ERRATUM

In Figure 3, the two outgoing dotted edges from the box labelled $h_b p, h_b g$ (upper left box) are incorrect and should be removed. The rest of the edges are correct (and have been computationally verified. The small Python program that we used to verify that the edges are correct can be found here: https://tinyurl.com/5ekjmsud).

A version of the paper where Figure 3 has been corrected can be found here: https://arxiv.org/abs/2303.13494

Attention! Dynamic Epistemic Logic Models of (In)attentive Agents

Gaia Belardinelli
University of Copenhagen
Copenhagen, Denmark
belardinelli@hum.ku.dk

Thomas Bolander
Technical University of Denmark
Kgs. Lyngby, Denmark
tobo@dtu.dk

ABSTRACT

Attention is the crucial cognitive ability that limits and selects what information we observe. Previous work by Bolander et al. (2016) proposes a model of attention based on dynamic epistemic logic (DEL) where agents are either fully attentive or not attentive at all. While introducing the realistic feature that inattentive agents believe nothing happens, the model does not represent the most essential aspect of attention: its selectivity. Here, we propose a generalization that allows for paying attention to subsets of atomic formulas. We introduce the corresponding logic for propositional attention, and show its axiomatization to be sound and complete. We then extend the framework to account for inattentive agents that, instead of assuming nothing happens, may default to a specific truth-value of what they failed to attend to (a sort of prior concerning the unattended atoms). This feature allows for a more cognitively plausible representation of the inattentional blindness phenomenon, where agents end up with false beliefs due to their failure to attend to conspicuous but unexpected events. Both versions of the model define attention-based learning through appropriate DEL event models based on a few and clear edge principles. While the size of such event models grow exponentially both with the number of agents and the number of atoms, we introduce a new logical language for describing event models syntactically and show that using this language our event models can be represented linearly in the number of agents and atoms. Furthermore, representing our event models using this language is achieved by a straightforward formalisation of the aforementioned edge principles.

KEYWORDS

Dynamic Epistemic Logic; Attention; Inattentional Blindness; Default Values; Syntactic Event Models; Succinctness

ACM Reference Format:
Gaia Belardinelli and Thomas Bolander. 2023. Attention! Dynamic Epistemic Logic Models of (In)attentive Agents. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 9 pages.

1 INTRODUCTION

Attention is the capacity of the mind to focus on a specific subset of available information. It limits and selects what we observe, to the extent that we may only consciously perceive events that receive our focused attention [16]. A fascinating family of phenomena suggesting that attention is necessary for visual awareness is the one where agents completely miss conspicuous events even when they happen at fixation. Inattentional blindness is one such phenomenon [13]. The name is suggestive of a form of cognitive blindness to external stimuli, which has been robustly replicated in the cognitive science literature. A famous experiment to test it is the so called *Invisible Gorilla* video [15], by Simons and Chabris [16]. It is an online video where subjects are asked to “Count how many times the players wearing white pass the basketball". While they focus on counting the ball passages, a clearly visible person in a gorilla costume crosses the scene. It is an unexpected appearance in such a situation, but it is an appearance right at fixation, as the gorilla passes in the middle of the group of players. Yet, Simons and Chabris found that about 50% of subjects do not perceive any gorilla [15, 16]. Interestingly, subjects are often surprised when they realise to have missed such a salient event. This surprise has been taken to reveal a metacognitive error about the completeness of visual awareness, or in other words, an incorrect belief about attention capacities [17, 18]. Indeed, researchers have shown that it is common for people to believe that they would notice much more than they in fact do, and that when they fail to attend something, they are not only uncertain about what they missed, but they often believe that what they did not notice, did not happen [16]. Attention and its limitations thus have substantial implications in people's belief dynamics, both in the sense that they severely limit what information is received, and in the sense that subjects often hold definite beliefs about events that they attend or fail to attend.

Dynamic Epistemic Logic (DEL) is a branch of epistemic logic that has been used to study the dynamics of knowledge and beliefs [19]. Only relatively recently has there been investigations into the notion of attention and related phenomena in the DEL literature. For example, Bolander et al. (2016) introduce a form of attention in a DEL framework, representing it as an atomic formula h_a that, if true, expresses that agent a pays attention to everything happening (any formula announced or any fact revealed) and, if false, that a pays attention to nothing at all [7]. That work might be considered as a first step towards modelling the rich and complex phenomenon of attention in DEL, and the present work may then be considered a second step. One of our contributions consists in generalizing the framework from [7] so that agents can pay attention to any subset of atomic formulas. We encode attention by means of attention atoms $\mathsf{h}_a p$, for each agent a and proposition p. For the dynamic part, we generalise the event models from [7] by first recasting them using a few and clear edge principles. Then, we gradually introduce different event models by building on this version of their model. What we do is the following: First, we account for agents that may have false beliefs about their attention (as in the inattentional blindness phenomenon above). Second, we account

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

for the dynamics of partial learning happening when agents only focus on a subset of the occurring events. In this version of the model, agents learn the part of the events that they are paying attention to, but keep intact their beliefs about what they did not attend to. As we have seen above, in reality, this is not always the case. In inattentional blindness for example, it often happens that inattentive agents change their beliefs to specifically account for the assumption that unattended events *did not occur*. Then, as a third step, we add default values as a parameter of event models, which are a sort of prior that agents have and use to update their beliefs in case they miss some information. This addition gives us a more cognitively plausible representation of the experimental findings mentioned above, as now agents can default to the non-existence of the gorilla in the video even if they were previously uncertain about it. We introduce a logic for the first model of propositional attention (without defaults), and prove its axiomatization sound and complete. Lastly, we show that our idea of representing edges of event models by edge principles can be generalised to a new type of syntactic event models where events and edges are specified using logical formulas. We show exponential succinctness of these syntactic event models as compared to standard (semantic) event models.

Besides providing insights into how human attention interacts with beliefs, this research also goes towards the improvement of human-AI interaction, as it may help e.g. robots to reason about humans, required in human-robot collaboration settings. As explained by Verbrugge [20], it's potentially dangerous if a robot in a human-robot rescue team makes too optimistic assumptions about the reasoning powers of human team members. The robot might for example falsely rely on a human to have paid attention to a certain danger, where in fact the human didn't. A proactively helpful robot should be able to take the perspective of the human and reason about what the human might or might not have paid attention to, and therefore which false beliefs the human might have. This requires that the robot has a model of the attention system of the human, and how this impacts her beliefs. We believe our models can be used in this way. Concretely, there has already been research on using epistemic planning based on DEL for human-robot collaboration [6], and since the models of this paper are also based on DEL, they lend themselves to immediate integration into such frameworks and systems. Full proofs of all results are in the supplementary material that can be found at https://tinyurl.com/2p8etvjy.

2 PROPOSITIONAL ATTENTION

2.1 Language

Throughout the paper, we use Ag to denote a finite set of *agents*, At to denote a finite set of *propositional atoms*, and we let $H = \{\mathsf{h}_a p \colon p \in At, a \in Ag\}$ denote the corresponding set of *attention atoms*. With $p \in At$, $a \in Ag$, $\mathsf{h}_a p \in H$ and $\mathcal{E}$ being a multi-pointed event model[1], define the language $\mathcal{L}$ by:[2]

$$\varphi ::= \top \mid p \mid \mathsf{h}_a p \mid \neg\varphi \mid \varphi \wedge \varphi \mid B_a\varphi \mid [\mathcal{E}]\varphi.$$

The attention atom $\mathsf{h}_a p$ reads "agent a is paying attention to whether p",[3] $B_a\varphi$ reads "agent a believes φ", and the dynamic modality $[\mathcal{E}]\varphi$ reads "after $\mathcal{E}$ happens, φ is the case". The formulas in $At \cup H \cup \{\top\}$ are called the *atoms*, and a *literal* is an atom or its negation. We often write $\bigwedge S$ to denote the conjunction of a set of formulas S. If S is empty, we take $\bigwedge S$ as a shorthand for $\top$. To keep things simple, we will assume that all consistent conjunction of literals are in a normal form where: (i) each atom occurs at most once; (ii) $\top$ doesn't occur as a conjunct, unless the formula itself is just $\top$; and (iii) the literals occur in a predetermined order (ordered according to some total order on $At \cup H$). This implies that given any disjoint sets of atoms P^+ and P^-, there exists a unique conjunction of literals (in normal form) containing all the atoms of P^+ positively and all the atoms of P^- negatively. For conjuncts that are *not* on this normal form, we assume them to always be replaced by their corresponding normal form. For any conjunction of literals $\varphi = \bigwedge_{1 \le i \le n} \ell_i$ and any literal ℓ, we say that φ *contains* ℓ if $\ell = \ell_i$ for some i, and in that case we often write $\ell \in \varphi$. For any conjunctions of literals φ, we define $Lit(\varphi)$ to be the set of literals it contains, that is, $Lit(\varphi) = \{\ell \mid \ell \in \varphi\}$. For an arbitrary formula φ, we let $At(\varphi)$ denote the set of propositional atoms appearing in it.

2.2 Kripke Model and Dynamics

We are going to model attention and beliefs using DEL [19], where static beliefs are modelled by pointed Kripke models, and attention-based belief updates are modelled by multi-pointed event models (our product update and satisfaction definitions will be slightly non-standard due to the multi-pointedness of the event models).

Definition 2.1 (Kripke Model). A *Kripke model* 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 $V : W \to \mathcal{P}(At \cup H)$ is a *valuation function*. Where w is the *designated world*, we call $(\mathcal{M}, w)$ a *pointed Kripke model*.

Definition 2.2 (Event Model). An *event model* 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 $e \in E$. Where $E_d \subseteq E$ is a set of *designated events*, $(\mathcal{E}, E_d)$ is a *multi-pointed event model*. When $Ag = \{a\}$ for some a, we usually refer to the single-agent event model (E, Q, pre) as (E, Q_a, pre).

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

Definition 2.3 (Product Update). Let $\mathcal{M} = (W, R, V)$ be a Kripke model and $\mathcal{E} = (E, Q, pre)$ be an event model. The *product update* of $\mathcal{M}$ with $\mathcal{E}$ is the Kripke model $\mathcal{M} \otimes \mathcal{E} = (W', R', V')$ where:

$W' = \{(w, e) \in W \times E \colon (\mathcal{M}, w) \vDash pre(e)\}$,[4]

$R'_a = \{((w, e), (v, f)) \in W' \times W' \colon (w, v) \in R_a \text{ and } (e, f) \in Q_a\}$,

$V'((w, e)) = \{p \in At \cup H \colon w \in V(p)\}$.

Given a pointed Kripke model $(\mathcal{M}, w)$ and a multi-pointed event model $(\mathcal{E}, E_d)$, we say that $(\mathcal{E}, E_d)$ is *applicable* in $(\mathcal{M}, w)$ iff there

[1] Defined further below. As usual in DEL, the syntax and semantics are defined by mutual recursion [19].
[2] So $\mathcal{L}$ takes the sets Ag and At as parameters, but we'll keep that dependency implicit throughout the paper.
[3] The h in the attention formula stands for hearing. It was proposed in [7], and we keep it as we take their framework as our starting point.
[4] We haven't yet defined satisfaction of formulas in $\mathcal{L}$. It's defined in Definition 2.4 below, where we again note the standard mutual recursion used in defining DEL [19].

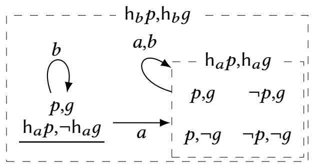

Figure 1: The pointed Kripke model $(\mathcal{M}, w)$. In the figure, p stands for "the players in the video pass the ball 15 times", g for "a clearly visible gorilla crosses the scene". We use the following conventions. Worlds are represented by sequences of literals true at the world. The model above has 5 worlds, 4 of which are inside the inner dashed box. Designated worlds are underlined. Whenever a world appears inside a dashed box, all the literals in the label of that box are also true in the world—and if the label is underlined, all worlds inside are designated. In this model, h_bp and h_bg hold in all worlds, and additionally, h_ap and h_ag hold in the worlds of the inner box. The accessibility relations are represented by labelled arrows. An arrow from (or to) the border of a dashed box means that there is an arrow from (or to) all the events inside the box.

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 $(\mathcal{M}, w) \otimes (\mathcal{E}, E_d) = (\mathcal{M} \otimes \mathcal{E}, (w, e))$ where e is the unique element of E_d satisfying $(\mathcal{M}, w) \vDash pre(e)$.

Definition 2.4 (Satisfaction). Let $(\mathcal{M}, w) = ((W, R, V), w)$ be a pointed Kripke model. For any $q \in At \cup H, a \in Ag, \varphi \in \mathcal{L}$ and any multi-pointed event model $\mathcal{E}$, satisfaction of $\mathcal{L}$-formulas in $(\mathcal{M}, w)$ is given by the following clauses extended with the standard clauses for the propositional connectives:

$$
\begin{array}{lll}
(\mathcal{M}, w) \vDash q & \text{iff} & q \in V(w); \\
(\mathcal{M}, w) \vDash B_a\varphi & \text{iff} & (\mathcal{M}, v) \vDash \varphi \text{ for all } (w, v) \in R_a; \\
(\mathcal{M}, w) \vDash [\mathcal{E}]\varphi & \text{iff} & \text{if } \mathcal{E} \text{ is applicable in } (\mathcal{M}, w) \text{ then} \\
 & & (\mathcal{M}, w) \otimes \mathcal{E} \vDash \varphi.
\end{array}
$$

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 $\vDash \varphi$.

Example 2.5. Ann and Bob are watching the Invisible Gorilla video [15]. Unbeknownst to Ann, Bob has already seen the video, so he knows the correct answer is 15 and that a clearly visible gorilla will pass by. Ann instead has no information about these things, as she has never seen that video. However, she likes riddles and tests of this sort, in which she gets absorbed very easily. Bob knows that, and thus he also knows that she will completely focus on counting the passages only, without realising that there is a gorilla, and thereby thinking to be paying attention to everything happening in the video, just as Bob. This situation is represented in Figure 1. We have $(\mathcal{M}, w) \vDash B_a h_a g \wedge \neg h_a g$: Ann believes she is paying attention to whether there is a gorilla or not, but she isn't.

3 PRINCIPLES FOR ATTENTION DYNAMICS

In this section, we first present the existing attention model [7]. We then propose an alternative representation using our edge principles, introduce a variant, and, finally, generalize to multiple propositions (capturing that agents can pay attention to subsets of At).

3.1 The Existing Model and our Version of it

As in [7], attention is represented as a binary construct where agents can either be paying attention to everything that happens or to nothing. The language they adopt is as the language above, except for their attention atoms h_a, $a \in Ag$, that are not relativised to propositional formulas. The intended meaning of such atoms is that the agent pays attention to everything, so they can be expressed in our language by letting h_a, $a \in Ag$, be an abbreviation of the formula $\bigwedge_{p \in At} h_a p$. Let $H' = \{h_a : a \in Ag\}$. Then $H' \cup At$ is the set of "atoms" on which their language is based. The static part of their model is a Kripke model, where it is assumed that agents are *attention introspective*, namely for all $w, v \in W, a \in Ag$, and $p \in At$, if $(w, v) \in R_a$ then $h_a(p) \in V(w)$ iff $h_a(p) \in V(v)$. The dynamics are given by the following event models. These event models represent situations in which any formula can be announced, true or false, and attentive agents will come to believe it.

Definition 3.1 (Event Model $\mathcal{E}(\varphi)$, [7]). Given a $\varphi \in \mathcal{L}$, the multi-pointed event model $\mathcal{E}(\varphi) = ((E, Q, pre), E \setminus \{s_\top\})$ is defined by:

$E = \{(i, J) : i \in \{0, 1\} \text{ and } J \subseteq Ag\} \cup \{s_\top\}$;

$Q_a = \{((i, J), (1, K)) : i \in \{0, 1\}, J, K \subseteq Ag \text{ and } a \in J\} \cup \{((i, J), s_\top) : i \in \{0, 1\}, J \subseteq Ag \text{ and } a \notin J\}$;

$pre : E \to \mathcal{L}$ is defined as follows, for $J \subseteq Ag$:

- $pre((0, J)) = \neg\varphi \wedge \bigwedge_{a \in J} h_a \wedge \bigwedge_{a \notin J} \neg h_a$;
- $pre((1, J)) = \varphi \wedge \bigwedge_{a \in J} h_a \wedge \bigwedge_{a \notin J} \neg h_a$;
- $pre(s_\top) = \top$.

This event model contains $2^{|Ag|+1} + 1$ events [7]. The preconditions of these events express whether the announced φ is true (i.e., whether it occurs positively or negatively in the precondition) and whether each agent a is attentive or not (i.e., whether h_a occurs positively or negatively in the precondition). We now briefly explain the intuition behind the edges of the model, but refer to [7] for more details. The elements of Q_a of the form $((i, J), (1, K))$ encode the following: Provided that agent a is attentive (i.e., $a \in J$), she believes that any event with precondition φ could be the actual one. The elements of Q_a of the form $((i, J), s_\top)$ then encodes: If instead she is not paying attention (i.e., $a \notin J$), she keeps the beliefs she had before the announcement (represented by the event $s_\top$ having the precondition $\top$. The $s_\top$ event induces a copy of the original model, thereby modeling the "skip" event where nothing happens).

In the following, for any set S, we use id_S to denote the identity function on S, i.e., $id_S(s) = s$, for all $s \in S$. From now on, most of our event models will be of a particular form where the set of events is a set of (conjunctive) formulas and where preconditions are given by the identify function on E, i.e., $pre = id_E$ (meaning that the events are their own preconditions). Our principle-based version of $\mathcal{E}(\varphi)$ is then the following.

Definition 3.2 (Principle-Based Event Model $\mathcal{E}'(\varphi)$). Given a $\varphi \in \mathcal{L}$, the multi-pointed event model $\mathcal{E}'(\varphi) = ((E, Q, id_E), E \setminus \{\top\})$ is:

$E = \{\psi \wedge \bigwedge_{a \in J} h_a \wedge \bigwedge_{a \notin J} \neg h_a : \psi \in \{\varphi, \neg\varphi\}, J \subseteq Ag\} \cup \{\top\}$;

Q_a is such that $(e, f) \in Q_a$ iff all the following are true:

- Basic Attentiveness: if $h_a \in e$, then $\varphi \in f$;
- Inertia: if $h_a \notin e$, then $f = \top$.

The *edge principles* of the model above are Basic Attentiveness and Inertia, describing the conditions under which there is an edge from e to f for agent a. Note that we have exactly the same set of event preconditions in $\mathcal{E}'(\varphi)$ as in $\mathcal{E}(\varphi)$. The difference is just that we define the events to be their own preconditions, which is possible since all pairs of events have distinct and mutually inconsistent preconditions. It's easy to check that $\mathcal{E}(\varphi)$ and $\mathcal{E}'(\varphi)$ also have the same edges, hence the models are isomorphic (we prove this in the Supplementary Material). What an agent considers possible after the announcement is now clearly and explicitly captured by the principles. By Basic Attentiveness, paying attention implies that, in all events considered possible, the announcement is true—and hence attentive agents believe what is announced. By Inertia, inattentive agents believe nothing happened, namely they maintain the beliefs they had before the announcement was made.

Compare the edge specification from $\mathcal{E}(\varphi)$ with the one from $\mathcal{E}'(\varphi)$. We are defining the same set of edges, but whereas the definition of Q_a in $\mathcal{E}(\varphi)$ does not make it immediately clear what those edges are encoding, we believe that our definition of Q_a in $\mathcal{E}'(\varphi)$ does. It is simply two basic principles, one specifying what events are considered possible by the agents paying attention (Basic Attentiveness), and another specifying the same for those not paying attention (Inertia). Even though from a technical viewpoint it is not a big step to introduce such principles, we find it helpful to be able to specify the relevant event models in a clear and concise manner. This makes it easier to use the model and build on it—as should become evident when we later generalise the event model.

3.1.1 Modified model. We now introduce a variant of the event model $\mathcal{E}'(\varphi)$ from Def. 3.2, one that is more appropriate for the types of scenarios that we would like to be able to model.

Truthful announcements. As the present work aims at modeling (noise-free) attention to external stimuli from the environment, in particular visual attention, the first assumption we give up is that announcements may be false. More precisely, we assume that if an agent pays attention to p and the truth-value of p is being revealed, then the agent sees the true truth-value of p. The new event model for announcing φ should then only contain events where φ is true:

$$E = \{\varphi \land \textstyle\bigwedge_{a\in J} \mathsf{h}_a \land \bigwedge_{\mathsf{h}_a \notin J} \neg \mathsf{h}_a : J \subseteq Ag\} \cup \{\top\}.$$

Learning that you were attentive. An assumption we have already given up is attention introspection, so in our models the agents may falsely believe to be paying attention (see Example 2.5). In this setting, it is very plausible to assume that, besides learning what the true event is, attentive agents also learn that they were attentive. This does not happen in the event model $\mathcal{E}'(\varphi)$. We thus substitute Basic Attentiveness with the following principle:

- Attentiveness: if $\mathsf{h}_a \in e$, then $\mathsf{h}_a, \varphi \in f$.

Summing up, the event model where announcements are truthful and attentive agents learn that they paid attention, looks as follows.

Definition 3.3 (Truthful and Introspective Event Model $\mathcal{E}''(\varphi)$). Given $\varphi \in \mathcal{L}$, the multi-pointed event model $\mathcal{E}''(\varphi) = ((E, Q, id_E), E \setminus \{\top\})$ is defined by:

$E = \{\varphi \land \bigwedge_{a\in J} \mathsf{h}_a \land \bigwedge_{a\notin J} \neg \mathsf{h}_a : J \subseteq Ag\} \cup \{\top\}$;

Q_a is such that $(e, f) \in Q_a$ iff all the following are true:

- Attentiveness: if $\mathsf{h}_a \in e$, then $\mathsf{h}_a, \varphi \in f$;
- Inertia: if $\mathsf{h}_a \notin e$, then $f = \top$;

The event model $\mathcal{E}''(p \land g)$ with $Ag = \{a, b\}$ is shown in Figure 2.

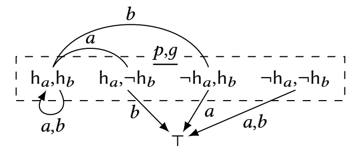

Figure 2: The event model $\mathcal{E}''(p \land g)$ with $Ag = \{a, b\}$. As our event models will have conjunctive preconditions, all distinct, and our events are their own preconditions, we can represent events by lists of formulas, the formulas contained in the event precondition. All other conventions are as for Kripke models (see Fig. 1).

3.2 Event Models for Propositional Attention

In this section, we introduce event models for agents that only pay attention to subsets of At. As our main aim is to model attention to external stimuli, we are interested in modeling the "announcement" of a conjunction of literals $(\neg)p_1 \land \cdots \land (\neg)p_n$, which we interpret as the parallel exposure to multiple stimuli (the truth value of all p_i being revealed concurrently). It could for instance be that we see a video that has 15 ball passes and a gorilla passing by, and that would correspond to the "announcement" $p \land g$, cf. Example 2.5.

Definition 3.4 (Propositional Attention Event Model $\mathcal{F}(\varphi)$). Let $\varphi = \ell(p_1) \land \cdots \land \ell(p_n) \in \mathcal{L}$, where for each p_i, either $\ell(p_i) = p_i$ or $\ell(p_i) = \neg p_i$. The multi-pointed event model $\mathcal{F}(\varphi) = ((E, Q, id_E), E_d)$ is defined by:

$$E = \{\bigwedge_{p\in S} \ell(p) \land \bigwedge_{a\in Ag} (\bigwedge_{p\in X_a} \mathsf{h}_a p \land \bigwedge_{p\in S\setminus X_a} \neg \mathsf{h}_a p) : S \subseteq At(\varphi) \text{ and for all } a \in Ag, X_a \subseteq S\}$$

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

- Attentiveness: if $\mathsf{h}_a p \in e$ then $\mathsf{h}_a p, \ell(p) \in f$;
- Inertia: if $\mathsf{h}_a p \notin e$ then $\ell(p) \notin f$;

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

In $\mathcal{F}(\varphi)$ we have, for each subset of literals in φ, an event containing those literals in the precondition. For those literals, the event also specifies whether each agent is paying attention to it or not. In this way, events account for all possible configurations of attention to any subset of the announcement and for the learning of truthful information regarding it. The edges are again given by two simple principles. Attentiveness states that if an agent pays attention to a specific atom, then she learns the literal in the announcement corresponding to it and that she was paying attention to it. Inertia says that if an agent doesn't pay attention to an atom, then she will not learn anything about it. As we take announcements as truthful revelations, the set of designated events only contains events where all the announced literals are true. The event model $\mathcal{F}(\varphi)$ with $\varphi = p \land g$ and $Ag = \{a, b\}$ is shown in Figure 3.

Example 3.5. Continuing Example 2.5, Ann and Bob have finished watching the Invisible Gorilla video (event model $\mathcal{F}(p \land g)$). As Bob expected, Ann learns that there are 15 ball passes, but she still doesn't know anything about whether there is a gorilla in the video, and believes Bob is in the same situation as herself. The pointed Kripke model $(\mathcal{M}', w') = (\mathcal{M}, w) \otimes \mathcal{F}(p \land g)$ in Figure 4

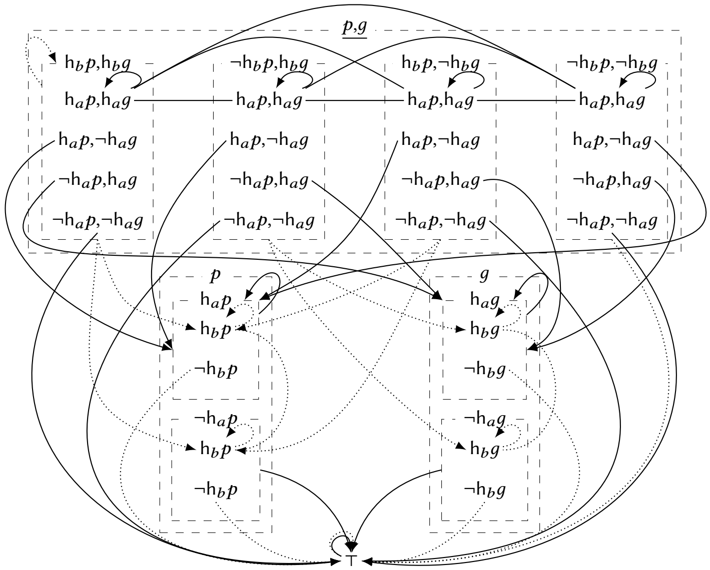

Figure 3: The event model $\mathcal{F}(p \wedge g)$ with $Ag = \{a, b\}$. Solid arrows are for agent a, dotted for agent b.

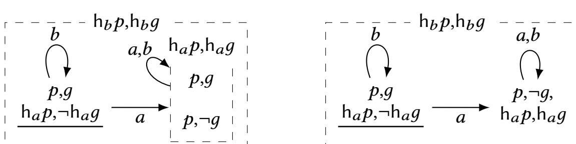

Figure 4: Pointed Kripke models $(\mathcal{M}', w') = (\mathcal{M}, w) \otimes \mathcal{F}(p \wedge g)$ (left) and $(\mathcal{M}'', w'') = (\mathcal{M}, w) \otimes \mathcal{E}(p \wedge g, d)$ (right), where the default map d is $d_a(p) = d_b(p) = \top$ and $d_a(g) = d_b(g) = \neg g$.

(left) represents the situation after exposure to the video, i.e., after the revelation of $p \wedge g$. Ann has only learnt about p and still has no information about g. We thus have $(\mathcal{M}', w') \models B_a p \wedge \neg B_a g \wedge \neg B_a \neg g$. Moreover, she wrongly believes Bob too hasn't received any information about the gorilla, so $(\mathcal{M}', w') \models B_b g \wedge B_a(\neg B_b g \wedge \neg B_b \neg g)$.

4 AXIOMATIZATION

We move to the axiomatization of our logic and show that it is sound and complete. The axiomatization is given by the set of axioms and inference rule of Table 1. It comprises standard axioms and inference rules for normal modal logic as well as reduction axioms. All propositional reduction axioms in Table 1 are for state-eliminating updates, in that they are relativized to the announced φ. Then, the only non-standard axiom we introduce is the one expressing the consequences of attention-dependent announcements for what concerns agents' beliefs. Where $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$ is

Table 1: The logic of propositional attention Λ. It is assumed that $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$ for some literals $\ell(p_i)$, $i = 1, \ldots, n$.

All propositional tautologies
$B_a(\varphi \to \psi) \to (B_a\varphi \to B_a\psi)$
$[\mathcal{F}(\varphi)]p \leftrightarrow (\varphi \to p)$
$[\mathcal{F}(\varphi)]\neg\psi \leftrightarrow (\varphi \to \neg[\mathcal{F}(\varphi)]\psi)$
$[\mathcal{F}(\varphi)](\psi \wedge \chi) \leftrightarrow ([\mathcal{F}(\varphi)]\psi \wedge [\mathcal{F}(\varphi)]\chi)$
$[\mathcal{F}(\varphi)]B_a\psi \leftrightarrow (\varphi \to \bigvee_{S \subseteq At(\varphi)}((\bigwedge_{p \in S} \mathsf{h}_a p \wedge \bigwedge_{p \in At(\varphi)\setminus S} \neg \mathsf{h}_a p) \to B_a([\mathcal{F}(\bigwedge_{p \in S} \ell(p))]\psi)))$
From φ and $\varphi \to \psi$, infer ψ
From φ infer $B_a\varphi$
From $\varphi \leftrightarrow \psi$, infer $\chi[\varphi/p] \leftrightarrow \chi[\psi/p]$[5]

the announced formula, the axiom is the following:

$$[\mathcal{F}(\varphi)]B_a\psi \leftrightarrow (\varphi \to \bigvee_{S \subseteq At(\varphi)} (\bigwedge_{p \in S} \mathsf{h}_a p \wedge \bigwedge_{p \in At(\varphi)\setminus S} \neg \mathsf{h}_a p) \to B_a([\mathcal{F}(\bigwedge_{p \in S} \ell(p))]\psi)))$$

The axiom can be read as saying that after exposure to the revelation of φ, agent a believes that only the conjunction of literals from φ to which she was paying attention to has been revealed.

Theorem 4.1. *The axiomatization in Tbl. 1 is sound and complete.*

Proof sketch. (Full proof in Supplementary Material.) Completeness proceeds by usual reduction arguments [19]. For soundness, the only complicated axiom is then one for attention-based belief update. The axiom relates updating with the event model $\mathcal{F}(\varphi)$ (left-hand side of biimplication) to updating with *another* event model $\mathcal{F}(\bigwedge_{p \in S} \ell(p))$ (right-hand side of biimplication). One of the crucial aspects of the proof is to show that these two event models update in a way that preserves modal equivalence for the worlds that are a-accessible from the designated worlds of the respective updates, which we show via bisimulation. This holds as the axiom specifies the attention profile of agent a with respect to φ, and we demonstrate it holds using a bisimulation construction. □

5 DEFAULTS

In event models for propositional attention, inattentive agents maintain their beliefs about what has been announced but they did not attend. Then, agents like Ann, who didn't hold any particular belief about the gorilla before watching the video and did not notice any while watching it, will not have any particular belief about it after having watched the video either. While this specific way of updating beliefs may be realistic and even rational in some cases, in many others, humans seem to update differently. As said in the introduction, in inattentional blindness situations agents that did not pay attention to an event and received no information about it often believe that the event did not happen. In these situations, agents seem to update their beliefs with respect to unattended events as well, regardless of whether their experience of the situation actually contained any evidence about them.

In this section we propose to account for these specific belief updates by introducing *default values.* A default value for an atom

q is either q, $\neg q$ or $\top$. If q has default value q for agent a in a given announcement, it means that, in lack of evidence about q, agent a will believe q to be true. If q means "the basketball players are wearing shoes", then an agent seeing the video might start to believe q even without actually having paid attention to q, but just assuming q to be true, as it would normally be true in such circumstances. Similarly, if q means "a gorilla is passing by", then agent a might have $\neg q$ as the default value: if the occurrence of a gorilla is not paid attention to, the agent will believe there was none. Finally, if q takes default value $\top$, it means that the agent doesn't default to any value, but preserves her previous beliefs. Maybe she has no strong beliefs about whether all the basket ball players are wearing white, and hence if q denotes that they are all wearing white, her default value for q would be $\top$. We can think of default values as representing some kind of qualitative priors: They encode what an agent believes about what normally occurs in a given situation, and where those beliefs are sufficiently strong to let agent update her beliefs using these priors even when no direct evidence for or against them is observed (paid attention to).

Definition 5.1 (Default Event Model $\mathcal{E}(\varphi, d)$). Suppose $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$, and suppose that d is a *default map*: to each agent a and atom p_i, d assigns a *default value* $d_a(p_i) \in \{p_i, \neg p_i, \top\}$. The *default event model* $\mathcal{E}(\varphi, d) = ((E, Q, id_E), E_d)$ is:

$$E = \{ \bigwedge_{p \in S} \ell(p) \wedge \bigwedge_{p \in At(\varphi) \setminus S} d_b(p) \wedge \bigwedge_{a \in Ag} \Big(\bigwedge_{p \in X_a} \mathsf{h}_a p \wedge \bigwedge_{p \in S \setminus X_a} \neg \mathsf{h}_a p \Big) : b \in Ag, S \subseteq At(\varphi) \text{ and for all } a \in Ag, X_a \subseteq S \}$$

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

- Attentiveness: if $\mathsf{h}_a p \in e$ then $\mathsf{h}_a p, \ell(p) \in f$;
- Defaulting: if $\mathsf{h}_a p \notin e$ then $d_a(p) \in f$.

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

Default event models differ from event models for propositional attention in that if an event in a default model does not contain a literal from the announced formula, then it contains its default value for one of the agents. Each event contains default values for one agent only, so that no event may contain contradicting default values. The accessibility relations are given by similar principles as above, with the difference that the second principle is now called Defaulting, and this principle implies that inattentive agents only consider possible the default values of what they left unattended. Note that defaults are common knowledge among the agents (the event model doesn't encode any uncertainty about the default map d). Figure 4 (right) illustrates the revised update of our initial model with the default event model representing Ann seeing the video. In lack of attention to g, she defaults to $\neg g$, the intuition being that she believes that she would see the gorilla had it been there. She comes to believe there is no gorilla: $(\mathcal{M}'', w'') \vDash B_a \neg g$.

Axiomatization. The axiomatization of the logic for propositional attention with defaults is given by the same axioms as in Table 1, except for the axiom for belief dynamics which is replaced by the following axiom where inattentive agents adopt the default option for the unattended atoms (where $\varphi = \ell(p_1) \wedge \cdots \wedge \ell(p_n)$). For $\varphi_{Sd} = \bigwedge_{p \in S} \ell(p) \wedge \bigwedge_{p \in At(\varphi) \setminus S} d_a(p)$, call the resulting table *Table 2*:

$$[\mathcal{E}(\varphi, d)] B_a \psi \leftrightarrow (\varphi \rightarrow \bigvee_{S \subseteq At(\varphi)} ((\bigwedge_{p \in S} \mathsf{h}_a p \wedge \bigwedge_{p \in At(\varphi) \setminus S} \neg \mathsf{h}_a p) \rightarrow B_a([\mathcal{E}(\varphi_{Sd}, d)]\psi)))$$

Theorem 5.2. *The axiomatization in Tbl. 2 is sound and complete.*

Proof sketch. (Full proof in Supplementary Material.) By a proof following the same structure as the proof of Theorem 4.1, we get that the resulting axiomatization is sound and complete. □

Example 5.3. In the introduction, we mentioned the potential application of our models for human-robot collaboration. Consider an emergency scenario with a mixed human-robot rescue team including a human doctor a and an assisting robot b. Suppose a is attending to an injured victim and that b is ready to assist. While she is attending to the victim, fire breaks out and creates a dangerous situation. The doctor, being absorbed in trying to help the victim, has not noticed the fire, and so it makes sense for the robot to inform her. This scenario is completely equivalent to the invisible gorilla example with p instead meaning, say, "the victim is injured" and g meaning "fire has broken out". The point is that the after fire has broken out, we are in the situation of Figure 4 (right) where $g \wedge B_b B_a \neg g$ holds: The robot correctly believes that the doctor has a false belief that there is no fire. A proactive robot should inform its human team members about any false beliefs that could lead to catastrophic outcomes. This requires the ability of the robot to model those false beliefs, including false beliefs arising due to inattentional blindness, which is exactly what our models provide.

6 SYNTACTIC EVENT MODELS

The event models introduced above are rather large. The event models for propositional attention grow exponentially with the number of agents: For each subset of agents $A \subseteq Ag$ and each announced atom p, it contains at least one event where all $\mathsf{h}_a p$, $a \in A$ occur positively, and all $\mathsf{h}_a p$, $a \in Ag \setminus A$ occur negatively. They also grow exponentially in the number of announced atoms: For each subset S of atoms in the announced formula φ, it contains at least one event in which the set of propositional atoms occurring is exactly S. However, note that we still managed to represent the event models in a relatively compact way in terms of a set of precondition formulas and a list of simple edge principles. This leads us to the following questions. Can we represent *any* event model—or at least a sufficiently general subclass of them—in terms of a set of precondition formulas and a set of edge principles? If so, can we then use this to define syntactically represented event models where the edges are defined by formulas representing the edge principles? This would give us a formally more precise way of handling principle-based event models. Would that then lead to more succinctly represented event models?

We are not the first to consider ways to represent event models succinctly and syntactically. Aucher [1] defined a language with special atoms p'_φ meaning "φ is the precondition of the current event". However, to be able to represent our edge principles via formulas, we need to be able to reason about the structure of the event preconditions, for instance when we want to say that some literal is contained in a precondition (like $\mathsf{h}_a p \in e$). Therefore it doesn't suffer for our purposes to introduce formulas where the preconditions are treated as atomic entities. Another approach is by Charrier and Schwarzentruber [9]. In their language, it is possible to reason about the precondition formulas, for instance the formula $(p_e \to p) \wedge (p_f \to \top)$ can be used to express that event e has precondition p and event f has precondition $\top$. They then represent edges by a program in PDL (propositional dynamic logic). This gives a very imperative representation of the edges, whereas we are here looking for a more declarative representation matching the edge principles introduced above.

We now introduce a new formal language to be used to describe event models. Where $\psi \in \mathcal{L}$, the *event language* $\mathcal{L}_{\mathcal{E}}$ is:

$$\varphi ::= \psi \Rightarrow \mathsf{e} \mid \mathsf{e} \Rightarrow \psi \mid \neg\varphi \mid \varphi \vee \varphi \mid \Box\varphi$$

The formula $\psi \Rightarrow \mathsf{e}$ is read as "ψ implies the precondition of the (current) event" and $\mathsf{e} \Rightarrow \psi$ as "the precondition of the (current) event implies ψ". We will use $e \Leftrightarrow \psi$ as shorthand for $\psi \Rightarrow \mathsf{e} \wedge \mathsf{e} \Rightarrow \psi$. Formulas of $\mathcal{L}_{\mathcal{E}}$ are to be evaluated in single-agent event models, since we are going to specify the edge principles for each agent a by a separate formula φ_a of $\mathcal{L}_{\mathcal{E}}$.

Definition 6.1 (Satisfaction). Let $\mathcal{E} = (E, Q, pre)$ be a single-agent event model over $\mathcal{L}$ (so $Q \subseteq E^2$). For any $e \in E$, satisfaction of $\mathcal{L}_{\mathcal{E}}$-formulas in $\mathcal{E}$ is given by the following clauses extended with the standard clauses for the propositional connectives:

$$\begin{array}{lll} (\mathcal{E}, e) \models \psi \Rightarrow \mathsf{e} & \text{iff} & \models \psi \to pre(e); \\ (\mathcal{E}, e) \models \mathsf{e} \Rightarrow \psi & \text{iff} & \models pre(e) \to \psi; \\ (\mathcal{E}, e) \models \Box\varphi & \text{iff} & (\mathcal{E}, f) \models \varphi \text{ for all } (e, f) \in Q. \end{array}$$

A formula ψ is called *valid* in $\mathcal{E} = (E, Q, pre)$ if $(\mathcal{E}, e) \models \psi$ holds for all $e \in E$. We then write $\mathcal{E} \models \psi$. To have a convient notation for reasoning about what holds true for a single event with precondition $\varphi \in \mathcal{L}$, we introduce the following notation, where $\psi \in \mathcal{L}_{\mathcal{E}}$:

$$\varphi \models \psi \quad \text{iff} \quad ((\{\varphi\}, \emptyset, id_{\{\varphi\}}), \varphi) \models \psi$$

Note that the e in the syntax is bound to the event e at which the formula is evaluated. So $\mathsf{e} \Rightarrow p \to \Box\mathsf{e} \Rightarrow \neg p$ means that if the precondition of the current event implies p, then the precondition of any accessible event implies $\neg p$. Concerning the notation $\varphi \models \psi$, note that we for instance have $p \wedge q \models \mathsf{e} \Rightarrow p \wedge \mathsf{e} \Rightarrow q$: Both p and q are implied by an event with precondition $p \wedge q$. Note that the $\Rightarrow \mathsf{e}$ operator is not truth-functional: For instance we have $\top \models \mathsf{e} \Rightarrow (p \vee \neg p)$, but we don't have $\top \models \mathsf{e} \Rightarrow p \vee \mathsf{e} \Rightarrow \neg p$.

Example 6.2. Consider the event model $\mathcal{E}'(\varphi)$ of Definition 3.2 for some $\varphi \in \mathcal{L}$ where $Ag = \{a\}$. By Inertia, if e is an event not containing h_a, then for any other event f with $(e, f) \in Q_a$, we have $f = \top$. We can express this using an $\mathcal{L}_{\mathcal{E}}$-formula: $\neg\mathsf{e} \Rightarrow \mathsf{h}_a \to \Box\mathsf{e} \Leftrightarrow \top$. The formula says: if h_a is not implied by the precondition of the current event, then any accessible event has a precondition equivalent to $\top$. The formula is simply Inertia expressed in $\mathcal{L}_{\mathcal{E}}$, and we have $\mathcal{E}'(\varphi) \models \neg\mathsf{e} \Rightarrow \mathsf{h}_a \to \Box\mathsf{e} \Leftrightarrow \top$.

When trying to come up with a new way of representing event models syntactically, there is a trade-off between generality and expressivity on one side and succinctness and elegance on the other. The more general a class of event models we want to be able to describe, the more complex the language might have to be and the longer and more complicated the formulas might become. Here we will aim for keeping things simple, even if it implies less generality. For instance, opposite the approach of [9], we decided not to include propositional atoms in $\mathcal{L}_{\mathcal{E}}$ for referring to the names of specific events. This limits expressivity, as then the language can only distinguish events by their preconditions and can not represent distinct events with the same precondition. However, for the event models of this paper, this is not a limitation.

We move to define our syntactic event models. To make the distinction clear, we will now refer to the standard event models of Definition 2.2 as *semantic event models*.

Definition 6.3. A *syntactic event model* is a pair $\mathcal{G} = (\psi_E, (\psi_a)_{a \in Ag})$, where all the ψ formulas belong to $\mathcal{L}_{\mathcal{E}}$. The semantic event model $\mathcal{H} = (E, Q, id_E)$ *induced* by $\mathcal{G}$ is defined as follows:

- $E = \{\varphi \in \mathcal{L} : \varphi$ is a conjunction of literals s.t. $\varphi \models \psi_E\}$;
- For all $a \in Ag$, Q_a is the largest subset of E^2 satisfying $(E, Q_a, id_E) \models \psi_a$. If such a unique largest set doesn't exist, let Q_a be the empty set.

Where $\psi_{E_d} \in \mathcal{L}_{\mathcal{E}}$, we call $(\mathcal{G}, \psi_{E_d})$ a *syntactic multi-pointed event model.* The *induced* multi-pointed event model of $(\mathcal{G}, \psi_{E_d})$ is $(\mathcal{H}, E_d)$ where $\mathcal{H}$ is the event model induced by $\mathcal{G}$ and $E_d = \{\varphi \in \mathcal{L} : \varphi$ is a conjunction of literals s.t. $\varphi \models \psi_{E_d}\}$.

Example 6.4. Consider again the event model $\mathcal{E}'(\varphi)$ of Def. 3.2, where we here let $\varphi = q$, assume $At = \{q\}$ and assume Ag to be any set of agents. Then $\mathcal{E}'(\varphi)$ is induced by the syntactic event model $\mathcal{G} = (\psi_E, (\psi_a)_{a \in Ag})$ defined as follows:

$$\psi_E = \mathsf{e} \Leftrightarrow \top \vee \big((\mathsf{e} \Rightarrow q \vee \mathsf{e} \Rightarrow \neg q) \wedge \textstyle\bigwedge_{a \in Ag}((\mathsf{e} \Rightarrow \mathsf{h}_a) \vee (\mathsf{e} \Rightarrow \neg\mathsf{h}_a)\big)$$
$$\psi_a = (\mathsf{e} \Rightarrow \mathsf{h}_a \to \Box\mathsf{e} \Rightarrow q) \wedge (\neg\mathsf{e} \Rightarrow \mathsf{h}_a \to \Box\mathsf{e} \Leftrightarrow \top).$$

The definition of ψ_E states that any event is either (equivalent to) $\top$ or else: 1) it implies either q or $\neg q$ and, 2) for all $a \in Ag$, it implies either h_a or $\neg\mathsf{h}_a$. Note that since the induced event model is always a model over a set of conjunctive preconditions, we can reformulate this as follows: ψ_E states that any event is either $\top$ or else 1) it contains either q or $\neg q$ and, 2) for all $a \in Ag$, it contains either h_a or $\neg\mathsf{h}_a$. Comparing with Definition 3.2, we see that this is exactly how we defined the set of events of this model. Concerning ψ_a, we earlier concluded that the second conjunct expresses Inertia. The first conjunct expresses Basic Attentiveness.

The *size* of a syntactic event model is the sum of the lengths of the formulas it consists of. We say that two semantic event models $\mathcal{E} = (E, Q, pre)$ and $\mathcal{E}' = (E', Q', pre')$ are *equivalent* if there exists $e \in E, e' \in E'$ such that for all pointed Kripke models $\mathcal{M} = ((W, V, R), w)$ and all formulas $\varphi \in \mathcal{L}$, $\mathcal{M} \otimes \mathcal{E}, (w, e) \models \varphi$ iff $\mathcal{M} \otimes \mathcal{E}', (w, e') \models \varphi$.[6] We can now prove an exponential succinctness result for syntactic event models. We show that for all $n \geq 1$, we can construct a particular syntactic event model $\mathcal{G}(n)$ that can't be represented by any semantic event model with less than 2^n events.

Proposition 6.5 (Exponential succinctness). *There exists syntactic event models $\mathcal{G}(n)$, $n \geq 1$, such that all of the following holds:*

- *$\mathcal{G}(n)$ has size $O(n)$.*
- *The semantic event model $\mathcal{H}(n)$ induced by $\mathcal{G}(n)$ has 2^n events (and is hence of size $\Omega(2^n)$).*
- *Any other semantic event model that is equivalent to $\mathcal{H}(n)$ will have at least 2^n events.*

[6] We could have defined this notion equivalently in terms of bisimulations [9], but as we haven't defined bisimulations in this paper, we choose this equivalent formulation [11].

Furthermore, we can construct the $\mathcal{G}(n)$ so that they use only one agent and where n is the number of atomic propositions.

Proof sketch. (Full proof in Supplementary Material.) The syntactic event model of Example 6.4 is of size $O(n)$ where n is the number of agents, and it induces the semantic event model $\mathcal{E}'(q)$ of Definition 3.2. However, [9] shows that any semantic event model equivalent to $\mathcal{E}'(q)$ has at least 2^n events. This construction uses n agents, but we can make the proof work with only one agent, by instead employing the semantic event model used in the succinctness proof for arrow updates [11]. □

In the proof above we refer to a result [9] showing that their succinct event models are exponentially more succinct than semantic event models. However, their representation of the event models of Definition 3.2 are of size $O(n^2)$ with n being the number of agents, whereas our syntactic event model are of size $O(n)$, hence even more compact (their PDL program for each agent has length $O(n)$, whereas our corresponding formula ψ_a is of constant length).

Note that the semantic event model induced by a syntactic event model is of a particular form where the event preconditions are conjunctions of literals. While this limits the kind of semantic event models we can represent using syntactic event models, it still covers a fairly generous class. Any semantic event model $\mathcal{H} = (E, Q, pre)$ where events are conjunctions of literals (as all event models of this paper are) can be turned it into a syntactic event model $\mathcal{G} = (\psi_E, (\psi_a)_{a \in Ag})$ by simply letting $\psi_E = \bigvee_{e \in E} \mathsf{e} \Leftrightarrow e$ and $\psi_a = \bigwedge_{e \in E} \big(\mathsf{e} \Leftrightarrow e \rightarrow \Box \bigvee_{(e,f) \in Q_a} \mathsf{e} \Leftrightarrow f\big)$.

We are here only using syntactic event models to provide simple and succinct representations of our semantic event models (that are otherwise of exponential size). However, it is relevant to mention that these new syntactic event models could potentially also be interesting for other reasons. As mentioned in [2], there has been quite a lot of resistance to DEL based on (semantic) event models, since one is "mixing syntax and semantics" (due to event models being semantic objects, but still appearing inside modal operators in the language). A syntactic event model clearly does not have this problem, as it's a representation using a sequence of formulas from the language $\mathcal{L}_{\mathcal{E}}$. All formulas are from the same language, so by slightly extending it, we could even represent an event model syntactically by a single formula of such an extended language.

7 RELATED AND FUTURE WORK

This work has built on a previous model for attentive agents [7], generalising the framework to model (1) agents who may pay attention to strict subsets of propositions; (2) agents who may default to specific truth values for the atomic formulas they failed to attend.

What we here call *attention* is similar to what has been called *observability* in the AI and DEL literature. Observability can be attached to different aspects of the world: to propositional atoms [8, 10], to actions [7], to actions of agents [5], or to particular actions [3], and the same holds for attention. However, conceptually, attention and observability are not exactly the same, and in this paper we have been focusing on representing attention to propositional atoms mainly as a starting point for a richer model of attention.

In addition, we proposed a syntactic description of event models that, besides working towards settling the mixture of syntax and semantics typical of DEL, allowed us to reach an exponential succinctness result. There is a clear relation to generalized arrow updates [12], but we conjuncture that our syntactic event models can be even more succinct than generalized arrow updates. We however leave this for future work.

The examples provided in this paper are arguably toy examples in the sense of involving few agents (2) and few propositional atoms (also 2). Since the semantic event models grow exponentially in both the number of agents and propositional atoms, the semantic representation doesn't scale well. However, the syntactic representation does, and in future work we'd like to consider whether we can define a product update directly in terms of syntactic event models to allow for better scalability of our framework.

We also plan to extend the model further to include more core features of attention, for instance an upper bound on the number of atomic formulas that an agent can pay attention to (a bound on the *attention capacity*). This is a simple tweak of the model, but it allows us to capture a lot more: attention as a bounded resource. This can be applied in at least two distinct ways. For instance in Example 5.3, whether the doctor pays attention to the fire breaking out or not might obviously depend on how busy she is attending to other things. This is similar to the invisible gorilla, where attending to the ball passes seems to consume all of the attention capacity. Adding attention capacities would allow the robot to have a more realistic model of human attention and when to intervene. The second use of attention capacities could be to apply it to allow robots to manage their own attention in order to save computational resources. In the DEL literature, attention as a cognitive resource has been explored by [4], where an attention budget and a subjective cost for formulas to be learnt are introduced in the model.

Attention may also relate to the notion of awareness [14], as both concepts can be thought as imposing some limitation on the set of propositions the agent entertains. However, the two also differ: Awareness seems to be more about the propositions that the agent can conceive and thus uses to reason, whereas attention (at least for how we formalised it) is a restriction on what agents perceive of an announcement. In this sense, attention seems to be more about learning dynamics, whereas awareness less so. Future work will explore their relationship.

ACKNOWLEDGMENTS

We gratefully acknowledge Rasmus K. Rendsvig for important inputs in the early phases of this project. We also gratefully acknowledge funding support by the Carlsberg Foundation through The Center for Information and Bubble Studies (CIBS). Finally, we thank the anonymous reviewers for helpful comments and feedback.

REFERENCES

[1] Guillaume Aucher. 2012. DEL-sequents for regression and epistemic planning. *Journal of Applied Non-Classical Logics* 22, 4 (2012), 337–367.
[2] Alexandru Baltag, Lawrence S Moss, and Slawomir Solecki. 2022. Logics for epistemic actions: completeness, decidability, expressivity. *arXiv preprint arXiv:2203.06744* (2022).
[3] Chitta Baral, Gregory Gelfond, Enrico Pontelli, and Tran Cao Son. 2012. An action language for reasoning about beliefs in multi-agent domains. In *Proceedings of the 14th International Workshop on Non-Monotonic Reasoning*, Vol. 4.
[4] Gaia Belardinelli and Rasmus K. Rendsvig. 2021. Epistemic Planning with Attention as a Bounded Resource. In *Logic, Rationality, and Interaction*, Sujata Ghosh and Thomas Icard (Eds.). Springer International Publishing, Cham, 14–30.

[5] Thomas Bolander. 2018. Seeing Is Believing: Formalising False-Belief Tasks in Dynamic Epistemic Logic. In *Outstanding Contributions to Logic*. Number 12. Springer, 207–236.

[6] Thomas Bolander, Lasse Dissing, and Nicolai Herrmann. 2021. 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)*.

[7] Thomas Bolander, Hans van Ditmarsch, Andreas Herzig, Emiliano Lorini, Pere Pardo, and François Schwarzentruber. 2016. Announcements to Attentive Agents. *Journal of Logic, Language and Information* 25 (2016), 1–35. https://doi.org/10.1007/s10849-015-9234-3

[8] Michael Brenner and Bernhard Nebel. 2009. Continual planning and acting in dynamic multiagent environments. *Autonomous Agents and Multi-Agent Systems* 19, 3 (2009), 297–331. https://doi.org/10.1007/s10458-009-9081-1

[9] Tristan Charrier and Francois Schwarzentruber. 2017. A Succinct Language for Dynamic Epistemic Logic.. In *AAMAS*. 123–131.

[10] Wiebe van der Hoek, Nicolas Troquard, and Michael Wooldridge. 2011. Knowledge and control. In *The 10th International Conference on Autonomous Agents and Multiagent Systems–Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 719–726.

[11] Barteld Kooi and Bryan Renne. 2011. Arrow update logic. *The Review of Symbolic Logic* 4, 4 (2011), 536–559.

[12] Barteld Kooi and Bryan Renne. 2011. Generalized arrow update logic. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*. ACM, 205–211. https://doi.org/10.1145/2000378.2000403

[13] A. Mack and I. Rock. 1998. *Inattentional blindness: Perception without attention*. Cambridge, MA: MIT Press.

[14] Burkhard C. Schipper. 2014. Awareness. In *Handbook of Epistemic Logic*, Hans van Ditmarsch, Joseph Y. Halpern, Wiebe van der Hoek, and Barteld P. Kooi (Eds.). College Publications.

[15] Daniel Simons and Christopher Chabris. [n.d.]. *selective attention test*. Youtube. https://www.youtube.com/watch?v=vJG698U2Mvo

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

[17] Daniel J. Simons and Christopher F. Chabris. 2011. What People Believe about How Memory Works: A Representative Survey of the U.S. Population. *PLOS ONE* 6, 8 (08 2011), 1–7. https://doi.org/10.1371/journal.pone.0022757

[18] Daniel J. Simons and Christopher F. Chabris. 2012. Common (Mis)Beliefs about Memory: A Replication and Comparison of Telephone and Mechanical Turk Survey Methods. *PLOS ONE* 7, 12 (12 2012), 1–5. https://doi.org/10.1371/journal.pone.0051876

[19] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. 2007. *Dynamic Epistemic Logic*. Springer Publishing Company. https://doi.org/10.1007/978-1-4020-5839-4

[20] Rineke Verbrugge. 2009. Logic and Social Cognition: The Facts Matter, and so Do Computational Models. *Journal of Philosophical Logic* 38, 6 (2009), 649–680. https://doi.org/10.1007/s10992-009-9115-9