A Logic of General Attention Using Edge-Conditioned Event Models

Gaia Belardinelli[1] , **Thomas Bolander**[2] , **Sebastian Watzl**[3]
[1]Stanford University
[2]Technical University of Denmark
[3]University of Oslo
gaiabel@stanford.edu, tobo@dtu.dk, sebastian.watzl@ifikk.uio.no

Abstract

In this work, we present the first general logic of attention. Attention is a powerful cognitive ability that allows agents to focus on potentially complex information, such as logically structured propositions, higher-order beliefs, or what other agents pay attention to. This ability is a strength, as it helps to ignore what is irrelevant, but it can also introduce biases when some types of information or agents are systematically ignored. Existing dynamic epistemic logics for attention cannot model such complex attention scenarios, as they only model attention to atomic formulas. Additionally, such logics quickly become cumbersome, as their size grows exponentially in the number of agents and announced literals. Here, we introduce a logic that overcomes both limitations. First, we generalize edge-conditioned event models, which we show to be as expressive as standard event models yet exponentially more succinct (generalizing both standard event models and generalized arrow updates). Second, we extend attention to arbitrary formulas, allowing agents to also attend to other agents' beliefs or attention. Our work treats attention as a modality, like belief or awareness. We introduce attention principles that impose closure properties on that modality and that can be used in its axiomatization. Throughout, we illustrate our framework with examples of AI agents reasoning about human attention, demonstrating how such agents can discover attentional biases.

1 Introduction

Attention is the crucial ability of the mind to select and prioritize specific subsets of available information [Watzl, 2017]. Research in psychology suggests that agents who do not pay attention to something do not update their beliefs about it [Simons and Chabris, 1999]. Therefore, while the ability to focus attention is a strength, as it helps agents ignore irrelevant information, restricting which information is learned can also introduce biases [Johnson, 2024]: an agent with limited attention may never learn about certain news, some aspects of a candidate's application, or what certain individuals (e.g. of a specific gender or nationality) have to say. Researchers have begun to investigate the significance of attention-driven biases in both humans and AI [Johnson, 2024; Munton, 2023]. AI systems that can robustly reason about attention could potentially detect such attentional biases and correct them.

As a step toward that aim, we provide the first logic of general attention, here captured in a dynamic epistemic logic (DEL) framework [Baltag *et al.*, 1998]. Attention is treated as a modality that restricts which parts of an event an agent learns. An agent who is not attending to a formula will mistake an event containing that formula for one that does not. Our work builds on earlier proposals that introduce DEL models of attention to capture its effects on agents' beliefs. However, the first proposal only modeled a notion of all-or-nothing attention [Bolander *et al.*, 2016], and later work extended it to apply to atomic propositions [Belardinelli and Bolander, 2023]. While the latter can treat cases where a member of a hiring committee systematically attends only to, say, the research parts of applicants' CVs, it cannot represent agents whose attention is biased against specific agents, e.g. paying attention to what candidate a has to say about a research topic, but not candidate b [Munton, 2023; Smith and Archer, 2020]. Or consider the distinction between an agent who aims to learn by paying attention to the world and an agent who aims to learn by attending to what others pay attention to – the latter arguably being an important component of social learning [Rendell *et al.*, 2010; Boyd *et al.*, 2011]. To model such cases, we need a notion of attention that applies to arbitrary formulas.

A key obstacle to such a richer notion is that models for attention quickly become highly complex and their size grows at least exponentially large, even for all-or-nothing attention. A fully general theory requires technical innovations that allow to capture attention and its impact on beliefs in a tractable way. To that goal, we adopt an edge-conditioned version of DEL [Bolander, 2018] and generalize it so that every edge carries both a source and a target condition. Source conditions are used to encode an agent's current attentional state, while target conditions specify what the agent learns about her own attention. We show that this single modification allows for an exponentially more succinct reformulation of earlier event models for propositional attention [Belardinelli and Bolander, 2023]. We also show that edge-conditioned event models are as expressive as standard event models, establishing exponential succinctness of edge-conditioned event models over standard event models. We further provide an axiomatization of edge-conditioned models and show that they are as expressive as generalized arrow updates [Kooi and Renne, 2011] and as succinct as them. Taken together, these results imply that edge-conditioned models provide a novel event model formalism that generalizes and unifies the two common alternative formalisms in DEL: standard event models and generalized arrow updates. Due to these properties, it seems a good candidate for a new standard formalism in DEL.

Thanks to the increased simplicity and clarity of the new formalisation, we can now generalize the framework along another dimension, namely to accommodate attention to arbitrary formulas (not merely atomic ones). In this richer framework attention becomes a modality, in that way resembling belief or awareness (though, unlike the latter, attention is intrinisically dynamic as it affects what agents learn from events, and it is not used to restrict their standing beliefs). We argue that the behaviour of that modality can be governed by a family of attention principles, namely closure conditions that can be used to axiomatize a specific logic of attention. The logic of general attention can model agents who attend to the world, to what other agents believe, or to what other agents attend to. It can thus be used to study social attention and learning and the attentional biases mentioned above in potentially complex multi-agent settings.

In summary, the paper makes the following contributions:

1. *We present the first general theory of attention*, allowing agents to attend to arbitrary formulas.
2. *We provide the first event model formalism that generalizes and unifies standard event models and generalized arrow updates*. We show that this new formalism is always at least as succinct as both, and can be exponentially more succinct.
3. *We provide the first principles for general attention* that can be used to axiomatize specific attention notions.

Due to page limits, we only provide proof sketches. *All full proofs can be found in the extended versions on Arxiv: https://arxiv.org/pdf/2505.14539*.

2 DEL and Propositional Attention

We are going to work with multiple distinct languages, in all of which we use Ag to denote a finite set of *agents* and P to denote a finite set of *propositional atoms*.[1] The symbol $\mathcal{L}$ is used to denote any language extending the *language of epistemic logic* $\mathcal{L}_{\text{EL}}$ given by the grammar:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_a\varphi,$$

with $p \in P$, $a \in Ag$, and $B_a\varphi$ reads "agent a believes φ". We define the other propositional connectives in the standard way. Every language $\mathcal{L}$ has a set of *atoms* $At(\mathcal{L})$ with $At(\mathcal{L}) \supseteq P$. A *literal* is an element of $At(\mathcal{L}) \cup \{\neg p\colon p \in At(\mathcal{L})\}$, and a *propositional literal* is an element of $P \cup \{\neg p\colon p \in P\}$. For $p \in At(\mathcal{L})$, we denote a literal by $\ell(p)$, where either $\ell(p) = p$ or $\ell(p) = \neg p$. For any formula φ, $At(\varphi)$ denotes the set of atoms appearing in it.

[1]So all our languages take the sets Ag and P as parameters, but this dependency is kept implicit throughout the paper.

DEL We now introduce the standard ingredients of DEL [Baltag *et al.*, 1998; van Ditmarsch *et al.*, 2007].

Definition 1 (Kripke model)**.** *A* Kripke model *for $\mathcal{L}$ is a tuple $\mathcal{M} = (W, R, V)$ where $W \neq \emptyset$ is a finite set of* worlds*, $R : Ag \to \mathcal{P}(W^2)$ assigns an* accessibility relation *R_a to each agent $a \in Ag$, and $V : W \to \mathcal{P}(At(\mathcal{L}))$ is a* valuation function*. Where w is the* actual world*, we call $(\mathcal{M}, w)$ a* pointed Kripke model*.*

Definition 2 (Standard event model)**.** *A* standard event model *for $\mathcal{L}$ is a tuple $\mathcal{E} = (E, Q, pre)$ where $E \neq \emptyset$ is a finite set of* events*, $Q : Ag \to \mathcal{P}(E^2)$ assigns an* accessibility relation *Q_a to each agent $a \in Ag$, and $pre : E \to \mathcal{L}$ assigns a* precondition *to each event. The set of event models for $\mathcal{L}$ is denoted by $\mathscr{E}(\mathcal{L})$. Where $E_d \subseteq E$ is a set of* actual events*, $(\mathcal{E}, E_d)$ is a* multi-pointed standard event model*. If $E_d = \{e\}$ for some* actual event *$e \in E$, then $(\mathcal{E}, E_d)$ is called a* pointed standard event model*, also denoted $(\mathcal{E}, e)$.*[2]

Intuitively, a (pointed) Kripke model represents an epistemic state, while a (pointed) event model represents an epistemic action or event happening. The *product update operator*, defined next, expresses how an epistemic state is updated as the consequence of an epistemic event.

Definition 3 (Standard product update)**.** *Let $\mathcal{M} = (W, R, V)$ be a Kripke model and $\mathcal{E} = (E, Q, pre)$ an event model, both for the same language $\mathcal{L}$. The* product update *of $\mathcal{M}$ with $\mathcal{E}$ is $\mathcal{M} \otimes \mathcal{E} = (W', R', V')$ where:*[3]

$W' = \{(w, e) \in W \times E\colon (\mathcal{M}, w) \vDash pre(e)\}$,
$R'_a = \{((w,e), (v,f)) \in (W')^2\colon (w,v) \in R_a, (e,f) \in Q_a\}$,
$V'((w, e)) = \{p \in At(\mathcal{L})\colon p \in V(w)\}$.

Given a pointed Kripke model $(\mathcal{M}, w)$ and a pointed or multi-pointed event model $(\mathcal{E}, E_d)$, we say that $(\mathcal{E}, E_d)$ is applicable *in $(\mathcal{M}, w)$ iff there exists a unique $e' \in E_d$ such that $(\mathcal{M}, w) \vDash pre(e')$. In that case, we define the* product update *of $(\mathcal{M}, w)$ with $(\mathcal{E}, E_d)$ as the pointed Kripke model $(\mathcal{M}, w) \otimes (\mathcal{E}, E_d) = (\mathcal{M} \otimes \mathcal{E}, (w, e'))$.*

We define the language of *DEL with standard event models* $\mathcal{L}_{\text{DEL}}$ as the language given by the grammar of $\mathcal{L}_{\text{EL}}$ extended with the clause $\varphi ::= [\mathcal{E}]\varphi$, where $\mathcal{E}$ is a pointed or multi-pointed standard event model. The formula $[\mathcal{E}]\varphi$ reads "after $\mathcal{E}$ happens, φ is the case". Notice that $At(\mathcal{L}_{\text{DEL}}) = P$.

Definition 4 (Satisfaction)**.** *Let $(\mathcal{M}, w) = ((W, R, V), w)$ be a pointed Kripke model for $\mathcal{L}_{DEL}$. Satisfaction of $\mathcal{L}_{DEL}$-formulas in $(\mathcal{M}, w)$ is given by the following clauses extended with the standard clauses for propositional logic:*

$(\mathcal{M}, w) \vDash p$	*iff*	$p \in V(w)$*, where* $p \in At(\mathcal{L}_{DEL})$
$(\mathcal{M}, w) \vDash B_a\varphi$	*iff*	$(\mathcal{M}, v) \vDash \varphi$ *for all* $(w, v) \in R_a$
$(\mathcal{M}, w) \vDash [\mathcal{E}]\varphi$	*iff*	*if* $\mathcal{E}$ *is applicable in* $(\mathcal{M}, w)$ *then* $(\mathcal{M}, w) \otimes \mathcal{E} \vDash \varphi$.

We say that a formula φ is valid *if $(\mathcal{M}, w) \vDash \varphi$ for all pointed Kripke models $(\mathcal{M}, w)$, and in that case we write $\vDash \varphi$.*

[2]We often denote event models by $\mathcal{E}$ independently of whether we refer to an event model (E, Q, pre), a pointed event model $((E, Q, pre), e)$ or a multi-pointed one, $((E, Q, pre), E_d)$. Their distinction will be clear from context.

[3]The meaning of $(\mathcal{M}, w) \vDash pre(e)$ depends on the semantics of $\mathcal{L}$. Semantics of specific languages $\mathcal{L}$ are introduced later.

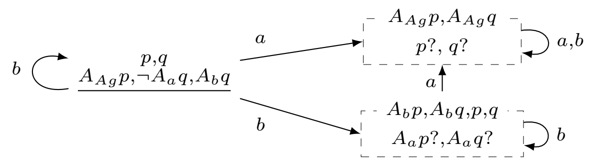

Figure 1: A pointed Kripke model $(\mathcal{M}, w)$ for $\mathcal{L}_{\text{PA}}$. In the figure, p stands for "the applicant has published several papers in top-tier journals", q for "the applicant has made significant contributions to diversity". There are two agents, Ann (a) and the AI agent (b), i.e. $Ag = \{a, b\}$. We use the following conventions. For a set of agents Ag', $A_{Ag'}p := \bigwedge_{a \in Ag'} A_a p$. Worlds are represented either by a sequence of literals true at the world, or by such a sequence of literals where some of the atoms are followed by question marks: $p_1?, \dots, p_n?, \ell(q_1), \dots, \ell(q_m)$. This is shorthand for the set of 2^n worlds corresponding to all possible truth-value assignments of the atoms $p_1, p_2, ..., p_n$, where $\ell(q_1), ..., \ell(q_m)$ are true at each of these worlds. When a world appears inside a dashed box, all the literals in the label of that box are also true at the world. The actual world is underlined. The accessibility relations are represented by labelled arrows. An arrow from (or to) the border of a box means that there is an arrow from (or to) all the events inside the box.

The Language of Propositional Attention We add new atomic proposition $A_a p$ called an *attention atom*, for all $a \in Ag$ and $p \in P$. The set of *attention atoms* is $H = \{A_a p \colon p \in P, a \in Ag\}$. The *language of propositional attention* $\mathcal{L}_{\text{PA}}$ is the language given by the grammar of $\mathcal{L}_{\text{DEL}}$ extended with the following clause: $\varphi ::= A_a p$, where $A_a p \in H$ [Belardinelli and Bolander, 2023]. We read $A_a p$ as "agent a pays attention to whether p".[4] We have $At(\mathcal{L}_{\text{PA}}) = P \cup H$. Satisfaction of formulas in $\mathcal{L}_{\text{PA}}$ is exactly as in Definition 4, except we replace $\mathcal{L}_{\text{DEL}}$ by $\mathcal{L}_{\text{PA}}$ everywhere (the difference only amounts to the addition of the new set of atoms H).

Example 1. *Ann is about to review the CV of an applicant for a position in her lab. To ensure an unbiased hiring process, her lab has adopted an AI agent to detect the presence of attentional biases and assess whether some aspects of the candidates' CVs receive more attention than others. The AI agent does not have any information about which aspects Ann is prioritizing, and only knows that she has not read the CV yet and so has no information about it. However, it correctly assumes that Ann overestimates her own attentional capacities: Ann believes that she (and everybody else) will pay attention to every aspect of the CV. In reality, Ann focuses more on the candidate's research track record, and does not pay so much attention to other important factors such as contributions to diversity. The AI agent has already read and paid attention to all aspects of the CV. This situation is represented in Fig. 1. We have e.g.* $(\mathcal{M}, w) \vDash A_a p \wedge \neg A_a q$*: Ann (a) pays attention to p and not to q. We also have* $(\mathcal{M}, w) \vDash B_b B_a(A_a p \wedge A_a q)$*: the AI agent (b) thinks that Ann believes to be paying attention to all aspects of the CV.*

Fig. 1 models the external perspective on the situation, for simplicity of presentation. To represent the case where it is the AI agent b itself using the logic to reason about the scenario, we would apply perspective shifts and consider the situation from the perspective of b only [Bolander *et al.*, 2021].

Standard Event Models for Propositional Attention The propositional event models of Belardinelli and Bolander [2023] are intended to represent the revelation of concurrent stimuli from the environment, of which an agent may pay attention to and receive only a portion.[5] Accordingly, a revelation is represented as a conjunction of propositional literals, that is, the conjunction of facts being revealed by the relevant stimuli. An agent may pay attention to only some aspects of the revealed stimuli and, hence, receive only some of these facts. Say agents a and b are shown a picture of c wearing a blue hat and a red shirt. Use p_1 for "c is wearing a blue hat" and p_2 for "c is wearing a red shirt". Then the formula revealed by the picture is $p_1 \wedge p_2$, with the intuition that $p_1 \wedge p_2$ *contains* the two pieces of information p_1 and p_2, and that a and b may attend to and receive different parts of it (e.g. agent a may only pay attention to the hat, p_1, and agent b only to the shirt, p_2).

We use the following additional conventions. For a set of formulas S, $\bigwedge S$ denotes their conjunction. For $S = \emptyset$, define $\bigwedge S := \top$. Conjunctions of literals are assumed to be in a normal form where each atom occurs at most once and in a specific order. A conjunction of literals $\varphi = \bigwedge_{1 \leq i \leq n} \ell_i$ is said to *contain* the literals $\ell_1, \dots, \ell_n$, and we write $\ell_i \in \varphi$.

The next definition recalls the propositional attention event models [Belardinelli and Bolander, 2023]. The function id_E used in the definition is the identity function on E, that is, for every $e \in E$, $id_E(e) = e$.

Definition 5 (Event model for propositional attention $\mathcal{F}(\varphi)$ [Belardinelli and Bolander, 2023])**.** *Let* $\varphi = \ell(p_1) \wedge \dots \wedge \ell(p_n)$ *with all* $p_i \in P$*. The* event model for propositional attention *representing the revelation of* φ *is the multi-pointed event model* $\mathcal{F}(\varphi) = ((E, Q, id_E), E_d)$ *for* $\mathcal{L}_{\text{PA}}$ *defined by:*

$$E = \{ \bigwedge_{p \in S} \ell(p) \wedge \bigwedge_{a \in Ag} \big(\bigwedge_{p \in X_a} A_a p \wedge \bigwedge_{p \in S \setminus X_a} \neg A_a p \big) \colon \\ S \subseteq At(\varphi) \text{ and for all } a \in Ag, X_a \subseteq S \}$$

Q_a *is such that* $(e, f) \in Q_a$ *iff the following holds for all* p*:*

- ATTENTIVENESS*: if* $A_a p \in e$ *then* $A_a p, \ell(p) \in f$*;*
- INERTIA*: if* $A_a p \notin e$ *then* $\ell(p) \notin f$*;*

$E_d = \{\psi \in E \colon \ell(p) \in \psi$*, for all* $\ell(p) \in \varphi\}$.

The definition is exemplified in Fig. 2. These event models grow exponentially in the number of revealed literals and the number of agents [Belardinelli and Bolander, 2023]. We now introduce *edge-conditioned event models*, which allow a more compact representation and easier generalization.

3 Edge-Conditioned Event Models

The idea behind edge-conditioned event models is to make the edges of event models conditional on formulas. For stan-

[4] Belardinelli and Bolander [2023] used instead the notation $h_a p$ derived from the "hearing" atoms h_a of Bolander *et al.* [2016]. The new notation $A_a p$ fits better with our generalizations in Section 5.

[5] We choose the term *revelation* over the term *announcement* to emphasize that the information is not necessarily disclosed by an agent (announced), but may simply be something that is seen or heard [van Ditmarsch, 2023]. Revelations are here always truthful.

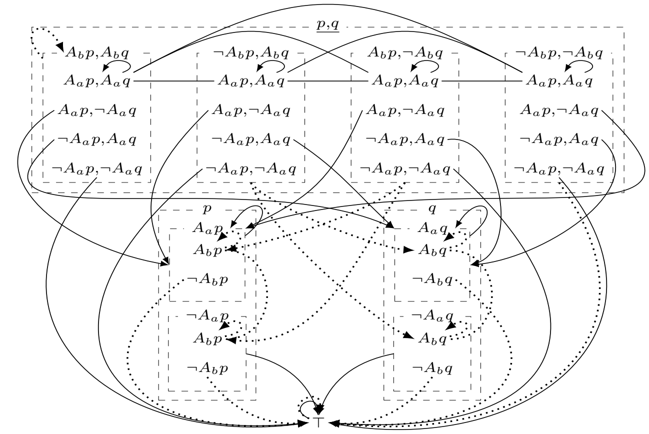

Figure 2: Event model $\mathcal{F}(p \wedge q)$ from [Belardinelli and Bolander, 2023]. Solid edges are for Ann (a), dotted for the AI agent (b). The figure adopts the same conventions as in the cited paper: An event is represented by a list of literals, corresponding to (some of the) literals that appear in the conjunctive precondition of the event itself. The formulas in the label of a dashed box are to be included as conjuncts in the precondition of all events inside the box. The convention for the edges is the same as in Fig. 1. See [Belardinelli and Bolander, 2023] for a detailed explanation of the figure.

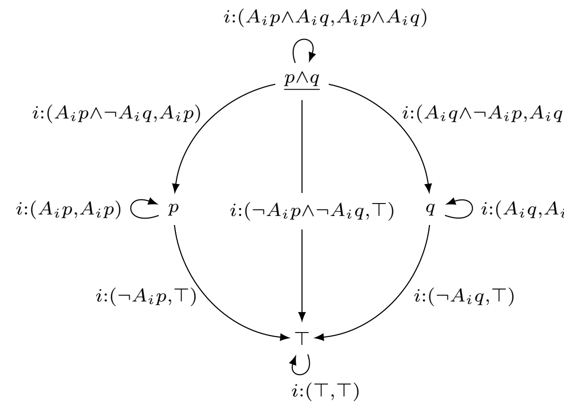

Figure 3: Edge-conditioned event model for propositional attention $\mathcal{H}(p \wedge q)$. Events are represented by conjunctive formulas corresponding to the event's own precondition. When for all agents $i \in Ag$, we have a (conditioned) edge $(e{:}\varphi_i, f{:}\psi_i) \in Q_i$, we add an arrow from e to f labelled by $i{:}(\varphi_i, \psi_i)$. This means that agent i has an edge from e to f with source condition φ_i and target condition ψ_i. For example, the arrow from event $p \wedge q$ to event p labelled by $i{:}(A_ip \wedge \neg A_iq, A_ip)$ corresponds to the edge $(p \wedge q{:}A_ip \wedge \neg A_iq, p{:}A_ip) \in Q_i$, for all $i \in Ag$. This edge models an agent who paid attention to p, but not to q, and therefore only learns p and that she paid attention to p.

dard event models, $(e, f) \in Q_a$ means that event f is accessible from event e by agent a, and whenever that is the case, we draw an a-edge from e to f. In edge-conditioned event models, what is accessible at an event has become conditioned by formulas: $(e{:}\varphi, f{:}\psi) \in Q_a$ means that f is accessible from e by a under the condition that φ is the case at the source e and ψ is the case at target f. As we will see, this simple modification has rather advantageous consequences, as it allows us to represent event models significantly more succinctly.

Definition 6 (Edge-conditioned event models)**.** *An* edge-conditioned event model *for $\mathcal{L}$ is a tuple $\mathcal{C} = (E, Q, pre)$ where E and pre are standard (i.e. as in Def. 2), and where $Q : Ag \to \mathcal{P}(E \times \mathcal{L} \times E \times \mathcal{L})$ assigns to each agent a set of quadruples (e, φ, f, ψ). For $(e, \varphi, f, \psi) \in Q_a$, we call (e, φ, f, ψ) a* conditioned edge*, where φ is the* source condition *(at e) and ψ is the* target condition *(at f). We often abbreviate (e, φ, f, ψ) as $(e{:}\varphi, f{:}\psi)$ to emphasize that it is an edge from e to f, where the source e has condition φ, and the target f condition ψ. The set of edge-conditioned event models for $\mathcal{L}$ is denoted by $\mathscr{C}(\mathcal{L})$. Where e is the* actual event*, $(\mathcal{C}, e)$ is a* pointed edge-conditioned event model*.*

Definition 7 (Edge-conditioned product update)**.** *Let $\mathcal{M} = (W, R, V)$ be a Kripke model and $\mathcal{C} = (E, Q, pre)$ an edge-conditioned event model, both for the same language $\mathcal{L}$. The* product update *of $\mathcal{M}$ with $\mathcal{C}$ is $\mathcal{M} \otimes \mathcal{C} = (W', R', V')$ where W' and V' are standard (Def. 3), and $R'_a = \{((w, e), (v, f)) \in (W')^2 \colon (w, v) \in R_a$ and $\exists \varphi, \psi \in \mathcal{L}$ such that $(e{:}\varphi, f{:}\psi) \in Q_a, (\mathcal{M}, w) \models \varphi$ and $(\mathcal{M}, v) \models \psi\}$. A pointed edge-conditioned event model $(\mathcal{C}, e)$ is called* applicable *in a pointed Kripke model $(\mathcal{M}, w)$ if $(\mathcal{M}, w) \models pre(e)$, and then the* product update *of $(\mathcal{M}, w)$ with $(\mathcal{C}, e)$ is the pointed Kripke model $(\mathcal{M}, w) \otimes (\mathcal{C}, e) = (\mathcal{M} \otimes \mathcal{C}, (w, e))$.*

Edge-conditioned event models were first introduced by Bolander [2018], but only with conditions at the source of edges. Adding conditions also at the target is technically straightforward and matches the conventions for generalized arrow updates [Kooi and Renne, 2011]. Despite the technical simplicity, it turns out to have significant advantages, including that this new formalism generalizes and is more succinct than both generalized arrow updates and standard event models, and that we acheive an exponential succinctness result for the event model for propositional attention that would not hold with only source conditions (Theorem 6 below).

We define the language of *DEL with edge-conditioned event models* $\mathcal{L}_{\text{ECM}}$ as the language given by the grammar of $\mathcal{L}_{\text{EL}}$ extended with the clause $[\mathcal{C}]\varphi$, where $\mathcal{C}$ is a pointed edge-conditioned event model. The semantics of $\mathcal{L}_{\text{ECM}}$ is defined as for $\mathcal{L}_{\text{DEL}}$ (Def. 4), except that the product update in the semantics of $[\mathcal{C}]\varphi$ uses Def. 7.

Theorem 1 (Soundness and completeness)**.** *Table 1 provides*

All prop. tautologies and $B_a(\varphi \to \psi) \to (B_a\varphi \to B_a\psi)$.
$[(\mathcal{C}, e)]p \leftrightarrow (pre(e) \to p)$
$[(\mathcal{C}, e)]\neg\psi \leftrightarrow (pre(e) \to \neg[(\mathcal{C}, e)]\psi)$
$[(\mathcal{C}, e)](\psi \wedge \chi) \leftrightarrow ([(\mathcal{C}, e)]\psi \wedge [(\mathcal{C}, e)]\chi)$
$[(\mathcal{C}, e)]B_a\psi \leftrightarrow (pre(e) \to \bigwedge_{(e:\chi, f:\chi') \in Q_a} (\chi \to B_a(\chi' \to [(\mathcal{C}, f)]\psi)))$
From φ and $\varphi \to \psi$, infer ψ. From φ infer $B_a\varphi$.
From $\varphi \leftrightarrow \psi$, infer $\chi[\varphi/p] \leftrightarrow \chi[\psi/p]$ (substitution).

Table 1: Axiomatization of DEL with edge-conditions ($\mathcal{L}_{\text{ECM}}$).

a sound and complete axiomatization of DEL with edge-conditioned event models.

Proof Sketch. The first line and last two lines of Table 1 provide a sound and complete axiomatization of the logic of the underlying epistemic language $\mathcal{L}_{\text{EL}}$ [Fagin *et al.*, 1995]. Completeness of the logic of the full language $\mathcal{L}_{\text{ECM}}$ then follows by standard reduction arguments [van Ditmarsch *et al.*, 2007]: we have reduction axioms (lines 2–5 of Table 1) for translating any formula involving the $[\mathcal{C}]$ modality into a formula without it, hence reducing the completeness proof to completeness of the underlying epistemic logic. For soundness, we verify the validity of the reduction axiom for the belief modality, the validity of the others being standard. □

We define $\mathcal{L}_{\text{PA}+}$ as the language given by the grammar of $\mathcal{L}_{\text{ECM}}$ extended with the clause $\varphi ::= A_a p$, with $A_a p \in H$. It is the language that uses edge-conditioned event models and includes the attention atoms. Note that $At(\mathcal{L}_{\text{PA}+}) = At(\mathcal{L}_{\text{PA}}) = P \cup H$. We exemplify Definition 6 by defining an edge-conditioned event model for propositional attention that corresponds to the standard event model of Definition 5.

Definition 8 (Edge-conditioned event model for propositional attention)**.** *Let $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$ with $p_i \in P$. The* edge-conditioned event model for propositional attention *representing the revelation of φ is the pointed edge-conditioned event model $\mathcal{H}(\varphi) = ((E, Q, id_E), \varphi)$ for $\mathcal{L}_{PA+}$ defined by:*

$$E = \{\textstyle\bigwedge_{p \in S} \ell(p) \colon S \subseteq At(\varphi)\},$$

$$Q_a = \{(\bigwedge_{p\in S} \ell(p) \colon \bigwedge_{p \in T} A_a p \wedge \bigwedge_{p \in S \setminus T} \neg A_a p, \bigwedge_{p \in T} \ell(p) \colon \bigwedge_{p \in T} A_a p) : T \subseteq S \subseteq At(\varphi)\}.$$

Figure 3 shows $\mathcal{H}(p \wedge q)$, i.e. the same revelation as in Figure 2 (when $Ag = \{a, b\}$). The standard and the edge-conditioned event models for propositional attention correspond to each other by being *update equivalent*.

Definition 9 (Update equivalence [Kooi and Renne, 2011; van Eijck *et al.*, 2012])**.** *Let $\mathcal{D}$ be a standard or edge-conditioned event model for $\mathcal{L}$ and $\mathcal{D}'$ a standard or edge-conditioned event model for $\mathcal{L}'$, where $At(\mathcal{L}) = At(\mathcal{L}')$. We say that $\mathcal{D}$ is* update equivalent *to $\mathcal{D}'$ if for all Kripke models $\mathcal{M}$ with atom set $At(\mathcal{L})$, $\mathcal{M} \otimes \mathcal{D}$ and $\mathcal{M} \otimes \mathcal{D}'$ are bisimilar.*[6]

Update equivalence entails semantic equivalence, as bisimilar models agree on all formulas [Blackburn *et al.*, 2001].

Theorem 2. *For any conjunction of propositional literals φ, $\mathcal{F}(\varphi)$ and $\mathcal{H}(\varphi)$ are update equivalent.*

Proof Sketch. Let $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$ and let $\mathcal{M}$ be any Kripke model. We define a bisimulation relation from $\mathcal{M} \otimes \mathcal{H}(\varphi) = (W, R, V)$ to $\mathcal{M} \otimes \mathcal{F}(\varphi) = (W', R', V')$ by $Z = \{((w, e), (w', e')) \in W \times W' \mid w = w'$ and for all $i = 1, \ldots, n : \ell(p_i) \in e$ iff $\ell(p_i) \in e'\}$. In other words, we match each event $\bigwedge_{p \in S} \ell(p)$ of $\mathcal{H}(\varphi)$ with all events of $\mathcal{F}(\varphi)$ that are of the form $\bigwedge_{p \in S} \ell(p) \wedge \cdots$. Then the proof proceeds by showing that Z satisfies the [Atom], [Forth] and [Back] conditions of being a bisimulation relation. □

[6]If $\mathcal{M}$ is a Kripke model for $\mathcal{L}$, it is also a Kripke model for any $\mathcal{L}'$ with $At(\mathcal{L}') = At(\mathcal{L})$, since its definition only depends on the set of atoms and agents, and the agent set is fixed. We also refer to such Kripke models as *Kripke models with atom set $At(\mathcal{L})$*. The notion of bisimulation is standard in modal logic [Blackburn *et al.*, 2001]. The definition is given in the Technical Appendix.

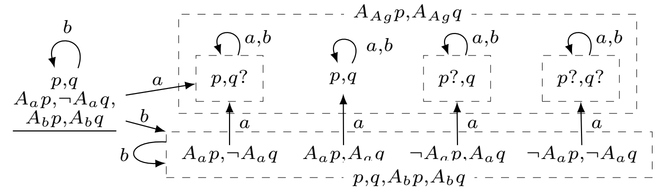

Figure 4: The pointed Kripke model $(\mathcal{M} \otimes \mathcal{H}(p \wedge q), (w, e))$ for $\mathcal{L}_{\text{PA}+}$, with $\mathcal{H}(p \wedge q)$ given in Fig. 3. We use the same conventions as before and omit worlds that are inaccessible by all agents.

This theorem implies that the event models of Figs. 2 and 3 are equivalent, but the latter is clearly simpler and easier to follow. As it turns out, it is also exponentially smaller (Thm. 6). Moreover, while $\mathcal{F}(\varphi)$ is multi-pointed, $\mathcal{H}(\varphi)$ is only pointed. Hence, by using edge-conditioned event models, we achieve several advantages in terms of simplicity and succinctness. There are also conceptual advantages. By separating the preconditions from the edge-conditions, these models introduce a conceptual distinction between the informational content pertaining to the event itself, which belongs to the event's preconditions, and information determining which events are considered possible by an agent, which is now a condition on the accessibility relation. Information about agents' attention is of the latter kind, as it pertains to the agent's perspective on the event (see also [Watzl, 2017], Ch. 13). It is not a part of the information that is revealed, but rather a condition on the accessibility relations (i.e., an edge-condition). Standard event models instead merge these distinct information types into a single construct, namely the preconditions.

Example 2. *Continuing Example 1, Ann has now reviewed the applicant's CV. She has learned that they published several papers in top-tier journals, but has not learned about their contribution to diversity, as she did not pay attention to it. The AI agent still has no information about what Ann paid attention to, and so thinks that Ann may have learned any part of the CV. This situation is depicted in Fig. 4, where e.g. $(\mathcal{M} \otimes \mathcal{H}(p \wedge q), (w, e)) \vDash B_a p \wedge \neg B_a q \wedge \neg B_a \neg q$: Ann has only learned about the candidate's publications, overlooking their contributions to diversity. We also have $(\mathcal{M} \otimes \mathcal{H}(p \wedge q), (w, e)) \vDash \neg B_b A_a p \wedge \neg B_b \neg A_a p \wedge \neg B_b A_a q \wedge \neg B_b \neg A_a q$: the AI agent still has no information about Ann's attention.*

4 Expressivity and Succinctness

Before introducing our logic of general attention, we examine key properties of edge-conditioned event models, which may be of independent interest for DEL and its applications.

Definition 10 (Transformation of standard event models into edge-conditioned event models)**.** *By mutual recursion, we define mappings $T_1 : \mathscr{E}(\mathcal{L}_{DEL}) \to \mathscr{C}(\mathcal{L}_{ECM})$ and $T_2 : \mathcal{L}_{DEL} \to$*

$\mathcal{L}_{ECM}$. Define T_1 by $T_1(E, Q, pre) = (E, Q', pre')$ where for each $a \in Ag$, $Q'_a = \{(e{:}\top, f{:}\top) \colon (e, f) \in Q_a\}$, and for each $e \in E$, $pre'(e) = T_2(pre(e))$. Define T_2 by:

$T_2(p) = p$, *for* $p \in P$ $\quad T_2(\neg\varphi) = \neg T_2(\varphi)$
$T_2(\varphi \wedge \psi) = T_2(\varphi) \wedge T_2(\psi)$ $\quad T_2(B_a\varphi) = B_a T_2(\varphi)$
$T_2([(\mathcal{E}, e)]\varphi) = [(T_1(\mathcal{E}), e)]T_2(\varphi)$

Theorem 3. *$\mathcal{E} \in \mathscr{E}(\mathcal{L}_{DEL})$ is update equivalent to $T_1(\mathcal{E})$.*

Proof Sketch. We show that for any Kripke model $\mathcal{M}$ and $\mathcal{E} \in \mathscr{E}(\mathcal{L}_{\text{DEL}})$, $\mathcal{M} \otimes \mathcal{E}$ is isomorphic to $\mathcal{M} \otimes T_1(\mathcal{E})$, i.e. they have the same worlds and accessibility relations. The proof is by induction on the depth of nesting of event model modalities inside $\mathcal{E}$ [Kooi and Renne, 2011]. It is straightforward, as the translation T_1 only replaces edges (e, f) with edges $(e{:}\top, f{:}\top)$ having trivial source and target conditions.[7] □

Theorem 3 shows that any standard event model can be equivalently represented as an edge-conditioned one. Hence, $\mathcal{L}_{\text{ECM}}$ is at least as expressive as $\mathcal{L}_{\text{DEL}}$. It turns out they are equally expressive, as any edge-conditioned event model can also be transformed into an equivalent standard one, using a construction inspired by Kooi and Renne [2011, Thm 4.7]:

Definition 11 (Transformation of edge-conditioned event models into standard event models)**.** *For any edge-conditioned event model (E, Q, pre) for $\mathcal{L}_{ECM}$ and any $e \in E$, let $\Phi(e)$ be the set of source and target conditions at e, i.e. $\Phi(e) = \{\varphi \in \mathcal{L}_{ECM} \colon (e{:}\varphi, f{:}\psi) \in Q_a\} \cup \{\psi \in \mathcal{L}_{ECM} \colon (f{:}\varphi, e{:}\psi) \in Q_a\}$. Set $\Phi'(e) = \Phi(e) \cup \{\neg\varphi \colon \varphi \in \Phi(e)\}$ and let $\mathsf{mc}(e)$ denote the set of maximally consistent subsets of $\Phi'(e)$. By mutual recursion, we define mappings $T'_1 : \mathscr{C}(\mathcal{L}_{ECM}) \to \mathscr{E}(\mathcal{L}_{DEL})$ and $T'_2 : \mathcal{L}_{ECM} \to \mathcal{L}_{DEL}$. The mapping T'_2 is as T_2 of Definition 10, except we replace T_i by T'_i and $\mathcal{E}$ by $\mathcal{C}$. We define T'_1 by $T'_1(E, Q, pre) = (E', Q', pre')$ where:*

$E' = \{(e, \Gamma) \colon e \in E, \Gamma \in \mathsf{mc}(e)\}$,
$Q'_a = \{((e, \Gamma), (e', \Gamma')) \in E' \times E' \colon (e{:}\varphi, e'{:}\varphi') \in Q_a, \varphi \in \Gamma, \varphi' \in \Gamma'\}$,
$pre'((e, \Gamma)) = T'_2(pre(e) \wedge \bigwedge \Gamma)$.

Theorem 4. *$\mathcal{C} \in \mathscr{C}(\mathcal{L}_{ECM})$ is update equivalent to $T'_1(\mathcal{C})$.*

Proof Sketch. Same proof idea as for Theorem 3, now using techniques from Kooi and Renne [2011, Thm 4.7]. □

We have now shown equal expressivity between standard event models and edge-conditioned ones, and hence also equal expressivity of $\mathcal{L}_{\text{DEL}}$ and $\mathcal{L}_{\text{ECM}}$. Next, we show that edge-conditioned event models are more succinct than standard ones (can be exponentially smaller, Thm. 6, and never more than linearly larger, Thm. 5). We use $|\mathcal{E}|$ for the size of an event model $\mathcal{E}$ (standardly defined and straightforwardly generalized to edge-conditioned event models, see the extended version on Arxiv).

Theorem 5. *For any $\mathcal{E} \in \mathscr{E}(\mathcal{L}_{DEL})$, $T_1(\mathcal{E})$ has size $O(|\mathcal{E}|)$.*

[7] A transformation of standard event models into edge-conditioned event models was first provided by [Li, 2023, p. 161]. Notice that their tranformation is only for event models that do not contain dynamic preconditions, whereas ours includes them too.

Proof Sketch. As mentioned in the proof sketch of Theorem 3, $T_1(\mathcal{E})$ is achieved from $\mathcal{E}$ by (recursively) adding $\top$ as both source and target condition on each edge (e, f), and this only gives a linear blowup in size. □

Theorem 6 (Exponential succinctness of edge-conditioned event models)**.** *Let $p \in P$ and let $n = |Ag|$. Then:*

1. *$\mathcal{H}(p)$ is of size $O(n)$.*
2. *Any standard event model that is update equivalent to $\mathcal{H}(p)$ has at least 2^n events.*

Proof Sketch. Item 1 follows from inspecting Fig. 3. To prove 2, we consider any standard event model $\mathcal{E}$ update equivalent to $\mathcal{H}(p)$. For each subset $Ag' \subseteq Ag$, we define a pointed Kripke model $(\mathcal{M}_{Ag'}, w_1) = ((W, R, V), w_1)$ with $W = \{w_1, w_2\}$, $R_a = \{(w_1, w_2)\}$ for all $a \in Ag$, $V(w_1) = \{p\} \cup \{A_a p : a \in Ag'\}$ and $V(w_2) = \emptyset$. Using Def. 7 and Def. 8, we show that the actual world of $(\mathcal{M}_{Ag'}, w_1) \otimes \mathcal{H}(p)$ has an a-edge to a $\neg p$-world iff $a \in Ag \setminus Ag'$. Since $\mathcal{E}$ is update equivalent to $\mathcal{H}(p)$, then also the actual world of $(\mathcal{M}_{Ag'}, w_1) \otimes \mathcal{E}$ has an a-edge to a $\neg p$-world iff $a \in Ag \setminus Ag'$. From this it follows that $\mathcal{E}$ has an actual event $e_d^{Ag'}$ with an a-edge to an event applicable in w_2 iff $a \in Ag \setminus Ag'$. For $Ag' \neq Ag''$, the events $e_d^{Ag'}$ and $e_d^{Ag''}$ must hence have different outgoing edges and thus be distinct. This implies that there are at least as many distinct events in $\mathcal{E}$ as there are subsets of Ag, i.e. at least $2^{|Ag|}$. □

This result shows that edge-conditioned event models can be exponentially smaller than their standard event model counterparts, and that exponential succinctness specifically holds for the event models for propositional attention (as comparing Figs. 2 and 3 also suggests).

Generalized Arrow Updates The *generalized arrow updates* of Kooi and Renne [2011] are a distinct class of event models that are similar to our edge-conditioned event models in including source and target conditions. However, generalized arrow updates do not have preconditions, and so they do not straightforwardly generalize standard event models as our edge-conditioned event models do. They however still have the same expressivity as standard event models [Kooi and Renne, 2011], and hence the same expressivity as edge-conditioned event models. A disadvantage of generalized arrow updates compared to edge-conditioned event models is that they can be less succinct than their standard event model counterparts [Kooi and Renne, 2011, Thm. 3.14]. In the extended version of the paper on Arxiv, we show that edge-conditioned event models are also at least as succinct as generalized arrow updates, via a result corresponding to Theorem 5. Therefore, edge-conditioned event models may be a good choice of event model formalism for DEL, as they are always at least as succinct as standard event models and generalized arrow updates, and sometimes exponentially more succinct. Furthermore, as they straightforwardly generalize standard event models, extensions such as with postconditions [van Ditmarsch and Kooi, 2008] become trivial for edge-conditioned event models, which is not true for generalized arrow updates.

5 A Logic for General Attention

We now generalize propositional attention to account for attention to, and revelation of, arbitrary formulas. The *language of general attention* $\mathcal{L}_{\text{GA}}$ is the language given by the grammar of $\mathcal{L}_{\text{ECM}}$ extended with the clause $\varphi ::= A_a\varphi$, where $a \in Ag$ and A_a is a new modal operator. The formula $A_a\varphi$ reads "agent a pays attention to φ". Note that, while in $\mathcal{L}_{\text{PA}}$ the formula $A_a p$ is an atom, in $\mathcal{L}_{\text{GA}}$ it is a modality applied to a propositional atom. Additionally, in $\mathcal{L}_{\text{PA}}$, the formula $A_a p$ reads "agent a is paying attention to *whether* p", whereas in $\mathcal{L}_{\text{GA}}$ it reads "agent a is paying attention *to* p".

Moving to general attention allows to formalize many new scenarios, namely all those where agents attend to more complex stimuli than just conjunctions of literals. For example, we may have $A_a((p \vee q) \to r)$, meaning that agent a is paying attention to the conditional $(p \vee q) \to r$. Such a conditional may represent the statement of a mathematical theorem, and $A_a((p \vee q) \to r)$ then says that agent a is paying attention to it. Following Belardinelli and Bolander [2023], we understand attention as being directed to truthful revelations. Attending to a theorem then means that, if the theorem is revealed, maybe as part of a larger revelation such as a research talk, the agent will learn its truth value.

Another application is one where agents may pay attention (or not) to the utterances of other agents. Say that agent a only pays attention to what agent b, but not agent c, says about p. In DEL, the truthful and public announcement of a formula φ by an agent i can be represented by the singleton event model where the actual event has precondition $B_i\varphi$ [van Ditmarsch, 2023]. Such an announcement makes all agents know that i believes φ. Then, to formalise the mentioned attention situation, we could use the formula $A_aB_bp \wedge A_aB_b\neg p \wedge \neg A_aB_cp \wedge \neg A_aB_c\neg p$: if agent b truthfully announces the (believed) truth-value of p, then agent a receives that announcement, but if agent c does the same, a receives nothing. Other scenarios can be modeled, such as attention to the attention of other agents. For example, the formula A_aA_bp represents that a pays attention to b paying attention to p. Besides added expressivity, it is also conceptually natural to treat attention as a modality, similarly to propositional attitudes such as belief, intention, and awareness. The language $\mathcal{L}_{\text{GA}}$ is interpreted in attention models:

Definition 12 (Attention model)**.** *An* attention model *is a tuple $\mathcal{M} = (W, R, V, \mathcal{A})$ where (W, R, V) is a Kripke model for $\mathcal{L}_{GA}$ and $\mathcal{A} : Ag \times W \to \mathcal{P}(\mathcal{L}_{GA})$ is an* attention function. *For an* actual world *w, $(\mathcal{M}, w)$ is a* pointed attention model.

The notion of edge-conditioned product update (Def. 7) immediately extends to attention models by defining $\mathcal{A}(a,(w,e)) = \mathcal{A}(a,w)$ for all (w,e) of the updated model. This simply means that each world preserves its attention assignments, similarly to how we treat the valuation function in the product update. The set of formulas that agent a is paying attention to at world w is the *attention set* $\mathcal{A}(a,w)$, also denoted by $\mathcal{A}_a(w)$. The truth of $\mathcal{L}_{\text{GA}}$ formulas is defined by the same clauses as for $\mathcal{L}_{\text{ECM}}$ with the following addition: $(\mathcal{M}, w) \vDash A_a\varphi$ iff $\varphi \in \mathcal{A}_a(w)$.

This setting has clear similarities with the logic of general awareness [Fagin and Halpern, 1988]. This does not mean that our framework reduces to a formalism for awareness, as the crucial aspect of attention separating it from awareness is that attention determines what agents learn, as we will see, which awareness does not.[8] Yet, the two frameworks are equivalent in terms of static language and models. We also do not place any restriction on agents' attention sets, similarly to what happens in the logic of general awareness. We may have that $\varphi \wedge \psi \in \mathcal{A}_a(w)$ but $\psi \wedge \varphi \notin \mathcal{A}_a(w)$, or we may have that $\varphi, \psi \in \mathcal{A}_a(w)$ but $\varphi \wedge \psi \notin \mathcal{A}_a(w)$. Whether this freedom is reasonable depends on the specific applications one has in mind, which may require to impose a range of closure properties on $\mathcal{A}_a$, called *attention principles*. Consider the following examples and suggested applications:

Conjunctive closure: *$\varphi \wedge \psi \in \mathcal{A}_a(w)$ iff $\varphi \in \mathcal{A}_a(w)$ and $\psi \in \mathcal{A}_a(w)$*. This principle makes sense when representing 'divided attention': every time an agent attends to $\varphi \wedge \psi$ it is as if she is dividing attention between φ and ψ (i.e. she attends to φ and attend to ψ separately) and vice versa. This principle may not be natural in resource-limited settings, where it may be possible to attend to φ and ψ separately but not together.

Commutativity: *$\varphi \wedge \psi \in \mathcal{A}_a(w)$ iff $\psi \wedge \varphi \in \mathcal{A}_a(w)$*. We may treat the order of conjuncts as irrelevant, e.g. when abstracting from the temporal order of information presented.

Sublanguage closure: *If $\varphi \in \mathcal{A}_a(w)$, and ψ is a formula constructed from atoms appearing in φ (i.e. $At(\psi) \subseteq At(\varphi)$), then $\psi \in \mathcal{A}_a(w)$*. This principle may represent agents who are interested in a specific issue, and so their attention focuses on anything that talks about it.

Subformula closure: *If $\varphi \in \mathcal{A}_a(w)$ and ψ is a subformula of φ, then $\psi \in \mathcal{A}_a(w)$*. This principle is justified when modeling information that is such that agents cannot pay attention to it unless they attend to all its components.

Agent a ignoring agent b: *$B_b\varphi \notin \mathcal{A}_a(w)$ for all φ (or all φ that are about a certain issue)*. This can model agents who systematically do not pay attention to the utterances of other agents, as in attention-driven social biases [Munton, 2023].

Agent a attending to agent b: *$B_b\varphi \in \mathcal{A}_a(w)$ for all φ*. This models agents who systematically pay attention to the utterances of other agents. This may also allow to model agents who engage in social attention and e.g. study social learning [Rendell *et al.*, 2010]

Attention introspection: *$(w, v) \in R_a$ implies $\mathcal{A}_a(w) = \mathcal{A}_a(v)$*. This principle models agents who have no doubts regarding what they pay attention to. It makes sense to assume it e.g. when agents are deliberately focusing on something.

Notice that each property discussed above can be easily turned into an axiom schema, and each axiom schema can be used in an axiomatization of the modelled attention notion. As a few examples, the axiom schema $A_a\varphi \wedge A_a\psi \leftrightarrow A_a(\varphi \wedge \psi)$ represents conjunctive closure; $A_a(\varphi \wedge \psi) \leftrightarrow A_a(\psi \wedge \varphi)$ represents commutativity; the formulas $A_i\varphi \to B_iA_i\varphi$ and $\neg A_i\varphi \to B_i\neg A_i\varphi$ represent attention introspection. Each of the considered principles corresponds to an axiom schema (or a combination of axioms). While many of these principles and their corresponding axioms are discussed in the aware-

[8]Awareness is a static notion, introduced to restrict the formulas that agents may reason and have explicit beliefs about, as a solution to the logical omniscience problem [Fagin and Halpern, 1988].

ness literature [Fagin and Halpern, 1988], in this settings they are interpreted differently, as attention is an intrinsically dynamic notion. Going beyond them, the correspondence to awareness logic allows to study the relation between attention and awareness, rarely discussed in the literature (but see [Fritz and Lederman, 2015; Belardinelli, 2023]).

Event Models for General Attention Suppose given a set $\Gamma \subseteq \mathcal{L}_{\text{GA}}$: the formulas that are revealed (or announced) by the occurring event. The intuition is now that every agent learns the subset of Γ that they are paying attention to. If $\psi \in \Gamma \cap \mathcal{A}_a(w)$, then agent a learns ψ at world w.

Definition 13 (Event model for general attention $\mathcal{R}(\Gamma)$)**.** *Let $\Gamma \subseteq \mathcal{L}_{GA}$ be a set of* revealed formulas*. The* event model for general attention *representing the revelation of Γ is the pointed edge-conditioned event model $\mathcal{R}(\Gamma) = ((E, Q, id_E), \bigwedge \Gamma)$ for $\mathcal{L}_{GA}$ defined by:*

$$E = \{\bigwedge S \colon S \subseteq \Gamma\},$$

$$Q_a = \{(\bigwedge S \colon \bigwedge_{\varphi \in T} A_a\varphi \wedge \bigwedge_{\varphi \in S \setminus T} \neg A_a\varphi, \bigwedge T \colon \bigwedge_{\varphi \in T} A_a\varphi) \colon T \subseteq S \subseteq \Gamma\}.$$

This event model contains, for each subset S of the revealed Γ, an event $\bigwedge S$. This represents the subset that an agent may learn by paying attention. The intuition behind the tuple $(\bigwedge S \colon \bigwedge_{\varphi \in T} A_a\varphi \wedge \bigwedge_{\varphi \in S \setminus T} \neg A_a\varphi, \bigwedge T \colon \bigwedge_{\varphi \in T} A_a\varphi)$ is that if $\bigwedge S$ is revealed at an event and agent a pays attention only to the subset T of S, then at that event agent a believes that only $\bigwedge T$ was revealed (as the event accessible from $\bigwedge S$ has precondition $\bigwedge T$), and a also learns that she was paying attention to T (as the event accessible from $\bigwedge S$ has (target) condition $\bigwedge_{\varphi \in T} A_a\varphi$). This is consistent with the intuition underlying propositional attention, but in a much more general setting. Lastly, notice that in the event model for general attention, $\bigwedge \Gamma$ is the actual event, representing the truthful revelation of Γ.

Notice that the attention introspection property is immediately preserved by this event model, as the target condition $A_a\varphi$, for all $\varphi \in T$, ensures that an agent knows what she attends to. The other attention principles are also preserved, as the dynamics do not modify agents' attention set.

Example 3. *Consider an enrichment of Example 2 and Fig. 4, where the AI now also pays attention to Ann's utterances regarding the CV. In particular, the AI agent is focusing on whether Ann has information regarding p and q, as that would mean that she paid attention to all aspects of the CV. More formally, we are considering a pointed attention model $(\mathcal{M}', w')$ for $\mathcal{L}_{GA}$ that is entirely like $(\mathcal{M} \otimes \mathcal{H}(p \wedge q))$ from Fig. 4, except that we have $A_bB_ap \wedge A_b\neg B_ap \wedge A_bB_aq \wedge A_b\neg B_aq$ true at all b-accessible worlds (see extended version for corresponding figure). Continuing the example, Ann has now submitted an online report to the AI agent with fillable text fields for research qualifications, teaching qualifications, contributions to diversity, etc. Ann writes p for research qualifications (cf. Fig. 1), but nothing for diversity, corresponding to the revelation of $\Gamma = \{B_ap, \neg B_aq\}$. From this, the AI agent correctly infers that Ann has not paid attention to diversity issues. This situation is represented in Fig. 5, where*

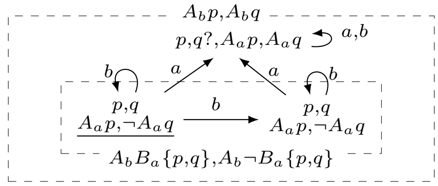

Figure 5: Attention model $(\mathcal{M}' \otimes \mathcal{R}(\{B_ap, \neg B_aq\}), (w', e))$ for $\mathcal{L}_{\text{GA}}$, where $(\mathcal{M}', w')$ is given in Example 3. We use the same conventions as before and additionally we let $B_a\{p, q\} := B_ap \wedge B_aq$. While earlier A_ap occurring at w represented that $A_ap \in V(w)$, here it represents that $p \in \mathcal{A}_a(w)$ (similarly for $\neg A_ap$).

$(\mathcal{M}' \otimes \mathcal{R}(\Gamma), (w', e)) \vDash A_ap \wedge \neg A_aq \wedge B_bA_ap \wedge B_b\neg A_aq$: the AI agent correctly believes that Ann has not paid attention to all parts of the CV.

This example illustrates how the logic of general attention may be used, for example by an AI agent, to reason about and discover the attentional biases of other agents. While the example is relatively basic, the underlying logical framework is very general, and can handle much more complex scenarios.

6 Concluding Remarks

In this paper, we first proposed a generalization of edge-conditioned event models that we used to model propositional attention. We showed that these models can be exponentially more succinct than standard event model, and never more than linearly larger. Then, we adopted them to capture attention to arbitrary formulas. In future work, we would like to further investigate the properties of our general attention framework and extend it to incorporate additional features of attention, such as capacity constraints. Limiting the attention resources of agents prompts questions about exactly which formulas the agent will learn. For example, consider an agent that has the capacity to attend to n formulas at most. If Γ reveals more than n formulas, which formulas will the agent prioritize? To address this issue, we need an ordering of the formulas that determines which ones are attended first and thus learned. We will explore this topic in our next paper.

A standard limitation of classical DEL, inherited by our framework, is that agents cannot recover from false beliefs: If an agent believes $\neg p$ and a public announcement of p occurs, the agent will come to believe every formula. The standard response is that these cases require to move to a richer framework, such as DEL based on plausibility models [Baltag and Smets, 2008]. Adapting our logic of general attention to that framework is another interesting direction for future work.

Acknowledgments

Gaia Belardinelli is funded by Independent Research Fund Denmark (grant no. 4255-00020B). Thomas Bolander is supported by the Independent Research Fund Denmark (grant no. 10.46540/4258-00060B). Sebastian Watzl is funded by the European Union's Horizon 2020 research and innovation program under grant agreement no. 101003208 (ERC Consolidator Grant 2020).

References

[Baltag and Smets, 2008] Alexandru Baltag and Sonja Smets. A qualitative theory of dynamic interactive belief revision. In Giacomo Bonanno, Wiebe van der Hoek, and Michael Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT7)*, volume 3 of *Texts in Logic and Games*, pages 13–60. Amsterdam University Press, 2008.

[Baltag *et al.*, 1998] Alexandru Baltag, Lawrence S. Moss, and Slawomir Solecki. The logic of public announcements and common knowledge and private suspicions. In Itzhak Gilboa, editor, *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-98)*, pages 43–56. Morgan Kaufmann, 1998.

[Belardinelli and Bolander, 2023] Gaia Belardinelli and Thomas Bolander. Attention! Dynamic Epistemic Logic Models of (In)Attentive Agents. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '23, page 391–399, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems.

[Belardinelli, 2023] Gaia Belardinelli. *Epistemic Logic Models of Awareness, Attention, and Implicit Reasoning*. PhD thesis, Københavns Universitet, 2023.

[Blackburn *et al.*, 2001] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, UK, 2001.

[Bolander *et al.*, 2016] Thomas Bolander, Hans van Ditmarsch, Andreas Herzig, Emiliano Lorini, Pere Pardo, and François Schwarzentruber. Announcements to Attentive Agents. *Journal of Logic, Language and Information*, 25:1–35, 2016.

[Bolander *et al.*, 2021] Thomas Bolander, Lasse Dissing, and Nicolai Herrmann. Del-based epistemic planning for human-robot collaboration: Theory and implementation. In *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning (KR 2021)*, 2021.

[Bolander, 2018] Thomas Bolander. Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic. In *Outstanding Contributions to Logic*, number 12, pages 207–236. Springer, 2018.

[Boyd *et al.*, 2011] Robert Boyd, Peter J Richerson, and Joseph Henrich. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(supplement_2):10918–10925, 2011.

[Fagin and Halpern, 1988] R. Fagin and J. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

[Fagin *et al.*, 1995] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. MIT Press, 1995.

[Fritz and Lederman, 2015] Peter Fritz and Harvey Lederman. Standard State Space Models of Unawareness (Extended Abstract). *Electronic Proceedings in Theoretical Computer Science*, 215:141–158, 2015.

[Johnson, 2024] Gabbrielle M Johnson. Varieties of bias. *Philosophy Compass*, 19(7):e13011, 2024.

[Kooi and Renne, 2011] Barteld Kooi and Bryan Renne. Generalized arrow update logic. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 205–211. ACM, 2011.

[Li, 2023] Lei Li. *Games, Boards and Play: A Logical Perspective*. PhD thesis, University of Amsterdam, 2023.

[Munton, 2023] Jessie Munton. Prejudice as the misattribution of salience. *Analytic Philosophy*, 64(1):1–19, 2023.

[Rendell *et al.*, 2010] Luke Rendell, Robert Boyd, Daniel Cownden, Marquist Enquist, Kimmo Eriksson, Marc W Feldman, Laurel Fogarty, Stefano Ghirlanda, Timothy Lillicrap, and Kevin N Laland. Why copy others? Insights from the social learning strategies tournament. *Science*, 328(5975):208–213, 2010.

[Simons and Chabris, 1999] Daniel J. Simons and Christopher F. Chabris. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception: Visual perception*, 28(9):1059–1074, 1999.

[Smith and Archer, 2020] Leonie Smith and Alfred Archer. Epistemic injustice and the attention economy. *Ethical Theory and Moral Practice*, 23(5):777–795, 2020.

[van Ditmarsch and Kooi, 2008] Hans van Ditmarsch and Barteld Kooi. Semantic results for ontic and epistemic change. In Giacomo Bonanno, Wiebe van der Hoek, and Michael Wooldridge, editors, *Logic and the Foundation of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games 3, pages 87–117. Amsterdam University Press, 2008.

[van Ditmarsch *et al.*, 2007] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer Publishing Company, 2007.

[van Ditmarsch, 2023] Hans van Ditmarsch. To be announced. *Information and Computation*, 292:105026, 2023.

[van Eijck *et al.*, 2012] Jan van Eijck, Ji Ruan, and Tomasz Sadzik. Action emulation. *Synthese*, 185:131–151, 2012.

[Watzl, 2017] Sebastian Watzl. *Structuring mind: The nature of attention and how it shapes consciousness*. Oxford University Press, 2017.