A Logic for General Attention Using Edge-Conditioned Event Models
Paper accepted at IJCAI 2025

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

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 and do not learn about it [Simons and Chabris, 1999]. Hence, while the ability to focus 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 or what certain individuals 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.

Earlier work introduced dynamic epistemic logic models (DEL models) of attention, capturing the effects of attention on agents' beliefs. However, the first proposal only modeled a notion of all-or-nothing attention [Bolander *et al.*, 2016], later extended to apply to atomic propositions [Belardinelli and Bolander, 2023]. While the latter of these can treat cases where a member of a hiring committee systematically attends only to, say, the research parts of applicants' CVs, it cannot represent an AI system that directs its attention to whether a committee member has such an attentional bias. It can also not represent agents whose attention is biased toward specific agents, e.g. paying attention to whether candidate a has achieved some result, but not whether b achieved the same [Munton, 2023; Smith and Archer, 2020]. A main obstacle to richer notions is that models for attention get highly complex, even for all-or-nothing attention, and that such models grow at least exponentially large. To obtain a fully general (and still manageable) theory of attention several technical and conceptual innovations were necessary.

In this work, we present the needed technical and conceptual innovations as well as the fully general logic of attention, where agents can attend to any arbitrary formulas. Attention is modelled as a modality that restricts which parts of an event an agent learns. An agent who is not attending to a formula is unable to distinguish an event containing that formula from one that does not. To formalize this, we adopt the framework of DEL, in particular its edge-conditioned version [Bolander, 2018]. We generalize edge-conditioned event models to have both source and target conditions. This simple modification allows to use source conditions to encode information about agents' current attention, while the target conditions encode what agents learn about their own attention. This representation brings both a conceptually cleaner and more tractable reformulation of earlier event models for propositional attention [Belardinelli and Bolander, 2023]. Indeed, in the main paper of this extended abstract [Belardinelli *et al.*, 2025], we show that while edge-conditioned DEL models are as expressive as standard event models [Baltag *et al.*, 1998], they are exponentially more succinct than them. Additionally, we provide a sound and complete axiomatization of edge-conditioned event models, show that they are as expressive as generalized arrow updates [Kooi and Renne, 2011], and as succinct as them. Finally, largely due to the gained simplicity of the new formalization, we managed to generalize the entire framework along another axis: We generalize the framework to handle attention to arbitrary formulas, not only propositional ones.

In this extended abstract, we focus on presenting edge-conditioned DEL and the logic of general attention.

Edge-conditioned DEL We use Ag to denote a finite set of *agents* and P to denote a finite set of *propositional atoms*. The *language of dynamic epistemic logic* $\mathcal{L}$ is given by the grammar: $\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_a\varphi \mid [\mathcal{C}]\varphi$, with $p \in P$, $a \in Ag$, and $\mathcal{C}$ an edge-conditioned event model (defined later). $B_a\varphi$ reads "agent a believes φ", and $[\mathcal{C}]\varphi$ reads "after $\mathcal{C}$ happens, φ is the case".

We now introduce DEL with edge-conditions. In the main paper [Belardinelli *et al.*, 2025], we also introduce standard DEL and *event models for propositional attention* [Belardinelli and Bolander, 2023], here omitted for brevity.

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 $V : W \to \mathcal{P}(P)$ is a* valuation function*. For $w \in W$, $(\mathcal{M}, w)$ is a* pointed Kripke model*.*

Definition 2 (Edge-conditioned event models)**.** *An* edge-conditioned event model *for $\mathcal{L}$ is $\mathcal{C} = (E, Q, pre)$ where $E \neq \emptyset$ is a finite set of* events*, $Q : Ag \to \mathcal{P}(E \times \mathcal{L} \times E \times \mathcal{L})$ assigns to each agent a set of quadruples (e, φ, f, ψ) (abbreviated $(e{:}\varphi, f{:}\psi)$), and $pre : E \to \mathcal{L}$. For $(e{:}\varphi, f{:}\psi) \in Q_a$, we call $(e{:}\varphi, f{:}\psi)$ a* conditioned edge*, where φ is the* source condition *(at e) and ψ is the* target condition *(at f). For $e \in E$, $(\mathcal{C}, e)$ is a* pointed edge-conditioned event model*.*

Intuitively, a Kripke model represents an epistemic state, while a event model represents an epistemic action/event happening. The idea behind edge-conditioned event models is to make the edges of event models, capturing accessible worlds, conditional on formulas. In edge-conditioned event models, the events that an agent can access are modeled by quadruples: $(e{:}\varphi, f{:}\psi) \in Q_a$ 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. The *product update operator*, defined next, expresses how an epistemic state is updated as the consequence of an epistemic event.

Definition 3 (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 $\mathcal{L}$. The* product update *of $\mathcal{M}$ with $\mathcal{C}$ is $\mathcal{M} \otimes \mathcal{C} = (W', R', V')$ where $W' = \{(w, e) \in W \times E \colon (\mathcal{M}, w) \models pre(e)\}$, $V'((w, e)) = \{p \in At(\mathcal{L}) \colon p \in V(w)\}$, and $R'_a = \{((w, e), (v, f)) \in (W')^2 \colon (w, v) \in R_a$ and $\exists \varphi, \psi \in \mathcal{L}$ s.t. $(e{:}\varphi, f{:}\psi) \in Q_a, (\mathcal{M}, w) \models \varphi$ and $(\mathcal{M}, v) \models \psi\}$. If $(\mathcal{M}, w) \models pre(e)$, the* product update *of $(\mathcal{M}, w)$ with $(\mathcal{C}, e)$ is the pointed Kripke model $(\mathcal{M}, w) \otimes (\mathcal{C}, e) = (\mathcal{M} \otimes \mathcal{C}, (w, e))$.*

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

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

In the main paper [Belardinelli *et al.*, 2025], we provide an axiomatization for edge-conditioned DEL that we show to be sound and complete. Moreover, we recast *event models for propositional attention* [Belardinelli and Bolander, 2023] in the edge-conditioned format, and we show that the result is equally expressive, but exponentially more succinct. Moreover, we compare the classes of standard DEL event models and edge-conditioned event models, showing the two equally expressive. We show the same for generalized arrow updates.

A logic for general attention We now apply edge-conditioned event models to provide a formalism for attention to, and revelation of, arbitrary formulas. The *language of general attention* $\mathcal{L}_{\text{GA}}$ is $\mathcal{L}$ 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". Following Belardinelli and Bolander [2023], we understand attention as being directed to truthful revelations.

A logic for general attention allows to formalize many more scenarios than the logic of propositional attention, 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 $(p \vee q) \to r$ which might e.g. be representing a mathematical theorem attended to. Agents may also pay attention to the utterances of specific agents only (or combinations of specific agents and topics). 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. We interpret $\mathcal{L}_{\text{GA}}$ in attention models.

Definition 5 (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 set of formulas that agent a is paying attention to at world w is the *attention set* $\mathcal{A}(a, w)$, also denoted $\mathcal{A}_a(w)$. The truth of $\mathcal{L}_{\text{GA}}$ formulas is defined by extending the above semantic clauses with: $(\mathcal{M}, w) \models A_a\varphi$ iff $\varphi \in \mathcal{A}_a(w)$. This setting has clear similarities with the logic for 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. We do not place any universal restriction on agents' attention sets, e.g. we may have that $\varphi \wedge \psi \in \mathcal{A}_a(w)$ but $\psi \wedge \varphi \notin \mathcal{A}_a(w)$. Some applications may however require to impose a range of closure properties on $\mathcal{A}_a$, called *attention principles* (see the main paper [Belardinelli *et al.*, 2025] for relevant examples).

Event models for general attention Suppose given $\Gamma \subseteq \mathcal{L}_{\text{GA}}$: the formulas that are revealed (or announced) by the occurring event. 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. In the definition below, id_E is the identity function on E, i.e. $id_E(e) = e$ for all $e \in E$.

Definition 6 (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:*

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

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

This event model contains, for each subset S of the revealed Γ, an event $\bigwedge S$: 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 in all events accessible from $\bigwedge S$, $\bigwedge T$ holds), and a also learns that she was paying attention to T (as in all events accessible from $\bigwedge S$, $\bigwedge_{\varphi \in T} A_a\varphi$ holds).

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, explained in more detail in the main paper [Belardinelli *et al.*, 2025].

References

[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 *et al.*, 2025] Gaia Belardinelli, Thomas Bolander, and Sebastian Watzl. A Logic for General Attention Using Edge-Conditioned Event Models. In *Proceedings of the 2025 International Joint Conference on Artificial Intelligence (IJCAI)*, IJCAI '25, 2025.

[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, 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.

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

[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.

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

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

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