvig (2021). Let $Ag \neq \emptyset$, $T \neq \emptyset$ and $\{R_t\}_{t \in T}$ be defined as in the previous section. We introduce the *attention cost function*:

$$c: (\Phi \cup \{\top\}) \times Ag \rightarrow \mathbb{N}$$

This function assigns a cost to each atomic formula for each agent, where $c(i_t^r, j)$ intuitively corresponds to the subjective cost for agent j to process the recommended post i_t^r, and $c(\top, j)$ intuitively corresponds to the trivial (zero) cost of accessing no new information. Note that the cost function is only defined for atomic formulas and not for all formulas of the language.[13] We add it as a parameter of the model as we are not interested in modifying it, but rather keep it fixed throughout updates (similarly to the set of agents, topics, and positions on topics). Additionally, we extend the opinion model with an *attention resource function*, capturing, for each agent, the amount of attention that she has available at each world. This function is introduced as an element of the tuple defining the model and not as one of its parameters, as its value will be modified when updating the model, that is, when agents learn new information and their attention span is decreased (see below for the update definition).

Definition 12 (Attention Opinion Model). *Given* Ag, T, $\{R_t\}_{t\in T}$, *and a cost function c, an attention opinion model is a tuple* $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j\in Ag}, \mathbf{V}, \mathbf{AT}\rangle$ *where* $\langle \mathbf{W}, \{\rightarrow_j\}_{j\in Ag}, \mathbf{V}\rangle$ *is an opinion model, and where:*

> $\mathbf{AT}$: $Ag \times W \rightarrow \mathbb{N}$ *is the attention resource function, indicating the attentional resources available to each agent at each world.*

The attention resource function represents how much attention span an agent has at each world. It can also be interpreted as a cognitive "budget" or as the agent's mental "energy." In accord with the focus on realistic agents, each agent's budget to learn new information is taken to be limited.

Note that, as they build on opinion models, attention models are also parametrized by one of the filtering conditions discussed above.

To design our language, we follow the idea in Solaki (2021) where agents are equipped with a cognitive capacity and their cognitive effort comes with a 'cost.' The cognitive capacity and cost we envision in this paper are tied to an agent's attention, hence we introduce a set of terms to account (syntactically) for attentional costs and attentional capacities.

Definition 13 (Terms). *Let* S_0 *be* $\{c_{j,i_t^r} : i_t^r \in \Phi, j \in Ag\} \cup \{c_{j,\top} : j \in Ag\} \cup \{\alpha_j : j \in Ag\}$. *The set of terms* S *is the smallest set satisfying (a)* $S_0 \subseteq S$*; (b) if* $s_1, \ldots, s_k$ *are terms, then* $s := z_1 s_1 + \ldots + z_k s_k$ *is a term, where* $z_1, \ldots, z_k \in \mathbb{Z}$. *Each* c_{j,i_t^r} *is a term representing the attention cost of processing* i_t^r

13 Determining the actual attentional costs is a matter of empirical research on human cognition in social networks; we introduce the cost function as a simplified way to represent quantitatively that receiving posts in social network platforms is an effortful and resource-consuming process.

for agent j, $c_{j,\top}$ *represents the attention consumed by j when j does not process any post, and each* α_j *is a term representing the attention resource of agent j.*

Intuitively, we use the attention resource α_j to express the attention capacity of agent j. Thus the formula $a_1\alpha_1 + ... + a_k\alpha_k \geq b$ states that each agent i (from 1 to k) has a specific attention 'budget' available and that together they satisfy this inequality.

With the aid of the given definition for syntactic terms, we can define our extended language.

Definition 14 (Attention language). *The attention language* $\mathcal{L}^A$ *is the language* $\mathcal{L}$ *of Definition 2 augmented with attention atoms (linear inequalities).*

The language $\mathcal{L}^A$ *is given by*

$$\phi ::= i_t^r \mid s \geq z \mid \mathsf{F}(j, i_t^r) \mid \neg\phi \mid \phi \wedge \phi \mid B_i\phi$$

where $i, j \in Ag$, $s \in S$, $i_t^r \in \Phi$ *and* $z \in \mathbb{Z}$.

Operators $\rightarrow$, $\vee$, $\leftrightarrow$, $\hat{B}$, $\leq$, $=$, $-$ are defined standardly in terms of the given language constructs. Moreover, inequality formulas such as $\alpha_j \geq c_{j,i_t^r}$ are also well-formed in this language since they abbreviate $\alpha_j + (-1)c_{j,i_t^r} \geq 0$. Such formulas read as: "agent j has enough attention to process the post i_t^r." As we will see, this perspective is sufficient to capture that an agent may only learn a small subset of the announcements that are available and, in principle, acceptable to her. Inequality formulas do not appear in the formulas $\mathsf{F}(j, i_t^r)$. This is because in social media, we usually do not have access to information about how much attention an agent has available.

Definition 15 (Term interpretation). *Let* $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_j\}_{j\in Ag}, \mathbf{V}, \mathbf{AT}\rangle$ *be an attention opinion model and w a world in it. Terms in S are interpreted recursively as follows* $c_{j,i_t^r}^{\mathbf{M},w} = c(i_t^r, j)$, $c_{j,\top}^{\mathbf{M},w} = c(\top, j)$, $\alpha_j^{\mathbf{M},w} = \mathbf{AT}(w, j)$ *and* $(z_1s_1 + ... + z_ks_k)^{\mathbf{M},w} = z_1s^{\mathbf{M},w} + ... + z_ks^{\mathbf{M},w}$.

The terms of the type c_{j,i_t^r} and $c_{j,\top}$ are interpreted by the cost function and the terms of the type α_j are interpreted by the attention resource function. Here we suppose that agents can always access and process their own posts without paying any attentional cost, denoted by $c(i_t^r, i) = 0$. Furthermore, if the platform does not push information to the agents, this information naturally does not consume agent j's attentional resources, which is denoted by $c(\top, j) = 0$. On the basis of the interpretation of terms, we give the formal truth conditions for inequalities as follows

Definition 16 (Semantic clause). *Given an attention opinion model* $\mathbf{M}$ *and world* $w \in \mathbf{W}$:

$$\mathbf{M}, w \models (s \geq z) \text{ iff } s^{\mathbf{M},w} \geq z$$

We define the truth of the other formulas in $\mathcal{L}^A$ in attention models as in Definition 3.

4.2 Dynamic logics of personalized announcements with attention

We now propose the Dynamic Logics of Personalized Announcements with Attention (DLPAA$_i$), by extending the dynamic definitions introduced in Section 2.2 to account for the fact that agents can only pay attention to some subset of announced formulas that are acceptable to them. Indeed, agents only have a finite amount of cognitive resources, so they will not learn everything that is in principle available to them.

Definition 17 (Attention action model for the personalized announcement of i_t^r). *An attention action model for the personalized announcements of* i_t^r *is a tuple* $\mathbf{D} := \langle \mathbf{E}, \{\mathbf{g_j}\}_{j \in Ag}, \mathbf{pre} \rangle$, *where* $\mathbf{E}$ *and* $\mathbf{pre}$ *are defined in Definition 5 and* $\mathbf{g_j}$ *is defined as follows.*

- $\mathbf{g_j} : E \times E \rightarrow \mathcal{L}^A$ *is the function representing two types of filtering.*

$$for\ j = i,\ \mathbf{g_j}(e, e') = \begin{cases} \top, & if\ e = e' = e_1 \\ \top, & if\ e = e' = e_0 \\ \bot, & otherwise \end{cases}$$

$$for\ j \neq i,\ \mathbf{g_j}(e, e') = \begin{cases} \mathsf{F}(j, i_t^r) \wedge (\alpha_j - c_{j,pre(e)} \geq 0), & if\ e = e' = e_1 \\ \neg(\mathsf{F}(j, i_t^r) \wedge (\alpha_j - c_{j,pre(e)} \geq 0)), & if\ e = e_1\ and\ e' = e_0 \\ \bot, & if\ e = e_0\ and\ e' = e_1 \\ \top, & if\ e = e' = e_0 \end{cases}$$

According to this definition, an agent processes a personalized announcement, if both types of filtering conditions, specified by the function $\mathbf{g_j}$, are satisfied. When an acceptable post is recommended to an agent, i.e., it passes the network's filter, *and* the agent has sufficient attention resources to process it, then the agent accesses its content. If instead the posted message is not acceptable *or* the agent does not have enough attention to pay for the cost of processing the posted message, then the agent will not access the post.

The following definition captures the effect of executing an event in a given possible world. It takes as input an attention opinion model and an attention action model and defines its product update. In the updated model the attention resources that are available to an agent are updated, depending on the cost she paid for processing a post.

Definition 18 (Attention product model). *Given an attention opinion model* $\mathbf{M} = \langle \mathbf{W}, \{\rightarrow_{\mathbf{j}}\}_{j\in A}, \mathbf{V}, \mathbf{AT}\rangle$ *and an attention action model* $\mathbf{D} = \langle \mathbf{E}, \{\mathbf{g_j}\}_{j\in Ag}, \mathbf{pre}\rangle$ *for a personalized announcement of* i_t^r, *the updated product model is given by* $\mathbf{M} \otimes \mathbf{D} := \langle \mathrm{W}', \{\rightarrow'_j\}_{j\in Ag}, \mathbf{V}', \mathbf{AT}'\rangle$ *where* $\mathbf{W}'$ *and* $\mathbf{V}'(w, i, t)$ *are updated as in Definition 7, and where*

- $(w, e) \rightarrow'_j (w', e')$ *iff* $w \rightarrow_j w'$ *and* $\mathbf{M}, w \vDash \mathbf{g_j}(e, e')$;
- $\mathbf{AT}'((w,e),j) = \begin{cases} \mathbf{AT}(w,j) - c(pre(e),j), \ \textit{if} \ \mathbf{M}, w \vDash \mathsf{F}(j, i_t^r) \wedge (\alpha_j - c_{j,pre(e)} \geq 0); \\ \mathbf{AT}(w,j), \ \textit{otherwise.} \end{cases}$

Note that the product update preserves the conditions imposed on attention opinion models. In particular, the preservation of seriality can be proven as we did in Lemma 1. To give reduction axioms later, we define the updated terms of attention capacity and costs in the language as follows.

- $\alpha_j^{(\mathbf{D},e)} = \begin{cases} \alpha_j - c_{j,pre(e)}, \ \text{if } \mathbf{M}, w \vDash \mathsf{F}(j, i_t^r) \wedge (\alpha_j - c_{j,pre(e)} \geq 0); \\ \alpha_j, \ \text{otherwise;} \end{cases}$

- $c_{j,i_t^r}^{(\mathbf{D},\mathbf{e})} = c_{j,i_t^r}$ and $c_{j,\top}^{(\mathbf{D},\mathbf{e})} = c_{j,\top}$

Example 4. *Consider the attention opinion model* $\mathbf{M} := \langle \mathbf{W}, \{\rightarrow_{\mathbf{j}}\}_{j\in A}, \mathbf{V}, \mathbf{AT}\rangle$ *where* $\langle \mathbf{W}, \{\rightarrow_{\mathbf{j}}\}_{j\in A}, \mathbf{V}\rangle$ *is as in Example 1 and where* $\mathbf{AT}(w, A) = \mathbf{AT}(w, J) = 10$ *and the same holds for worlds u and v. Consider Helen's post about her opinion h on topic D, i.e.,* H_D^h. *We illustrate the model and dynamics by considering three situations, that differ either in the value of the threshold for the structural filter, or in the cost assigned to learning* H_D^h.

- *Let* $c(H_D^h, A) = c(H_D^h, J) = 5$ *and suppose that the filter is User-Based Push with threshold* $\theta = 2$, *as in Example 3. Call* $\mathbf{D}$ *the attention action model for the personalized announcement of* H_D^h *based on this filter. In this case, both Alice and John have enough attention to cover for the cost of processing the formula* H_D^h, *at all worlds, i.e.,* $\mathbf{M}, w \vDash \alpha_i - c_{i,H_D^h} \geq 0$ *for all w and agents* $i = A, J$. *However, while Alice can actually pay attention to the post* *and learn it, as the post satisfies the User-Based filtering condition for Alice at w and it is acceptable for her at w* (*i.e.*, $\mathbf{M}, w \vDash \mathsf{F}(A, H^h_D)$),[14] *for John this is not the case. Indeed, the post does not satisfy the filtering condition for John at w* (*i.e.*, $\mathbf{M}, w \vDash \neg\mathsf{F}(J, H^h_D)$),[15] *so even if he could in principle learn it, as it is coherent with his beliefs in w and he has enough attention for it, he does not learn it. Hence, the updated model* $\mathbf{M} \otimes \mathbf{D} = \langle \mathbf{W}', \{\to'_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}', \mathbf{AT}'\rangle$ *is such that* $\langle \mathbf{W}', \{\to'_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}''\rangle$ *is as in Fig. 6, where Alice learns about* H^h_D *and John does not, i.e.,* $\mathbf{M} \otimes \mathbf{D}, (w, e_1) \vDash B_A H^h_D \wedge \neg B_J H^h_D$. *Additionally, Alice here spends some attention to learn the formula, i.e.,* $\mathbf{AT}'((w, e_1), A) = 10 - 5 = 5$ *while John does not, i.e.,* $\mathbf{AT}'((w, e_1), J) = 10$.

- *Now let the cost of learning* H^h_D *be* $c(H^h_D, A) = 5$ *and* $c(H^h_D, J) = 15$, *and suppose that the filter is User-Based Push with threshold* $\theta = 0$. *Call* $\mathbf{D}'$ *the attention action model for personalized announcement of* H^h_D *based on this filter and the new threshold. Because the threshold is zero, we now have that both Alice and John find the post acceptable, that is,* $\mathbf{M}, w \vDash \mathsf{F}(A, H^h_D) \wedge \mathsf{F}(J, H^h_D)$. *However, while Alice has enough attention resources to cover for its cost, John does not, i.e., for all w,* $\mathbf{M}, w \vDash \alpha_A - c_{A,H^h_D} \geq 0$ *and* $\mathbf{M}, w \vDash \alpha_J - c_{J,H^h_D} \leq 0$. *This means that the updated model* $\mathbf{M} \otimes \mathbf{D}' = \langle \mathrm{W}', \{\to'_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}', \mathbf{AT}'\rangle$ *is still such that* $\langle \mathbf{W}', \{\to'_{\mathbf{j}}\}_{j\in Ag}, \mathbf{V}'\rangle$ *is as in Fig. 6, where* $\mathbf{M} \otimes \mathbf{D}', (w, e_1) \vDash B_A H^h_D \wedge \neg B_J H^h_D$, *but this result is now obtained because of John's lack of attention resources, i.e., because* $\mathbf{M}, w \vDash (\alpha_J - c_{J,H^h_D} \leq 0)$. *As before, Alice spends some attention but John does not, i.e.,* $\mathbf{AT}'((w, e_1), A) = 10 - 5 = 5$ *and* $\mathbf{AT}'((w, e_1), J) = 10$.

Definition 19 (Dynamic Attention Language). *The dynamic attention language* $\mathcal{L}^{A+}$ *is the language* $\mathcal{L}^{A}$ *augmented with dynamic operators.*

The language $\mathcal{L}^{A+}$ *is given by*

$$\phi ::= i^r_t \mid s \geq z \mid \mathsf{F}(j, i^r_t) \mid \neg\phi \mid \phi \wedge \phi \mid B_i\phi \mid [\mathbf{D}, e]\phi$$

where $i, j \in Ag$, $s \in S$, $i^r_t \in \Phi$ *and* $z \in \mathbb{Z}$. *The symbol* $\mathbf{D}$ *stands for an attention action model for personalized announcements and e is an event of it.*

14 It is the case that $\mathbf{M}, w \vDash \mathsf{F}(A, H^h_D)$ as $\mathbf{M}, w \vDash H^h_D$ with $w \to_{\mathsf{A}} w$, and $|\mathbf{V}(w, H, D) \cap \mathbf{V}(w, A, D)| \geq 2$, so the post H^h_D satisfies the User-Based filtering condition for Alice at w.

15 It is the case that $\mathbf{M}, w \vDash \neg\mathsf{F}(\mathsf{J}, H^h_D)$ as $|\mathbf{V}(w, H, D) \cap \mathbf{V}(w, A, D)| \leq 2$, so the post H^h_D does not satisfy the User-Based filtering condition for John at w.

Definition 20 (Dynamic semantics). *Consider an attention opinion model* $\mathbf{M}$, *a world* w *in* $\mathbf{M}$, *and an attention action model* $\mathbf{D}$ *for personalized announcements. We define the truth of* $\phi \in \mathcal{L}^{A+}$ *at* w *in* $\mathbf{M}$ *inductively as in Definitions 3 and 16 with the additional clause*

$$\mathbf{M}, w \vDash [\mathbf{D}, e]\phi \text{ iff } \mathbf{M}, w \vDash pre(e) \text{ implies } \mathrm{M} \otimes \mathrm{D}, (w, e) \vDash \phi.$$

4.3 Axiomatization of SLPAA$_i$ and DLPAA$_i$

In this section, we present our axiomatization results, which follows the approach of other axiomatizations in the literature that include inequalities Fagin *et al*. (1990); Fagin and Halpern (1994); Halpern (2017); Solaki (2021).

Note that *Non-negative attention*, *Self-attention*, and *No-attention* in Table 4 correspond to the three assumptions about attention made in Definition 12 (Non-Negative Attention) and below Definition 15 (Self-attention and No-attention). Specifically, the agent's attention capacity is non-negative, the agent does not consume attention to process their own posts, and no attention is consumed when the agent does not process any new information. The *INEQ* axioms in Table 4 are introduced to accommodate the linear inequalities.

The validity of the axioms can be proved using standard methods, and in particular, the proofs for the new reduction axioms in Table 5 are similar to the proofs in Section 3.2; here we only pay attention to the reduction axiom for linear inequalities. We prove that the axiom $[\mathbf{D}, e](z_1 s_1 + ... + z_k s_k \geq z) \leftrightarrow pre(e) \rightarrow (z_1\, s_1^{(\mathbf{D},e)} + ... + z_k\, s_1^{(\mathbf{D},e)} \geq z)$ is valid below.

By the given definition, it is sufficient to prove that $\mathbf{M} \otimes \mathbf{D}, (w, e) \vDash z_1 s_1 + ... + z_k s_k \geq z$ iff $\mathbf{M}, w \vDash z_1\, s_1^{(\mathbf{D},e)} + ... + z_k\, s_k^{(\mathbf{D},e)} \geq z$, which is $z_1\, s_1^{\mathbf{M}\otimes\mathbf{D},(w,e)} + ... + z_k\, s_k^{\mathbf{M}\otimes\mathbf{D},(w,e)} \geq z$ iff $z_1\, (s_1^{(\mathbf{D},e)})^{\mathbf{M},w} + ... + z_k\, (s_k^{(\mathbf{D},e)})^{\mathbf{M},w} \geq z$. We claim that for $1 \leq t \leq k$, $s_t^{\mathbf{M}\otimes\mathbf{D},(w,e)}$ iff $(s_t^{(\mathbf{D},e)})^{\mathbf{M},w}$. We have two cases.

- If s_t is of the form α_j, then
  –Suppose that $\mathbf{M}, w \vDash \mathsf{F}(j, i_t^r) \wedge (\alpha_j - c_{j,pre(e)} \geq 0)$. By Definition 18 and the updated values of terms, we have $s_t^{\mathbf{M}\otimes\mathbf{D},(w,e)} = (s_t^{(\mathbf{D},e)})^{\mathbf{M},w} = \mathbf{AT}(w, j) - c(pre(e), j)$.
  –Suppose not, then $s_t^{\mathbf{M}\otimes\mathbf{D},(w,e)} = (s_t^{(\mathbf{D},e)})^{\mathbf{M},w} = \mathbf{AT}(w,j)$.
- If s_t is the form of c_{j,i_t^r} or $c_{j,\top}$, since such terms and the corresponding cost function remain the same after the update, then $s_t^{\mathbf{M}\otimes\mathbf{D},(w,e)} = (s_t^{(\mathbf{D},e)})^{\mathbf{M},w} = c(pre(e), j)$.

Definition 21. *The axiom system of $SLPAA_i$ is given by the reduction axioms and rules of Tables 1 and 4, extended with one axiom from Table 2, corresponding to the selected filtering condition i.*

Theorem 3. *The axiom system of $SLPAA_i$ is sound and complete with respect to the class of attention models with the selected filtering condition i.*

Proof. The soundness result is straightforward as mentioned. For the completeness, we can take a similar approach to that in the proof of Theorem 1. Using the attention axioms, it can be verified that the canonical model fulfils the properties of attention models and using the properties of maximal consistent sets and *INEQ*, it can be shown that the truth lemma case for the inequality formulas holds as well. □

Definition 22. *The axiom system of $DLPAA_i$ is given by the reduction axioms and rules of Tables 1, 4 and 5, extended with one axiom from Table 2 corresponding to the selected filtering condition i.*

Theorem 4. *The axiom system of $DLPAA_i$ is sound and complete with respect to the class of attention models with selected filtering condition i and attention action models for personalized announcements.*

Table 4. The additional axioms of SLPAA_i.

Non-negative attention	$\alpha_i \geq 0$
Self-attention	$c_{i,i_t^r} \geq 0 \wedge (-1)\, c_{i,i_t^r} \geq 0$
No-attention	$c_{i,\top} \geq 0 \wedge (-1)c_{i,\top} \geq 0$
INEQ	All instances of valid formulas about linear inequalities

Table 5. The additional reduction axioms and rules of DLPAA_i.

Reduction axioms
$[\mathbf{D}, e]i_t^r \leftrightarrow pre(e) \rightarrow i_t^r$
$[\mathbf{D}, e](z_1 s_1 + ... + z_k s_k \geq z) \leftrightarrow pre(e) \rightarrow (z_1\, s_k^{(\mathbf{D},e)} + ... + z_k\, s_k^{(\mathbf{D},e)} \geq z)$
$[\mathbf{D}, e](\phi_1 \wedge \phi_2) \leftrightarrow [\mathbf{D}, e]\phi_1 \wedge [\mathbf{D}, e]\phi_2$
$[\mathbf{D}, e]\neg\phi \leftrightarrow pre(e) \rightarrow \neg[\mathbf{D}, e]\phi$
$[\mathbf{D}, e]\mathsf{F}(j, i_t^r) \leftrightarrow pre(e) \rightarrow \mathsf{F}(j, i_t^r)$
$[\mathbf{D}, e]B_i\phi \leftrightarrow pre(e) \rightarrow \bigwedge_{e' \in E} (g_i(e, e') \rightarrow B_i[\mathbf{D}, e]\phi)$
Necessitation rule of $(\mathbf{D}, e)$
From ϕ infer $[\mathbf{D}, e]\phi$

Proof. The soundness result is also straightforward, as mentioned. For the completeness, the result follows as in Theorem 2, i.e., by combining the validity of the reduction axioms of Table 5 with Theorem 3. □

Similar to Section 3, a number of validities can be derived in the DLPAA$_i$ setting as well. Most importantly, if a post i_t^r is acceptable for agent *j and* the agent has sufficient attentional resources, then the agent will actually come to believe the post when the recommender system shows it in her feed:

$$\mathrm{F}(j, i_t^r) \wedge (\alpha_j \geq c_{j,i_t^r}) \rightarrow [\mathbf{D}, e_1]B_j i_t^r$$

The validities of our frameworks reflect the interplay of structural and attentional limitations and their impact on the formation of beliefs of agents in social networks. Their implications can be even more salient in combined social-doxastic logics (Baltag *et al*., 2019; Smets and Velázquez-Quesada, 2019a) that additionally model the friendship, followership, or closeness relations among agents. In particular, if DLPA$_i$ or DLPAA$_i$ is used as the underlying doxastic logic of the social network logic, instead of standard doxastic logics, it will result in a more realistic modelling of the diffusion of opinion and the formation of clusters and cascades; the limits in the agents' access to information and their effects on the agents' beliefs about each other (captured in the current setting) would in turn result in different decisions on which opinions to adopt and how to form or manipulate one's social network relations.

5. Discussion and Future Work

Our focus in this paper has been on studying the belief dynamics of agents on social platforms triggered by the personalized information they gain access to, taking into account two crucial factors: *structural filtering* and *attention-driven filtering*. We first developed a logical framework to characterize the belief dynamics arising from the first factor. Subsequently, we extended the framework to incorporate the second factor.

From a technical perspective, we employed the following main ingredients: (i) a variation of event models and corresponding product updates (ii) quantitative formulas to express attention bounds, and (iii) the common completeness-via-reduction procedure to obtain the two sound and complete axiomatizations. While each of these ingredients has been previously used, they originate from quite 'disparate' applications. Similar variations of event models were used for the logical study of false belief tasks, coming from developmental psychology (Bolander, 2018), and the axiomatizations of languages with linear inequalities originate from the logical study of knowledge and probability (Fagin *et al*., 1990; Fagin and Halpern, 1994). In our framework, the combination of these ingredients, suitably adjusted and re-contextualized for personalization and attention, delivers simple sound and complete logics which can be valuable in that they provide a common and new application avenue of these constructs on the study of social network phenomena.

The *simplicity* of the resulting logics could contribute to the broader enterprise of using logical and philosophical methods to study these phenomena. Avoiding spurious, artificially contrived, or opaque constructs and techniques is crucial for the *metatheoretical* study of personalization and attention, which requires simplicity and well-chosen abstrations, in what is otherwise a field dominated by opaque machine learning techniques. Therefore, to have simple, human-readable, and easily interpretable symbolic statements (without forfeiting soundness and completeness) is, in fact, a desirable feature. This simplicity not only strengtens the case for the value of philosophical analysis in understanding network phenomena but also highlights the potential of logical systems to address explainability-related challenges. These challenges are becoming increasingly urgent, particularly in terms of the accountability and responsibility for the platforms in question, with significant regulatory implications (Wiewiórowski, 2024).

Our proposed framework to model the dynamics for personalization and attention in social networks provides a basis on which further work can be built. We envision different avenues to extend our setting. First, note that our modelling of the doxastic layer of a social network leaves room for further development. So far, our treatment of the static logic of beliefs was based on the weak axiom system known as *KD* without imposing further idealizing properties, such as unlimited introspection, on cognitively constrained agents. Similarly, the event models we have considered fall within the class of models already studied within DEL, but are here applied to the class of scenarios in which we reason about the acceptability of personal announcements that pass the structural filtering conditions of recommender systems and the agent's attention-driven filtering. As such, our current approach can be integrated with the generalized study of epistemic actions, which, in the future, could be handy to obtain sound and complete axiomatizations for possibly extended settings. For example, Bjorndahl & Nalls (Bjorndahl and Nalls, 2021) propose to make the edges of the event model world-dependent, thereby 'endogenizing' the event model within the primary Kripke structures. That is, each structure strture comes with a function that maps world-agent pairs to relations on events, which is then used to define the updated accessibility relations. The function effectively 'arbitrates' the informational access of each agent to the event model and offers greater flexibility to model scenarios intuitively and simply, without the proliferation of spurious events. This approach is similar to ours in that our acceptability and filtering functions (**f** and **g**) play an analogous arbitrating role and help define the product update, based on whether a condition is satisfied and per type of agent. When it comes to the axiomatization, Bjorndahl & Nalls introduced additional primitive syntactic items in the language, $\xi_{j,\sigma,\sigma'}$ constants (where j stands for agents, and σ, σ' stand for events), interpreted with the use of the endogenous function. Then dynamic axiomatizations can be obtained in the standard way, where the constants are especially useful for the reduction axiom of epistemic attitudes. This is similar to the use of $\mathsf{F}(j,\ i_t^r)$ and $\alpha_j \geq c_{j,\, i_t^r}$ formulas, needed for our reduction axioms, which are nonetheless not introduced ad-hoc as but they are *already* included in the object language and have meaningful readings in terms of personalization (Section 2) and attention (Section 4).

In our follow-up work, we plan to enrich the current setting by taking on board other epistemic attitudes, such as different types of knowledge and belief (including conditional beliefs), to extend our semantics to plausibility models for belief (Baltag and Smets, 2016; van Benthem, 2007). Right now the agents in our model cannot accept information that directly contradicts their prior beliefs. As such the type of belief dynamics that is considered in this paper falls under 'belief expansion.' Allowing for scenarios that lift this constraint, i.e., when working with agents that can also 'accept' surprising information that contradicts their prior beliefs, requires a different notion of 'acceptance.' The use of plausibility models is a natural choice when studying richer notions of 'acceptance' that can involve a contingency plan for belief revision (which our current model in this paper does not include). The modelling of events in plausibility models will also provide a richer setting to model the different types of events considered in this paper as well as other variations that would lift the mentioned transparency constraints.

Social relationships among users play a crucial role in information distribution and can be affected by it as well. Hence the natural next step is to combine the modelling of the doxastic layer with social network models with "friendship" or "neighbour" relations for a full-fledged social-doxastic framework, as in (Seligman *et al*., 2013; Solaki *et al*., 2016; Smets and Velázquez-Quesada, 2017, 2018, 2019a,b, 2020; Baltag *et al*., 2019), but made sensitive to structural and attentive limitations. We expect that such a framework will help us shed light on phenomena like the formation of echo chambers or informational cascades in online networks, since it will be closer to the reality of agents' behaviour in such networks. It will similarly be interesting to study the decidability and other computational properties of these richer logics. Other directions for future work include enabling agents to send various types of composite posts, instead of only allowing them to send posts that reflect atomic information, or further exploring aspects such as agents being able to modify or retract their own previously expressed opinions.

Taking a broader perspective, we can also address other kinds of filtering. For example, when agents send posts, they can proactively open or block information channels leading to specific groups. Likewise, when agents access information, they can actively choose to open or block information channels originating from specific groups. Furthermore, in our work, we focus on a single type of entity, which is agents. However, some social platforms may involve multiple types, such as information sources solely dedicated to news distribution. At the same time, the design of user ranks can also be a crucial component of certain social platforms, such as premium and free users. Hence, from a logical perspective, to investigate the role of user types and user ranks on information distribution within social platforms is also a part of our future work.

Acknowledgements

Gaia Belardinelli is funded by the Independent Research Fund Denmark (grant no. 4255-00020B).

The authors thank Johan van Benthem, the anonymous referees of this paper as well as the audiences of the various workshops where earlier versions of this work were presented, for their valuable comments and feedback.

References

Aggarwal, C. C. (2016). *Recommender systems*, volume 1. Springer.

Alechina, N., Logan, B., Nguyen, H. N., and Rakib, A. (2011). Logic for coalitions with bounded resources. *Journal of Logic and Computation*, 21(6), 907–937.

Baltag, A., Christoff, Z., Rendsvig, R. K., and Smets, S. (2019). Dynamic epistemic logics of diffusion and prediction in social networks. *Studia Logica*, 107(3), 489–531.

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

Baltag, A., Moss, L. S., and Solecki, S. (1998). The logic of public announcements, common knowledge, and private suspicions. In *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*, 43–56, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Baltag, A. and Renne, B. (2016). Dynamic epistemic logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2016 edition.

Baltag, A. and Smets, S. (2016). *A qualitative theory of dynamic interactive belief revision*, 813–858. Springer International Publishing, Cham.

Belardinelli, G., Bolander, T., and Watzl, S. (2025). A logic for general attention using edge-conditioned event models. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, IJCAI '25.

Belardinelli, G. and Rendsvig, R. K. (2021). Epistemic planning with attention as a bounded resource. In *International Workshop on Logic, Rationality and Interaction*, 14–30. Springer.

Bjorndahl, A. and Nalls, W. (2021). Endogenizing epistemic actions. *Studia Logica*, 109(5), 1049–1091.

Blackburn, P., de Rijke, M., and Venema, Y. (2001). *Modal logic*. Cambridge University Press.

Bolander, T. (2018). *Seeing is believing: formalising false-belief tasks in dynamic epistemic logic*, 207–236. Springer International Publishing.

Cherniak, C. (1986). *Minimal rationality*. MIT Press.

Engesser, T., Mattmüller, R., Nebel, B., and Thielscher, M. (2021). Game description language and dynamic epistemic logic compared. *Artificial Intelligence*, 292, 103433.

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., and Sandvig, C. (2015). I always assumed that i wasn't really that close to [her]: Reasoning about invisible algorithms in the news feed. In *Proceedings of the 33rd annual SIGCHI conference on human factors in computing systems*, 153–162, NY. ACM.

Fagin, R. and Halpern, J. Y. (1994). Reasoning about knowledge and probability. *Journal of the ACM*, 41(2), 340–367.

Fagin, R., Halpern, J. Y., and Megiddo, N. (1990). A logic for reasoning about probabilities. *Information and Computation*, 87(1), 78–128. Special Issue: Selections from 1988 IEEE Symposium on Logic in Computer Science.

Flaxman, S., Goel, S., and Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1), 298–320.

Halpern, J. Y. (2017). *Reasoning about uncertainty*. The MIT Press.

Halpern, J. Y. and Pucella, R. (2011). Dealing with logical omniscience: Expressiveness and pragmatics. *Artificial Intelligence*, 175(1):220–235.

Jain, A. and Gupta, C. (2018). *Fuzzy logic in recommender systems*, 255–273. Springer International Publishing.

Jannach, D., Zanker, M., Felfernig, A., and Friedrich, G. (2010). *Recommender systems: An Introduction*. Cambridge University Press.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Li, L. (2023). *Games, boards and play: A logical perspective*. PhD thesis, University of Amsterdam.

Litman, L. and Reber, A. S. (2005). Implicit cognition and thought. In Holyoak, K. and Morrison, B., editors, *The Cambridge Handbook of Thinking and Reasoning*, 431–453. Cambridge University Press.

Liu, F., Wang, W., and Yang, S. (2024). Towards a logical approach to recommendations. In Liao, B., Pang, J., and Rienstra, T., editors, *Logics for New-Generation AI 2024: Proceedings of the Fourth International Workshop*, 47–65, Hangzhou.

Olsson, E. (2023). Coherentist theories of epistemic justification. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2023 edition.

Paul, K. (2019). Instagram tests hiding how many people like a post. that has influencers worried. https://www.theguardian.com/technology/2019/nov/15/instagram-likes-influencers-social-media.

Proferes, N. (2017). Information flow solipsism in an exploratory study of beliefs about twitter. *Social Media plus Society*, 3(1), 1–17.

Schwitzgebel, E. (2010). Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, 91(4), 531–553.

Seligman, J., Liu, F., and Girard, P. (2013). Facebook and the epistemic logic of friendship. *arXiv preprint arXiv:1310.6440*.

Simon, H. A. (1971). Designing organizations for an information-rich world. *Computers*, *communications*, *and the public interest*, 72, 37.

Smets, S. and Velázquez-Quesada, F. R. (2017). How to make friends: A logical approach to social group creation. In Baltag, A., Seligman, J., and Yamada, T., editors, *Logic*, *Rationality*, *and Interaction*, 377–390. Springer Berlin Heidelberg.

Smets, S. and Velázquez-Quesada, F. R. (2018). The creation and change of social networks: A logical study based on group size. In Madeira, A. and Benevides, M., editors, *Dynamic Logic. New Trends and Applications*, 171–184. Springer International Publishing.

Smets, S. and Velázquez-Quesada, F. R. (2019a). A logical analysis of the interplay between social influence and friendship selection. In *International Workshop on Dynamic Logic*, 71–87. Springer.

Smets, S. and Velázquez-Quesada, F. R. (2019b). A logical study of group-size based social network creation. *Journal of Logical and Algebraic Methods in Programming*, 106, 117–140.

Smets, S. and Velázquez-Quesada, F. R. (2020). A closeness-and priority-based logical study of social network creation. *Journal of Logic*, *Language and Information*, 29(1), 21–51.

Solaki, A. (2021). *Logical models for bounded reasoners*. PhD thesis, University of Amsterdam.

Solaki, A., Terzopoulou, Z., and Zhao, B. (2016). Logic of closeness revision: Challenging relations in social networks. In *Proceedings of the ESSLLI*, 123–134.

Stalnaker, R. (1991). The problem of logical omniscience, I. *Synthese*, 89(3), 425– 440.

van Benthem, J. (2007). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17, 129–155.

van Benthem, J. (2008). Logic and reasoning: Do the facts matter? *Studia Logica: An International Journal for Symbolic Logic*, 88(1), 67–84.

van Benthem, J. (2011). *Logical dynamics of information and interaction*. Cambridge University Press.

van Ditmarsch, H. and Labuschagne, W. (2007). My beliefs about your beliefs: A case study in theory of mind and epistemic logic. *Synthese*, 155(2), 191–209.

van Ditmarsch, H., van der Hoek, W., and Kooi, B. (2008). *Dynamic epistemic logic*. Springer.

Verbrugge, R. (2009). Logic and social cognition. *Journal of Philosophical Logic*, 38(6), 649–680.

Vīķe-Freiberga, V., Däubler-Gmelin, H., Hammersley, B., and Maduro, L. M. P. P. (2013). A free and pluralistic media to sustain european democracy. European Commission.

Wiewiórowski, W. (2024). TechDispatch European Data Protection Supervisor: Explainable Artificial Intelligence. https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2023-11-16-techdispatch-22023-explainable-artificial-intelligence_en.

Young Pedersen, M., Smets, S., and Ågotnes, T. (2021). Modal logics and group polarization. *Journal of Logic and Computation*, 31(8), 2240–2269.

Zhou, X., Xu, Y., Li, Y., Josang, A., and Cox, C. (2012). The state-of-the-art in personalized recommender systems for social networking. *Artificial Intelligence Review*, 37, 119–132.

About the Author

Gaia Belardinelli received her Ph.D. degree at the University of Copenhagen in 2023. She is currently a postdoctoral fellow at Stanford University in California, USA.

✉ gaiabel@stanford.edu

Lei Li received a Ph.D. degree jointly awarded by Tsinghua University and the University of Amsterdam in 2023. He is currently a Lecturer at Shaanxi Normal University.

✉ lileity@snnu.edu.cn

Sonja Smets received her Ph.D. degree at the Vrije Universiteit Brussel in 2001, she is currently a professor of logic and epistemology at the Institute for Logic, Language and Computation at the University of Amsterdam in the Netherlands.

✉ s.j.l.smets@uva.nl

Anthia Solaki received her Ph.D. degree at the University of Amsterdam (Institute for Logic, Language, and Computation) in 2021 and she is currently working as a research scientist at the Netherlands Organization for Applied Scientific Research (TNO). The research included in this manuscript was initiated at ILLC, University of Amsterdam and pursued independently since.

✉ ansolaki@gmail.com