The Logical Development of Pretense Imagination*

Aybüke Özgün[a] and Tom Schoonen[b]

[a] Institute for Logic, Language, and Computation, University of Amsterdam, The Netherlands

[b] Centre for Advanced Studies in the Humanities *Human Abilities*, Humboldt-Universität zu Berlin, Germany

This is a final draft of an article in Erkenntnis
https://doi.org/10.1007/s10670-021-00476-9
**** Please cite the published version.****

Abstract

We propose a logic of imagination, based on simulated belief revision, that intends to uncover the logical patterns governing the *development* of imagination in pretense. Our system complements the currently prominent logics of imagination in that ours in particular formalises (1) the algorithm that specifies what goes on *in between* receiving a certain input for an imaginative episode and what is imagined in the resulting imagination, as well as (2) the goal-orientedness of imagination, by allowing the context to determine, what we call, the *overall topic* of the imaginative episode. To achieve this, we employ well-developed tools and techniques from dynamic epistemic logic and belief revision theory, enriched with a *topicality* component which has been exploited in the recent literature. As a result, our logic models a great number of cognitive theories of pretense and imagination (cf. Currie & Ravenscroft 2002; Nichols & Stich 2003; Byrne 2005; Williamson 2007; Langland-Hassan 2012, 2016).

Keywords: [Logic of Imagination, Pretense Imagination, Dynamic Epistemic Logic, Belief Revision, Aboutness, Topicality]

1 Introduction

Consider the following example of the phenomenon known as *pretense*:

*The authors contributed equally.

> The child is encouraged to 'fill' two toy cups with 'juice' or 'tea' or whatever the child designated the pretend contents of the bottle to be. The experimenter then says, 'Watch this!', picks up one of the cups, turns it upside down, shakes it for a second, then replaces it alongside the other cup. The child is then asked to point at the 'full cup' and at the 'empty cup' (both cups are, of course, really empty throughout). (Leslie, 1994, p. 223)

Children from as young as two years old already consistently point to the cup *that has been turned upside down* when asked to point at the 'empty cup' (cf. Leslie 1994; Nichols & Stich 2003). This indicates that children, at a very young age, are able to engage in pretense that goes against what they believe the world to actually be like. One of the main questions that arises is *how* we develop such a pretend scenario that seems so rational, but is often in contradiction with our actual beliefs: the children actually believe that both cups are empty, yet they behave in pretense *as if* one of the cups is full. They imagine this non-actual scenario in a *reality-oriented* way. Which logical rules, if any, govern the development of such a pretense scenario?

In this paper, we will provide a novel formal model of, what we call, *pretense imagination* based on tools from dynamic epistemic logic and belief revision theory. In the first two sections we review some of the current theories of pretense imagination and distil some of its essential features. We do this to inform our model, empirically and conceptually, of the current philosophical theories of pretense imagination. Then, in Section 4, we introduce branching-time belief revision structures in which the target notion of pretense imagination and a related notion of belief are formalised by means of normal and classical modal operators. In particular, we will use these structures to extract the imaginative episode that an agent is engaged in and that records what the agent *has imagined up until the current moment*. (In this sense, our pretense imagination operator adopts the natural language reading 'the agent has imagined that', rather than 'the agent imagines that'.) In Section 5, we enrich the initially proposed models with a *topicality* component in order to overcome the idealisations that result from the former framework and shortcomings of some previous logics of imagination. Throughout, we discuss a worry for the logic of imagination of Berto (2018, 2021) and show that our model is sufficiently rich to overcome this issue and, thus, provides us with a further step in the right direction toward developing an adequate formalisation of imagination in pretense.

2 Imagination, Belief, and Topics

It is common consensus that the notion of 'imagination' is highly ambiguous and used in many different ways (cf. Kind 2013; Balcerak Jackson 2016). In this paper, we study the kind of imagination that is involved in pretense and pretend play from a logical perspective (see, e.g., Currie & Ravenscroft 2002; Nichols & Stich 2003; and Langland-Hassan 2012, 2016). Therefore we call the kind of imagination that we are interested in *pretense imagination*.[1] In the following subsections, we discuss some essential features of and, in particular, several important factors that restrict pretense imagination.

2.1 Beliefs in Pretense

Consider again the example of the pretend tea-party, as described above. There seem to be two crucial ways in which pretense imagination is *restricted* by the agent's *beliefs*. First, pretense imagination seems to follow *belief-like* patterns, which explains the rational behaviour with respect to which cups are full. Secondly, *background beliefs* can be imported into the imaginative episode, which explains, e.g., the beliefs about the workings of gravity in the pretense.[2] We will discuss these in turn.

The two main theories of pretense – that of Nichols & Stich (2003) and of Langland-Hassan (2012) – have it that the development of pretense is very closely related to our ability to reason about our beliefs.[3] To capture this in our formalisation, we focus on belief and belief revision, where the latter is of *hypothetical* nature hinting at real belief changes were the pretend scenario to be actual. In this sense, it is sufficient to use models and operators that describe a situation where the objective facts of the world do not change but only the pretend belief state of the imagining agent changes. This belief revision process follows, roughly, Ramsey's (1929) famous pattern:

> First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent).
>
> (Stalnaker, 1968, p. 102)

[1] It is an interesting question whether this kind of imagination also features in other cognitive activities besides pretense, such as future planning, decision making, and risk assessment. We believe that it does but neither the conceptual arguments nor the adequacy of the formal model presented in this paper hinges on this view (we will get back to this point in our concluding remarks).

[2] Note that one can imagine recalcitrant situations with respect to both of these restrictions if the agent *explicitly intervenes*. We will set this complication aside for now and address the details of this in the next section, when we elaborate on the theory of imagination at play here.

[3] In the literature concerning pretense and the relevant imagination involved, there is a debate between those claiming that there is nothing over and above the cognitive attitudes belief and desire that is needed to account for what is going on during pretend play (the use of 'desire' here is meant in a non-technical, pre-theoretical sense) and those claiming that there is a specific cognitive capacity, distinct from belief and desire, that is involved (a pretense- or imagination-attitude). The former support the *Single Attitude* (SA) account and the latter support the *Distinct Cognitive Attitude* (DCA) account. Ultimately, the models of this paper can be interpreted in line with either theory.

The second important factor that restricts pretense imagination is the agent's *background beliefs* about the actual world. As Williamson notes, "[o]ne's imagination should not be completely independent of one's knowledge of what the world is like" (2016, p. 114). For example, in the tea-party pretense scenario, the participants continue the pretense imagining that tea falls downwards as opposed to upwards because they *import* their background beliefs about gravitational forces – that unsupported objects fall towards the centre of the Earth – into the pretense. We will use the phrase 'taking on board' to refer to those beliefs that the agent incorporates into the pretense and uses to further the pretend scenario. The agent takes on board, in the imagination, that when full cups are turned upside down their contents fall down. However, why is it that some other background beliefs, for example that Paris is the capital of France, are not taken on board?

We argue that one of the reasons why the agent does not imagine Paris being the capital of France in the tea-party situation is simply that this is *off-topic* and *irrelevant* to the pretend tea-party. This suggests a natural way to separate the background beliefs that can be taken on board in the pretense from the ones that are not: we select the relevant background beliefs to import into pretense based on what they are about, in other words, based on their *topics* (cf. Berto 2018, 2021). We will briefly discuss the theory of topicality, as this will be the second component – next to belief revision – that will feature heavily in our formal framework.

2.2 Aboutness: topicality

In a series of work, Berto (2018, 2021); Berto & Hawke (2018); and Hawke (2018) have developed a general theory of *topic-sensitive* propositional content, which they have applied, in epistemic contexts, to address the problems of logical omniscience (cf. Hawke et al. 2020; Özgün & Berto 2020). We briefly recap the main components of their proposal and refer to the aforementioned sources for a more detailed presentation.

Within pretense imagination, we focus only on *propositional* imagination: imagining *that* such and such is the case. Imagination, as a mental attitude towards propositions, thus ranges over propositional contents that are generally taken to be sets of possible worlds. However, treating propositional content this way leads to too crude an identification of propositions, which causes serious idealisation problems – such as the problems of logical omniscience – for formal representations of mental attitudes. Here is an example. Since 'Extremally disconnectedness is not a hereditary property of topological spaces' and 'Jane is a logician or she is not a logician' are true at exactly the same possible worlds (namely, all), they are treated to represent the same proposition. However, they obviously do not say the same thing as they differ in *topic*: the latter is about **Jane**, **Jane's profession** etc., whereas only the former is about, e.g., **extremally dis-**

connectedness, **hereditary properties**, **topology**, but not about **Jane**. One can grasp facts about Jane without having even heard of what a topology is. Therefore, arguably, we can imagine, believe, or know the latter without imagining, believing, or knowing the former and *vice versa*. While it is difficult, if at all possible, to capture this by using *only* the standard possible worlds semantics and Hintikka's (1962) way of modelling (propositional) mental attitudes as quantification over possible worlds, supplementing the standard possible worlds semantics with an account of *aboutness* – "the relation that meaningful items bear to whatever it is that they are on or of or that they address or concern" (Yablo, 2014, p. 1) – solves the problem to a great extent (cf. Berto 2018, 2021; Berto & Hawke 2018).[4] The content of an interpreted sentence then becomes a pair of its intension and topic. Thus, in particular, imagining a proposition requires also knowing what it is about, i.e., having grasped its topic.

It is common consensus in theories of partial content that truth functional logical connectives do not add anything to the topic of a sentence, that is, they are *topic-transparent* (cf. Fine 1986, 2016; Hawke 2018). Whatever is on topic with 'Jane is a logician' is also on topic with 'Jane is *not* a logician' and *vice versa*. They are about exactly the same things, e.g., **Jane** and **Jane's profession**. Similarly, the topic of 'Jane is a logician and Kate is a philosopher' is the same as 'Jane is a logician or Kate is a philosopher'. It is a *fusion* of the topic of 'Jane is a logician' and 'Kate is a philosopher'. The topic of 'Kate is a philosopher' is *part of* the topic of 'Jane is a logician and/or Kate is a philosopher'. That is, topics can be fused together and include other topics as their proper parts. They stand in a *mereological relation*. All of this will be reflected in the formal models in Section 5.

Overall Topics

Berto (2018, 2021) presents a formalisation of propositional imagination that incorporates a topicality component that represents the topic-sensitivity of (propositional) mental states. While his logics of imagination successfully employ (conditional) modal operators that can discern logically and necessarily equivalent propositions, they fall short of representing the *overall topic* of an imaginative episode, an important factor affecting the development of pretense imagination. To see what we mean by 'overall topic' and how this affects the imagination, consider the following two situations:

> **Context A:**
> You are flying to Australia the day after tomorrow to take a well-deserved holiday. That evening, when watching the news, you find out that it is likely

[4] Some problems remain open, one of which we raise below and aim to address in our model.

> that there will be a tornado in Indonesia and that nothing else is known at this point. You wonder whether the tornado might affect your flight.
>
> **Context B:**
> You have a friend living in Singapore, who lives right by the coast. That evening, when watching the news, you find out that it is likely that there will be a tornado in Indonesia and that nothing else is known at this point. You wonder whether this might affect your friend.

In order to help you evaluate the effects of the tornado in each case, you engage in an imaginative exercise.[5] In particular, in both cases, you use the following explicit input

(1) There is a tornado in Indonesia,

and start the imaginative process to determine the effects thereof. As **Context A** involves holiday planning and **Context B** is concerned with your friend living close to a tornado zone in Indonesia, the imaginings resulting from (1) could be different in **Context A** and **Context B**. For example, imagining 'Booking a flight through the US rather than Indonesia is safer' seems to be *off-topic* in **Context B**, whereas it is *on-topic* in **Context A**.

The above example is no exception, so a formal model of imagination should be able to account for the fact that different contexts – based on the exact same explicit input and background beliefs – might give rise to different imaginative episodes solely due to their distinct *overall topics*. Berto's (2018, 2021) logics of imagination, however, are unable to do so, as these logics only focus on the relationship between the topics of particular input/output propositions and overlook the idea that there might be overall topics to exercises of imagination.[6] We suggest a way forward by not only focusing on the topic of the particular propositions involved, but also adding, what we will call, an *overall topic* to our model of pretense imagination.[7]

---

[5] In both scenarios it is stipulated that a tornado is highly probable but not absolutely certain. This is to make sure that the cognitive exercise at play here is pretense imagination and not mere belief revision, as the agent might not actually believe that there will be a tornado. We thank an anonymous reviewer for pointing this out.

[6] This is inspired by an objection raised against Berto's work by Timothy Williamson when the former presented some of his work at the 'Philosophy of Imagination' conference at the Ruhr Universität Bochum in March 2018.

[7] The particular form of our models is not essential to this enrichment. The same solution could also be implemented in Berto's (2018, 2021) models of imagination.

3 Pretense Imagination

By now we have described two important factors that restrict pretense imagination: belief and topicality. However, we have not yet discussed *what* pretense imagination is, how it functions, and what its crucial features are. We will do so in this section. In pretense imagination, for example in the tea-party example above, the entire episode is made up out of a number of (temporally) shorter instances: the pretending that the cup is being turned upside-down, that the tea is being poured, *et cetera*. These are all 'part' of the entire tea-party pretense. It seems obvious that some of these are explicitly 'intended' by the agent, while others, e.g., that the tea falls to the floor after the cup being held upside-down, develop without any intentions from the agent. Also, it seems that pretense is full of choices from the agents that might go beyond what usually happens at a tea-party; the agent might, for example, imagine a butler coming in to join the party.[8] We discuss these features in turn.[9]

Explicit Input

We consider an *imaginative episode* – e.g., the pretend tea-party – as a sequence of individual *imaginative stages* – e.g., pouring the tea; keeping the cup upside down, *et cetera*. Such an imaginative episode always starts with a particular input, which Langland-Hassan (2016, pp. 65-67) argues to be an *intention* of the agent. The intention that starts the imaginative episode consists of two parts. On the one hand, the intention provides the proposition that starts the imaginative episode. This is the proposition that makes up the first stage in the sequence of imaginative stages. On the other hand, the intention seems to play a role in demarcating what the imaginative episode (as a whole) *is about*. We use the term *input proposition* to refer to the former and *overall topic* to refer to the latter in order to keep these two components clearly separated. An input proposition and overall topic together form the *explicit input* of an imaginative episode.

Internal Development

Given an explicit input, the imaginative episode unfolds. In the case of the pretend tea-party, the development of this kind of imagination seems to follow a pattern that

[8]See Nichols & Stich (2003, pp. 23-24) for empirical evidence that people do make such choices in pretense.

[9]Though most of what is said here is taken from the work of Langland-Hassan (2012, 2016), the resulting general picture (and thus our model thereof) captures most theories of pretense (e.g., that of Currie & Ravenscroft 2002 and Nichols & Stich 2003) and is compatible with particular theories of imagination (e.g., that of Byrne 2005; Williamson 2007, and Berto 2021).

is very similar to that of rational belief revision. As Langland-Hassan puts it: "imagination [...] has its own norms, logic, or algorithm that shapes the sequence of i_x after the initiation of an imagining by a top-down intention" (Langland-Hassan, 2016, p. 67). The development of the imaginative episode is governed by the very same mechanisms that guide the inferences we make in rational belief updates (cf. Byrne 2005; Williamson 2007; Langland-Hassan 2016; Williamson 2016; Berto 2021). We call this kind of development the *internal development* of the imaginative episode. In terms of the tea-party example, this development leads to the agent *automatically* imagining that the tea falls towards the ground when the cup is turned upside down. This nicely allows us to explain some of the features of imagination relating to the *reality-oriented* development of pretense. Moreover, the involuntariness of the internal development explains the non-arbitrary nature of imagination: we are not free to imagine whatever we want given a certain input and topic (cf. Byrne 2005; Kind 2016; Williamson 2016).

Cyclical Interventions

Pretense imagination is generally thought to be extremely flexible, playing an important role in "guiding creative endeavours" of agents (Langland-Hassan, 2016, p. 72). Relatedly, "there is much to be said," Langland-Hassan points out, "for the idea that imagination allows us to audition a *variety* of ways things might go, in order to choose a best course of action" (ibid., original emphasis). In order to explain this flexibility of pretense imagination, we need more than only the internal development, since, given an input p in a situation s, we would expect the internal development to always lead to the same outcome, namely whatever the result of a belief revision with p in s is. This way, we could never test the variety of options given p in s through imagination, nor account for its flexibility.

These variations might occur because agents actively *intervene* into the imaginative episode. They add additional content forcefully (in that it does not necessarily follow from the previous imaginative stage) and this content can go beyond what the agent otherwise would have imagined. So, when testing the variety of potential outcomes given p in s, the agent actively intervenes somewhere in the imaginative episode with additional contents (e.g., q_1, q_2, etc.).[10]

[10] For those who worry about phenomenology of an imaginative episode and the lack of 'active choice' that seems to be involved, note that most of this intervening happens sub- or unconsciously. "What we might pre-theoretically think of as a single imaginative episode could in fact involve many such top-down 'interventions.' These interventions would allow for the overall imagining to proceed in ways that stray from what would be generated if one never so intervened" (Langland-Hassan, 2016, pp. 74-75).

3.1 Essential Features of Pretense Imagination

From this discussion, toward a more systematic approach, we distil the following central features of a theory of pretense (all of which are present in most work on pretense and imagination, e.g., Nichols & Stich, 2003; Langland-Hassan, 2012, 2016; Berto, 2021). We intend to capture all of these in our formal framework.[11]

PI: The imagination involved in pretense is strictly *propositional imagination*. That is, imagining *that* such and so is the case (as opposed to, e.g., sensory or objectual imagination; see Langland-Hassan, 2016; Balcerak Jackson, 2018).[12]

ESP: The pretense always has an *explicit starting point*. This can either be in the form of an explicit external input ('Let's imagine that...') or activated by something that caught the imaginer's attention (e.g., looking at an airplane might start off an imaginative episode where one pretends to be able to fly) (cf. Langland-Hassan 2016).

QU: A crucial feature is, what has been called, *quarantining*. Pretense does not interfere with one's actual beliefs. One can pretend that p is the case irrespective of whether they believe that p or not (cf. Nichols & Stich 2003; Langland-Hassan 2012).

CHO: "[P]retence is full of *choices* that are not dictated by the pretence premise, or by the scripts and background knowledge that the pretender brings to the pretence episode [...] these choices typically get made quite effortlessly" (Nichols & Stich, 2003, p. 35).

RAT: Within the pretense, the agent reasons/behaves *rationally*; pretense seems to follow a 'belief-like' inference pattern (cf. Nichols & Stich 2003; Langland-Hassan 2012, 2016). That is, the agent responds to information in the pretense in a way very similar to how they would respond if the information were actual.

ROI: Pretense involves a form of *reality-oriented imagination*. The imagination involved in pretense is the kind that is, in a sense, restricted by (known) causal laws and that is the same as the imagination that is used to evaluate certain

[11]'Pretense' is usually used to denote the imaginative episode *in combination with* the appropriate physical actions. So, in the case of the tea-party, when one moves their arm in the motion as if sipping tea from an empty cut, this is part of (and often the defining part of) the pretense episode. However, for our purposes, we ignore this part and only focus on the *imagination* that is involved in such pretense.

[12]This is not to say that there is no imagery involved in pretense, what we mean is that the kind of imagination that allows us to explain the pretense behaviour is propositional imagination.

conditionals (e.g., 'what would happen if...') (cf. Byrne 2005; Williamson 2007; Kind 2016).[13]

In the next section, we take a first step towards a full-blown formal model of the logical development of pretense imagination.

4 The Logical Development of Pretense

We propose a formal model of pretense imagination from which we can read off sequences of individual imaginative stages, denoted by $(i_1, \ldots, i_n)$, that form imaginative episodes, $\mathcal{I}$. As the pretense imagination follows 'belief-like' inference patterns and develops in stages, we use a variant of *branching-time belief revision models* introduced by Bonanno (2007). These models "provide a way of modeling the evolution of an agent's beliefs over time in response to informational inputs" (Bonanno, 2012, p. 206). In our framework, the imagined propositions play the role of 'informational inputs' and a simulated belief revision function characterises the way the agent changes their *beliefs in pretense* (called *simulated beliefs*) in light of a new input. We then, in Section 5, enrich these structures with a topicality component, following Berto (2018, 2021), in order to render the imagining agent in question non-omniscient with respect to what they believe and imagine.

4.1 Syntax and (idealised) Semantics

Let $\mathsf{Prop} = \{p_1, p_2, \ldots\}$ be a countable set of propositional variables and $\mathcal{L}$ be the language of classical propositional logic defined on Prop. The language $\mathcal{L}_{\mathsf{BI}}$ of the logic of belief and imagination is then defined by the grammar:

$$\varphi := p_i \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid B_{@}\psi \mid B\psi \mid I\psi$$

where $p_i \in \mathsf{Prop}$ and $\psi \in \mathcal{L}$. We often use $p, q, r, \ldots$ for propositional variables. We will follow the usual rules for the elimination of the parentheses. We employ $\vee, \rightarrow, \leftrightarrow$ as the usual abbreviations as follows: $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \rightarrow \psi := \neg\varphi \vee \psi$, and $\varphi \leftrightarrow \psi := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. We define $\top$ as $p \vee \neg p$ for any $p \in \mathsf{Prop}$, and $\bot := \neg\top$.

[13] There is a very interesting and intricate relationship between ROI and RAT. For our purposes, focusing on 'tea-party-like' examples of pretense imagination, the distinction is relatively intuitive. However, when one considers cases of pretense imagination involving more 'exotic' cases (e.g., 'imagine there is a monster under the bed' or 'imagine that I am Luke Skywalker') more needs to be said. We thank an anonymous reviewer for pressing us on this issue. We will return to it in the conclusion of this paper.

We read '$B_@\psi$' as 'the agent actually believes that ψ'; '$B\psi$' as 'the agent believes in pretense that ψ'; and '$I\psi$' as 'the agent has imagined that ψ'. Note that the modalities $B_@$, B, and I range only over Booleans. That is, our language of belief and imagination expresses only *first-order* attitudes, in line with both the way the cognitive science and philosophy literature examine imagination (cf. Currie & Ravenscroft 2002; Nichols & Stich 2003; Byrne 2005; Williamson 2007; Langland-Hassan 2012, 2016) and the way that AGM theory formalises rational belief revision (cf. Alchourrón et al. 1985; Bonanno 2012). While $B_@$ represents the agent's actual beliefs (outside of pretense, at the initial stage), B refers to the agent's *simulated* beliefs that they come to possess at the stages of an imaginative episode.

We interpret $\mathcal{L}_{\mathsf{BI}}$ in (a version of) branching-time belief revision models, the first component of which consists of a forward-looking branching-time structure.

Definition 4.1. **Next-time Branching Frame**
A *next-time branching frame* is a pair $(S, \rightarrowtail)$, where S is a nonempty set of *stages* and $\rightarrowtail$ is a binary relation on S such that for all $s_1, s_2, s_3 \in S$,

1. if $s_1 \rightarrowtail s_3$ and $s_2 \rightarrowtail s_3$, then $s_1 = s_2$ (no branching to the past);
2. if $(s_1, \ldots, s_n)$ is a sequence in S such that $s_i \rightarrowtail s_{i+1}$ for every $i = 1, \ldots, n-1$, then $s_n \neq s_1$ ($\rightarrowtail$ is strictly a next time relation).

Bonanno (2007) calls the elements of S 'instants' or 'dates', however, we prefer to call them 'stages', as we think of them as imaginative stages in which the agent could be. We read '$s \rightarrowtail s'$' as "s' is an *immediate successor* of s" or "s is *the immediate predecessor* of s'". Every stage has at most a unique immediate predecessor (see Definition 4.1.1), but can have several immediate successors. We work with *rooted* next-time branching frames in order to explicitly mark the actual belief state of the agent. To define a rooted frame, we let $\rightarrowtail^+$ denote the transitive closure of $\rightarrowtail$. A next-time branching frame $(S, \rightarrowtail)$ is *rooted* if there is $s_0 \in S$ such that $s_0 \rightarrowtail^+ s'$ for all $s' \in S$ with $s_0 \neq s'$. We call such an s_0 *the initial stage*. While the root s_0 represents the agent's actual belief state *before* the imaginative episode has started, its $\rightarrowtail^+$-successors are the possible imaginative stages the agent can reach via *simulated* belief revision.

Definition 4.2. **Branching-time Belief Revision Model**
A *branching-time belief revision model* (in short, *model*) is a tuple $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V\rangle$, where

1. W is a nonempty set of *possible worlds*;

2. $\{\preceq_s\}_{s\in S}$ is a set of well-preorders on W, where $\preceq_s$ denotes the well-preorder assigned to stage s. A *well-preorder* on W is a reflexive and transitive binary relation such that every nonempty subset of W has a minimal element, where the set of minimal elements $Min_{\preceq_s}(P)$ for any $P \subseteq W$ is defined as

$$Min_{\preceq_s}(P) = \{w \in P : w \preceq_s v \text{ for all } v \in P\}.$$[14]

3. Let WPre_W be the set of all well-preorders on W. Then, $\mu : \mathsf{WPre}_W \times \wp(W) \to \mathsf{WPre}_W$ is a *simulated revision function* such that for any $\preceq \in \mathsf{WPre}_W$ and nonempty $P \subset W$:

$$\mu(\preceq, P) = (Q \times W\backslash Q) \cup (\preceq \cap (Q \times Q)) \cup (\preceq \cap (W\backslash Q \times W\backslash Q)),$$

for some nonempty $Q \subseteq P$ such that for any $w, v \in P$ if $v \in Q$ and $w \preceq v$, then $w \in Q$ (we call such a Q a *downward* $\preceq$*-closed* subset of P). And, $\mu(\preceq, \emptyset) = \preceq = \mu(\preceq, W)$. We denote $\mu(\preceq, P)$ simply by $\preceq^P$.

4. $(S, \rightarrowtail)$ is a *rooted next-time branching frame* such that if $s \rightarrowtail s'$, then $\preceq_{s'} = \preceq_s^P$ for some $P \subseteq W$.

5. $V : \mathsf{Prop} \to \wp(W)$ is a valuation function that maps every propositional variable in Prop to a set of possible worlds.

'$\preceq_s$' is the *plausibility order at stage* s and represents the arrangement of worlds to the degree that the agent considers them plausible at s. We read '$w \preceq_s v$' as 'w is at least as plausible as v at stage s'. We say 'w and v are *equally plausible* at stage s', denoted by $w \approx_s v$, if $w \preceq_s v$ and $v \preceq_s w$. We define *strict plausibility*, denoted by $\prec_s$, in a usual way as $w \prec_s v$ iff $w \preceq_s v$ and $w \not\approx_s v$. The set $Min_{\preceq_s}(W)$ forms the set of possible worlds the agent considers *most plausible* at s and represents the agent's (simulated) belief set at s. Since each $\preceq_s$ is well-founded, every nonempty subset of W has a minimal element with respect to each $\preceq_s$ – i.e., for all $s \in S$ and $P \subseteq W$ such that $P \neq \emptyset$, $Min_{\preceq_s}(P) \neq \emptyset$.[15] So, for each $s \in S$, $(W, \preceq_s, V)$ constitutes a standard plausibility model for belief (cf. Baltag & Smets 2006; van Benthem 2007).

[14] Every well-preorder $\preceq_s \subseteq W \times W$ is a total order: either $w \preceq_s v$ or $v \preceq_s w$ for all $w, v \in W$.

[15] This condition guarantees that the agent never believes a blatant contradiction and, in turn, never imagines a blatant contradiction such as $\bot$. Note, however, that we think that inconsistent pretence is possible in principle. It is just that the current framework cannot deal with it in a completely satisfactory way. One way to do so, would be to add *impossible* worlds or states to the models (see for example, respectively, Berto 2017; Saint-Germier 2021). See below for more about imagining contradictory propositions within a single imaginative episode. Thanks to an anonymous reviewer for urging us to stress this point.

A branching-time belief revision model is intended to represent the evolution of an agent's simulated beliefs and imagination over time: the root of the model represents the stage the agent has not yet started the imaginative process and the branches of the model represents the possible ways the agent's imagination *can* develop, following the revision policy described by μ (see Definition 4.2.4). We can therefore see the plausibility structure *at the initial stage* as the model that represents the agent's *actual belief state* and their *pretend* or *simulated belief states* are represented by the further stages in a branching-time belief revision model. (When it is contextually clear which belief state of the agent we refer to, we usually drop the phrases "actual" and "simulated" and say "belief state" only.) We emphasise that the only component of the model that varies from stage to stage is the plausibility ordering and that the valuation of the atoms stays the same throughout the stages of a branching-time structure. This means that our branching-time structures *represent simulated belief changes of the imagining agent in a world where the objective facts do not change.*

The simulated revision function μ represents the agent's belief revision policy that they follow throughout the process of pretense imagination. It takes a plausibility order on W and a subset P of W and returns another plausibility order such that, for some downward $\preceq$-closed subset Q of P, all Q-worlds become strictly more plausible than all $W \backslash Q$-worlds, and the ordering remains the same within the Q and $W \backslash Q$ zones. Downward closed-ness of Q in P guarantees that the most plausible P-worlds before the revision becomes the most plausible worlds after revision by P. This corresponds to a more general version of the well-known lexicographic upgrade policy. This level of generality allows us to remain agnostic as to which *specific* belief revision policy an agent should/can adopt through an imaginative episode, while constraining what is imagined after the first imaginative input by the agent's actual beliefs (as $\mu(\preceq, P)$ copies $\preceq$ to some extent) and some rationality constraints (see (1) and (2) below). To explain the former, one can put conditions on Q in the definition of μ and model agents who revise their beliefs with respect to a specific, fixed belief revision policy throughout the imaginative episode. For example, when $P = Q$, belief revision function μ corresponds to lexicographic upgrade, which models the belief revision policy of an agent who tends to change their beliefs rather radically, accepting the incoming information without much deliberation. On the other extreme, taking Q to be only the most plausible P-worlds corresponds to a relatively conservative belief revision policy (cf. Boutilier (1994)'s *minimal revision*) that keeps as much of the previous ordering, i.e., of the (simulated or actual) beliefs as possible (cf. van Benthem 2007; van Benthem & Liu 2007; Liu 2008). To explain the latter, it is easy to see, by the definition of μ, that (1) $\mu(\preceq, \emptyset) = \preceq$, (2) $Min_{\mu(\preceq,P)}(W) \subseteq P$ when $P \neq \emptyset$. (1) means that the agent does not revise their (actual or simulated) belief state in light of a contradictory input and (2) guarantees that the new order obtained after revision by a consistent proposition

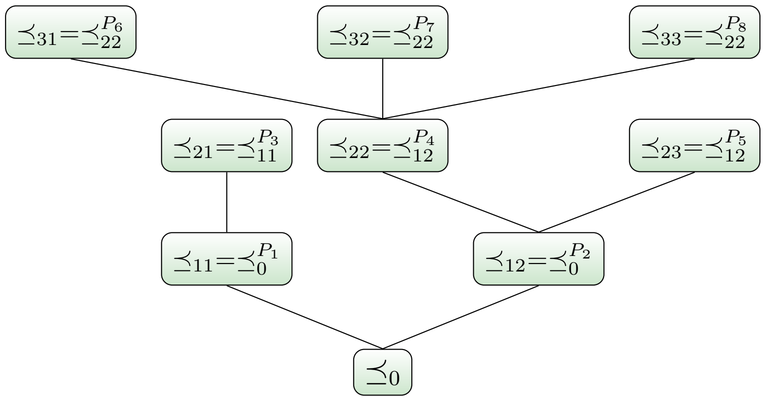


Figure 1: An example of a branching-time belief revision model $\langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V\rangle$, where $P_i \subseteq W$.

P has only P-worlds as most plausible worlds. This corresponds to a restricted version of the *Success Postulate* of the AGM belief revision theory, which claims that the input proposition is believed after revision. In fact, since the weakest form of μ corresponds to minimal revision, μ satisfies all eight AGM postulates (see Boutilier 1994, Theorem 3.25 and Boutilier 1996, Theorem 1).[16] (See Figure 1 for an illustration of a branching-time belief revision model (where $\preceq_{ij}$ represents the plausibility ordering at stage s_{ij}).)

The kind of imagination that we are interested in can be read off from the actual development of the pretense scenario, represented by a finite sequence of linear stages, called *history*, $h = (s_0, s_1, \ldots, s_n)$, such that $s_i \rightarrowtail s_{i+1}$, $n > 0$, and s_0 is the root of the underlying next-time branching frame. We call s_n the *current stage* and s_0 is the *initial stage*. We impose that $n > 0$ since we are interested in the development of pretense imagination and s_0 represents the agent's actual belief state *before* the imaginative episode has started. History h thus keeps track of the past stages, but does not tell us anything about the future. Given a branching-time belief revision model $\mathcal{M}$ and a history $h = (s_0, s_1, \ldots, s_n)$, we will be able to extract the corresponding imaginative episode $\mathcal{I} = (i_1, \ldots, i_n)$ as described in Section 4.2.

We now have the required tools to give the semantics for our language. Formulas of

[16]If we eliminate the downward closed-ness condition of Q in Definition 4.2.3, the agent in principle can follow a belief revision policy such that after revision by a consistent proposition P, some of the initially less plausible P-worlds become the most plausible ones. In this case μ violates the minimality constraints of the classical AGM belief revision theory (AGM3 and AGM7) as well as principles of informational economy under consistent revision (AGM4 and AGM8) (see Bonanno 2012, Section 3 for a similar comparison). We leave the investigations of the conceptual underpinnings of different belief revision policies involved in imagination to future work and here adopt the AGM-like policy μ as a first pass.

$\mathcal{L}_{\mathsf{BI}}$ are evaluated not only with respect to possible worlds, but with respect to world-history pairs of the form $\langle w, h \rangle$. Thus, the *intension of* φ with respect to h in $\mathcal{M}$ is $|\varphi|^h_{\mathcal{M}} := \{w \in W : \mathcal{M}, \langle w, h \rangle \Vdash \varphi\}$ (we omit the subscript $\mathcal{M}$ and superscript h when the model and actual history are clear from the context).

DEFINITION 4.3. **$\Vdash$-Semantics for $\mathcal{L}_{\mathsf{BI}}$**
Given a model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V \rangle$ and a world-history pair $\langle w, h \rangle$ such that $h = (s_0, s_1, \ldots, s_n)$, the semantics for $\mathcal{L}_{\mathsf{BI}}$ is defined recursively as follows:

$$\begin{array}{lll}
\mathcal{M}, \langle w, h \rangle \Vdash p & \text{iff} & w \in V(p) \\
\mathcal{M}, \langle w, h \rangle \Vdash \neg\varphi & \text{iff} & \text{not } \mathcal{M}, \langle w, h \rangle \Vdash \varphi \\
\mathcal{M}, \langle w, h \rangle \Vdash \varphi \wedge \psi & \text{iff} & \mathcal{M}, \langle w, h \rangle \Vdash \varphi \text{ and } \mathcal{M}, \langle w, h \rangle \Vdash \psi \\
\mathcal{M}, \langle w, h \rangle \Vdash B_@\varphi & \text{iff} & Min_{\preceq_{s_0}}(W) \subseteq |\varphi|^h_{\mathcal{M}} \\
\mathcal{M}, \langle w, h \rangle \Vdash B\varphi & \text{iff} & Min_{\preceq_{s_n}}(W) \subseteq |\varphi|^h_{\mathcal{M}} \\
\mathcal{M}, \langle w, h \rangle \Vdash I\varphi & \text{iff} & \exists k < n(\preceq_{s_{k+1}} = \preceq^{\varphi}_{s_k} \text{ and } \mathcal{M}, \langle w, h[k+1] \rangle \Vdash B\varphi)
\end{array}$$

where $\preceq^{\varphi}_{s_k} = \preceq^{|\varphi|^h_{\mathcal{M}}}_{s_k}$ and $h[k] = (s_0, \ldots, s_k)$ is the initial segment of h of length $k+1$. For any $\Sigma \subseteq \mathcal{L}_{\mathsf{BI}}$ and $\varphi \in \mathcal{L}_{\mathsf{BI}}$, φ is said to be a *logical consequence* of Σ, denoted by $\Sigma \vDash \varphi$, if for all models $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V \rangle$ and all world-history pairs $\langle w, h \rangle$ of $\mathcal{M}$: if $\mathcal{M}, \langle w, h \rangle \Vdash \psi$ for all $\psi \in \Sigma$, then $\mathcal{M}, \langle w, h \rangle \Vdash \varphi$. For single-premise entailment, we write $\psi \vDash \varphi$ for $\{\psi\} \vDash \varphi$. As a special case, *logical validity*, $\vDash \varphi$, truth at all world-history pairs of all models, is $\emptyset \vDash \varphi$, entailment by the empty set of premises.

It is not difficult to see that the truth of Booleans in a given model is time and history *independent*, that is, their truth values depend only on the actual world.

LEMMA 4.1. For every model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V \rangle$, world-history pairs $\langle w, h_1 \rangle$ and $\langle w, h_2 \rangle$ in $\mathcal{M}$, and $\varphi \in \mathcal{L}$, we have $|\varphi|^{h_1}_{\mathcal{M}} = |\varphi|^{h_2}_{\mathcal{M}}$.

Proof. The proof is straightforward by subformula induction on φ. □

We stress the difference between the modality $B_@$ for *actual* beliefs and the modality B for *pretense* beliefs. The latter represents the *simulated* beliefs of the agent that they come to possess at the current stage of the imaginative episode, *after* the imaginative process has started. The former, on the other hand, represents the *actual* beliefs of the agent – i.e., the beliefs of the agent at the initial stage s_0. Moreover, the truth of sentences involving only the simulated belief modality do not depend on the whole history, but only on the plausibility ordering at the current stage. Similarly, the truth of a sentence involving only $B_@$ as its modality depends on the plausibility ordering only at the initial stage, not on the ones at the other stages in a given history.

Imagination, on the other hand, is dependent on both w and the whole history h. According to the proposed semantics, an agent has imagined φ if they have successfully revised their belief state with φ at some earlier stage in the history.[17] In other words, we take what an agent has imagined up to the current stage to be the cumulative collection of propositions by which they have revised their belief state at some stage before the current one. A less terse and more appropriate reading of $I\varphi$, then, is that "the agent has taken φ on board at some stage of the imaginative episode". In this sense, the imagination operator I is a backward looking modality that keeps track of the informational inputs the agent uses through an imaginative episode. Moreover, although the agent never imagines $\bot$ (see footnote 17), due to the definition of the simulated revision function μ, nothing stands in the way of taking φ on board while believing (either actually or in pretense) $\neg\varphi$, or taking φ and $\neg\varphi$ on board in one imaginative episode *but at different stages*.

4.2 Imagination, Axiomatic Properties, and Idealised Imaginers

Recall that Langland-Hassan (2016) distinguishes between imaginative stages that follow internally from their predecessors and those that are added through intervention. Our model allows us to capture this distinction very nicely. Given a history $h = (s_0, \ldots, s_n)$ and $k > 0$, we define the kth imaginative stage i_k, *the set of sentences the agent has imagined up to stage k*, as

$$i_k = \{\varphi \in \mathcal{L} : \langle w, h[k] \rangle \Vdash I\varphi\}.$$

This way we extract the imaginative stages through the actual history and define the corresponding imaginative episode $\mathcal{I} = (i_1, \ldots, i_n)$ as a sequence of sets of sentences in $\mathcal{L}$. An imaginative episode starts with an input proposition, forming the first imaginative stage i_1, and then develops into the full imaginative episode. In order to distinguish

[17] The agent is said to have *successfully* revised their beliefs by φ at some stage s in the given history if they believe φ in the next stage, after revision by φ. This corresponds to the *Success Postulate* of the AGM belief revision theory (Alchourrón et al. 1985) and, as B ranges only over Booleans, our framework is not subject to problems concerning higher-order beliefs such as the Moorean phenomena (cf. Holliday & Icard 2010). Due to the second conjunct in the semantic clause of $I\varphi$ in Definition 4.3 (that is, $\mathcal{M}, \langle w, h[k+1] \rangle \Vdash B\varphi$), our imagination operator is always concerned with the so-called successful revisions (for the sake of brevity, we usually drop the phrase "successful"). In fact, as stated above, the simulated revision function μ by definition always leads to successful revisions as long as the intension of the new informational input is nonempty. Since $Min_{\preceq_s}(W) \neq \emptyset$ for all s in every model, both $\neg B_{@}\bot$ and $\neg B\bot$ are validities with respect to the proposed semantics. This means that the agent never believes (actually or in pretense) nor imagines blatant contradictions (where the latter is guaranteed by the above mentioned component in the semantic clause of $I\varphi$).

between stages that follow through internal development and stages that are added through intervention, we introduce two distinct operators into our language: $I_i\varphi$ and $I_a\varphi$. The former concerns *i*nternally developed stages and the latter concerns *a*dded content through intervention. Given a model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s\in S}, \mu, V\rangle$ and a world-history pair $\langle w, h\rangle$ with $h = (s_0, \ldots, s_n)$, we interpret these two modalities as follows:

$$\begin{array}{lll} \langle w,h\rangle \Vdash I_i\varphi & \text{iff} & \exists k < n((\preceq_{s_{k+1}} = \preceq^{\varphi}_{s_k} \text{ and } \langle w, h[k+1]\rangle \Vdash B\varphi) \text{ and} \\ & & (\text{if } k = 0, \text{then } \langle w, h[k+1]\rangle \Vdash B_{@}\varphi; \text{otherwise } \langle w, h[k]\rangle \Vdash B\varphi))^{18} \\ \langle w,h\rangle \Vdash I_a\varphi & \text{iff} & \exists k < n((\preceq_{s_{k+1}} = \preceq^{\varphi}_{s_k} \text{ and } \langle w, h[k+1]\rangle \Vdash B\varphi) \text{ and} \\ & & (\text{if } k = 0, \text{then } \langle w, h[k+1]\rangle \nVdash B_{@}\varphi; \text{otherwise } \langle w, h[k]\rangle \nVdash B\varphi)) \end{array}$$

Semantically, $I_i\varphi$ states that 'the agent takes φ on board at some stage of the actual history *where they already believe it*'.[19] The proposition expressed by φ is in this sense part of the internal development. On the other hand, $I_a\varphi$ says that 'the agent takes φ on board at some stage of the actual history and φ was not believed at that stage'. This implies that φ was imagined not as a result of the agent's simulated revision process, but added 'externally' to the imaginative episode. The proposition expressed by φ is in this sense *added* content through intervention.[20]

4.2.1 Properties of $B_@$, B, and I

A couple of words on the axiomatic properties of our modal operators $B_@$, B, and I are in order. First of all, notice that both $B_@$ and B are interpreted in terms of *truth in most plausible worlds*. The only difference between the semantic clauses of these operators consists in the *stage* with respect to which they are evaluated. Therefore, it is not surprising that they satisfy the same axiomatic properties. For this reason,

[18] It is easy to see that the second conjunct in the semantic clause of $I_i\varphi$ is redundant: $\langle w, h[k]\rangle \Vdash B\varphi$ (or $\langle w, h[k+1]\rangle \Vdash B_@\varphi$, if $k = 0$) guarantees that $|\varphi|_{\mathcal{M}} \neq \emptyset$, thus, $\preceq_{s_{k+1}} = \preceq^{\varphi}_{s_k}$ implies that $\langle w, h[k+1]\rangle \Vdash B\varphi$ since μ leads to successful revision by φ as long as $|\varphi|_{\mathcal{M}} \neq \emptyset$.

[19] If the stage after which φ is taken on board is the initial stage (i.e., $k = 0$), "believe" in the reading of I_i refers to the agent's actual beliefs. Otherwise, it is the agent's simulated beliefs at stage k. The same applies to I_a.

[20] Note that, in theory, an agent might 'intervene' content that they already believe in the pretense. Such 'interventions' are not captured by our semantic clause of I_a and our model would label such a transition between two stages as internally developed. This might seem like a flaw in the definition, yet we would argue that this is in fact as it should be. The interventions that make pretense imagination have CHO – the fact that an agent can make choices in pretense imagination – as a characteristic feature are *not* 'interventions' with something the agent already believes (in the pretense). These latter instances of 'intervening' are neither philosophically interesting nor the kind of interventions that authors discussing CHO seem to have in mind (e.g., Nichols & Stich 2003).

we state the principles of interest only in terms of B and note that everything we say about B also holds for $B_@$. To this end, in this section, whenever we say "the agent believes" we refer to the agent's simulated beliefs.

As standard for the belief modality interpreted as *truth in the most plausible worlds*, our simulated belief modality B is a normal modal operator that satisfies the axiom of consistency D, formulated as $B\varphi \to \neg B \neg \varphi$, or, equivalently, as $\neg B \bot$ (see Blackburn et al. 2001 for a general introduction to basic modal logic and see, e.g., van Benthem 2007 for a logic of belief on plausibility models). This, in particular, means that our agent never has inconsistent beliefs at a stage, believes all logical truths, and their beliefs are closed under believed implications:

Consistency of *B*: $\models B\varphi \to \neg B \neg \varphi$

Omniscience rule for *B*: if $\models \varphi$, then $\models B\varphi$

Closure under believed implications: $\models B(\varphi \to \psi) \to (B\varphi \to B\psi)$

As a consequence of the above principles, we also have that the agent believes all logical consequences of what they believe and their beliefs are closed under logical equivalences:

Closure under valid implications for *B*: if $\models \varphi \to \psi$, then $\models B\varphi \to B\psi$

Closure under valid equivalences for *B*: if $\models \varphi \leftrightarrow \psi$, then $\models B\varphi \leftrightarrow B\psi$

The agent in question is therefore highly idealised, in the sense that they are logically omniscient with respect to their beliefs.

The imagination operator I is weaker, namely, a classical modal operator (see, e.g., Chellas 1980 and Pacuit 2017 for classical modal logics). The agent does not necessarily imagine all logical truths, their imagination is not closed under imagined implications and they do not necessarily imagine all logical consequences of what they imagine, i.e., the following *fail*:

Omniscience rule for *I*: if $\models \varphi$, then $\models I\varphi$

Closure under imagined implications: $I(\varphi \to \psi) \to (I\varphi \to I\psi)$

Closure under valid implications for *I*: if $\models \varphi \to \psi$, then $\models I\varphi \to I\psi$

Counterexample 1: Consider the model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V \rangle$ in Figure 2, where $W = \{w_1, w_2, w_3\}$ such that $V(q) = \{w_1\}$ and $V(p) = \{w_2\}$ and where μ is the

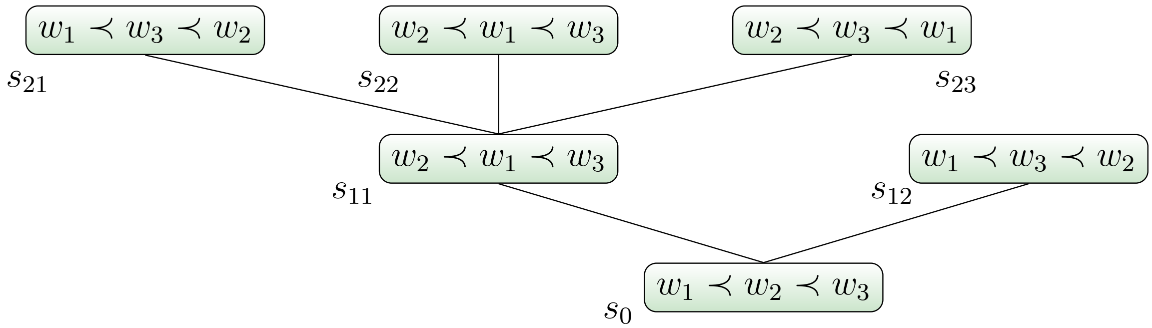


Figure 2: Counterexample 1; the plausibility ordering of each stage is given in the corresponding box.

so-called lexicographic upgrade operator defined on W.[21] The rest of the model is as depicted in Figure 2. For the sake of this argument, it is sufficient to focus on the branches that include stage s_{11}. For omniscience rule for I: $p \vee \neg p$ is a logical validity, but $\langle w_1, (s_0, s_{11})\rangle \nVdash I(p \vee \neg p)$ since $\preceq_{s_{11}} \neq \preceq_{s_0}^{(p\vee\neg p)} = \preceq_{s_0}$. Moreover, for closure under valid implications, we have $p \to (p \vee \neg p)$ logically valid and $\langle w_1, (s_0, s_{11})\rangle \Vdash Ip$ since $\preceq_{s_{11}} = \preceq_{s_0}^{p}$ and $\langle w_1, (s_0, s_{11})\rangle \Vdash Bp$. However, $\langle w_1, (s_0, s_{11})\rangle \nVdash I(p \vee \neg p)$ as shown above. As a counterexample for closure under imagined implications, consider the world-history pair $\langle w_1, (s_0, s_{11}, s_{21})\rangle$: we have $\langle w_1, (s_0, s_{11}, s_{21})\rangle \Vdash I(p \to q)$ since $\preceq_{s_{21}} = \preceq_{s_{11}}^{(p\to q)}$ as $|p \to q| = \{w_1, w_3\}$, and $\langle w_1, (s_0, s_{11}, s_{21})\rangle \Vdash B(p \to q)$. Moreover, $\langle w_1, (s_0, s_{11}, s_{21})\rangle \Vdash Ip$ since $\preceq_{s_{11}} = \preceq_{s_0}^{p}$ and $\langle w_1, (s_0, s_{11})\rangle \Vdash Bp$. However, $\langle w_1, (s_0, s_{11}, s_{21})\rangle \nVdash Iq$ since $\preceq_{s_{11}} \neq \preceq_{s_0}^{q}$ and $\preceq_{s_{21}} \neq \preceq_{s_{11}}^{q}$ – i.e., the sequence (s_0, s_{11}, s_{21}) cannot be obtained via an upgrade by q.

However, if φ and ψ are logically or necessarily equivalent, imagining one automatically leads to imagining the other. In other words, the following principle does hold:

Closure under valid equivalences for I: if $\vDash \varphi \leftrightarrow \psi$, then $\vDash I\varphi \leftrightarrow I\psi$

This is because the simulated revision policy characterised by μ cannot distinguish logically or necessarily equivalent propositions: $\preceq_s^{\varphi} = \preceq_s^{\psi}$ if $\vDash \varphi \leftrightarrow \psi$. Therefore, although weaker than belief, the operator I still renders our agents unrealistically idealised with respect to their imagination. For example, according to the proposed semantics, if the agent imagines at a stage that Jane is a logician or she is not, they also imagine that

[21] The lexicographic upgrade of a preorder $\preceq \subseteq W \times W$ by a subset $P \subseteq W$ makes all P-worlds strictly more plausible than all $W \backslash P$-worlds and keeps the ordering the same within those two zones. Our simulated belief revision function μ is the lexicographic upgrade when $P = Q$ in Definition 4.2.3. Even though the lexicographic upgrade does not play an essential role in our conceptual arguments, for the sake of simplicity, we take μ to be a lexicographic upgrade operator in all our (counter)examples.

$2+2=4$. Intuitively, we can imagine or believe the former without imagining or believing the latter and *vice versa*. In addition, while the former might be on-topic with an imaginative episode about Jane, the latter is not necessarily so. Consider again the tea-party example from Section 2. The agent does imagine that one of the cups is full, however, they do not imagine that one of the cups is full *and 2 + 2 = 4*, even though these two sentences are logically equivalent. In a similar vein, they do not import their beliefs about Paris being the capital of France to the imaginative episode as this might be completely off-topic. The model of Section 4.1 is unable to account for such cases.

5 What's it all About: Adding Topicality

This section aims at refining the formal models of Section 4 in a way that the modal operators $B_@$, B and I, as well as the simulated revision function μ, become more sensitive to distinctions between logically equivalent contents. To do so, we endow branching-time belief revision models with (an enriched version of) topic models introduced in Berto (2018). This way we can evade the problems concerning the aforementioned idealisations.

Definition 5.1. **Topic Model for $\mathcal{L}$**
A *topic model* $\mathcal{T}$ is a tuple $\langle T, \oplus, t\rangle$, where

1. T is a nonempty set of *possible topics*;

2. $\oplus : T \times T \to T$ is a binary idempotent, commutative, associative operation: *topic fusion*. We assume unrestricted and complete fusion, that is, for all $T' \subseteq T$, $\bigoplus T' \in T$.

3. t : $\mathsf{Prop} \to T$ is a *topic function* assigning a topic to each element in Prop. t extends to the whole $\mathcal{L}$ by taking the topic of a sentence φ as the fusion of the topics of the atomic propositions occuring in it. I.e.,

$$t(\varphi) = \bigoplus \mathsf{AT}(\varphi) = t(p_1) \oplus \cdots \oplus t(p_n),$$

where $\mathsf{AT}(\varphi) = \{p_1, \ldots, p_n\}$ is the set of propositional variables occuring in φ.

In the metalanguage we use variables $\mathbf{a}, \mathbf{b}, \mathbf{c}$ $(\mathbf{a}_1, \mathbf{a}_2, \ldots)$ ranging over possible topics. We define *topic parthood*, denoted by $\sqsubseteq$, in a standard way as

$$\forall \mathbf{a}, \mathbf{b}(\mathbf{a} \sqsubseteq \mathbf{b} \text{ iff } \mathbf{a} \oplus \mathbf{b} = \mathbf{b}).$$

Thus, $(T, \oplus)$ is a *complete join semilattice* and $(T, \sqsubseteq)$ a *poset*. The *strict topic parthood*, denoted by $\sqsubset$, is defined as usual: $\mathbf{a} \sqsubset \mathbf{b}$ iff $\mathbf{a} \sqsubseteq \mathbf{b}$ and $\mathbf{b} \not\sqsubseteq \mathbf{a}$. Another important operator is the so-called *topic intersection* $\sqcap : T \times T \to T$ such that $\mathbf{a} \sqcap \mathbf{b} = \bigoplus\{\mathbf{c} \in T : \mathbf{c} \sqsubseteq \mathbf{a} \text{ and } \mathbf{c} \sqsubseteq \mathbf{b}\}$. In words, $\mathbf{a} \sqcap \mathbf{b}$ is the fusion of all topics that are a common part of both $\mathbf{a}$ and $\mathbf{b}$. Finally, topics of complex sentences φ are defined from their primitive components in $\mathsf{AT}(\varphi)$, where all the logical connectives, as argued in Section 2.2, are topic-transparent. We therefore have that for all $\varphi, \psi \in \mathcal{L}$, $t(\neg\varphi) = t(\varphi)$ and $t(\varphi \wedge \psi) = t(\varphi) \oplus t(\psi)$.[22]

Topic models provide an abstract and objective (i.e., agent independent) representation of the mereological structure of topics assigned to Boolean sentences and, in turn, help us make distinctions between logically equivalent contents (Berto 2018, 2021). However, as we argued in Section 2, Berto's theory is too coarse-grained in that it cannot account for the possibility that exactly the same explicit input can lead to different imaginative episodes due to their distinct overall topics (recall the example about the tornado in Indonesia from Section 2.2). The reason why Berto's account is unable to deal with this issue is, we suggest, that his topic models include neither a representation of the overall topic of the imaginative episode nor the totality of topics the agent has mastered already (though the latter has been employed in recent work, cf. Hawke et al. 2020; Özgün & Berto 2020). We add these two components to our models in order to overcome the aforementioned issues.

We can now define a *topic-sensitive* version of branching-time belief revision models:

Definition 5.2. **Topic-sensitive model**
A *topic-sensitive model* is a tuple $\langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, T, \oplus, t, \mathfrak{b}, \mathfrak{i}, V\rangle$

1. $\langle T, \oplus, t\rangle$ is a topic model;

2. $\mathfrak{b}$ and $\mathfrak{i}$ are designated elements of T such that $\mathfrak{b}$ represents 'the totality of topics the agent has grasped' and $\mathfrak{i}$ represents 'the overall topic of the imaginative episode'.

3. $\langle S, \rightarrowtail, W, \{\preceq_s\}_{s \in S}, \mu, V\rangle$ is a model such that for all $s, s' \in S$, if $s \rightarrowtail s'$ then $\preceq_{s'} = \preceq_s^{\varphi}$ for some $\varphi \in \mathcal{L}$ with $t(\varphi) \sqsubseteq \mathfrak{b} \sqcap \mathfrak{i}$.

A topic-sensitive model is equipped with an overall topic of the imagination exercise, $\mathfrak{i}$, and the totality of the topics the agent has already grasped, that is, $\mathfrak{b}$. These two components together impose a topicality filter on what the agent believes (actually or

[22] Note that this straightforwardly generalises to other two-place connectives.

in pretense) and imagines, thus, resolving the issues regarding idealisations noted at the end of the previous section.

Component $\mathfrak{b}$ makes sure that the agent cannot believe those propositions whose topic they have not grasped. Intuitively, you do not believe that extremally disconnectedness is a hereditary property if you have never heard of, e.g., the topological properties 'extremally disconnecteness' or 'being hereditary'. Believing a proposition requires having a grasp of its topic. The designated element $\mathfrak{b}$ allows us to account for this. Secondly, as mentioned above, one does not imagine *everything* they believe. Some of our beliefs might be off-topic with regard to the given imaginative episode and a purposeful imaginative exercise requires keeping the imaginings within the subject matter of this imaginative episode. The component $\mathfrak{i}$ – i.e., the overall topic of the imaginative episode – helps us to capture this formally. So, $\mathfrak{b}$ and $\mathfrak{i}$ together make it that pretense imagination is topic-restricted in the following way: the topic of what the agent imagines is a common part of both the totality of the topics the agent grasps *and* the overall topic of the imaginative episode. This is formalised by using the topic intersection operator $\sqcap$. Finally, the constraint in Definition 5.2.3 states that, in pretense, an agent revises their beliefs only according to the revision policy defined by μ and *only with those propositions whose topics they have mastered and that fall under the overall topic of the imaginative episode.*

These features will be better understood when we present the new, topic-sensitive semantics for $\mathcal{L}_{\mathsf{BI}}$. While the semantics of the Booleans remain as they were before, the semantics of $B_{@}\varphi$, $B\varphi$, and $I\varphi$ are made stronger in the appropriate way with topicality constraints.

Definition 5.3. **$\Vdash$-Semantics for $\mathcal{L}_{\mathsf{BI}}$**
Given a topic-sensitive model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s\in S}, \mu, T, \oplus, t, \mathfrak{b}, \mathfrak{i}, V\rangle$ and world-history pair $\langle w, h\rangle$ such that $h = (s_0, s_1, \ldots, s_n)$, the semantics for $\mathcal{L}_{\mathsf{BI}}$ is as given in Definition 4.3 for the components in $\mathcal{L}$, plus:

$$\begin{array}{lll} \mathcal{M}, \langle w, h\rangle \Vdash B_{@}\varphi & \text{iff} & Min_{\preceq_{s_0}}(W) \subseteq |\varphi|_{\mathcal{M}} \text{ and } t(\varphi) \sqsubseteq \mathfrak{b} \\ \mathcal{M}, \langle w, h\rangle \Vdash B\varphi & \text{iff} & Min_{\preceq_{s_n}}(W) \subseteq |\varphi|^{h}_{\mathcal{M}} \text{ and } t(\varphi) \sqsubseteq \mathfrak{b} \\ \mathcal{M}, \langle w, h\rangle \Vdash I\varphi & \text{iff} & \exists k < n(\preceq_{s_{k+1}} = \preceq^{\varphi}_{s_k} \text{ and } \langle w, h[k+1]\rangle \Vdash B\varphi) \text{ and } t(\varphi) \sqsubseteq \mathfrak{b} \sqcap \mathfrak{i} \end{array}$$

According to the topic-sensitive semantics, the agent believes φ at stage s iff (1) φ is true at all the most plausible worlds at s and (2) the agent has already grasped the topic of φ, i.e., the topic of φ is included in $\mathfrak{b}$ (Özgün & Berto, 2020). Therefore, the agent cannot believe φ (actually or within a pretense) if they have not grasped its topic. Imagination, on the other hand, is restricted, additionally, by the overall topic of the imaginative exercise. The agent has imagined φ if they have revised their belief state with φ at some earlier stage in the history and *the topic of φ is included* *in the intersection of the overall topic of the imaginative episode and the topic of the agent's belief state.* In topic-sensitive models with a singleton T, the semantics given in Definitions 4.3 and 5.3 coincide.[23]

Let us now see to what extent the topic-sensitivity solves the aforementioned problems concerning idealisation and overall topic of an imaginative episode.

5.1 Idealisations, Tea-Parties, and Tornadoes

As in Section 4.2.1, we focus on the operator B and note that the same properties also hold for $B_@$. Topic-sensitivity allows us to model agents who do not believe all logical truths and whose beliefs are not closed under logical implications. That is, topic-sensitive models *invalidate* the following principles:

Omniscience rule for B: if $\vDash \varphi$, then $\vDash B\varphi$

Closure under valid implications for B: if $\vDash \varphi \to \psi$, then $\vDash B\varphi \to B\psi$

Moreover, our agents can imagine/believe φ without imagining/believing ψ even when they are logically or necessarily equivalent. That is the following principles no longer hold in topic-sensitive models.

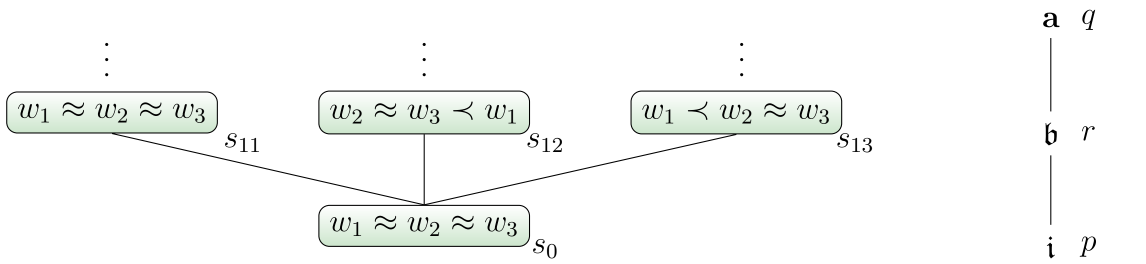
Closure under valid equivalences for B: if $\vDash \varphi \leftrightarrow \psi$, then $\vDash B\varphi \leftrightarrow B\psi$

Closure under valid equivalences for I: if $\vDash \varphi \leftrightarrow \psi$, then $\vDash I\varphi \leftrightarrow I\psi$

Counterexample 2: Consider the topic sensitive model $\mathcal{M} = \langle S, \rightarrowtail, W, \{\preceq_s\}_{s\in S}, \mu, T, \oplus, t, \mathfrak{b}, \mathfrak{i}, V\rangle$ in Figure 3, where $(S, \rightarrowtail)$ and $\preceq_s$ are as given in Figure 3(a), $W = \{w_1, w_2, w_3\}$, $T = \{\mathbf{a}, \mathfrak{b}, \mathfrak{i}\}$ with the topic lattice as depicted in Figure 3(b), and μ as a lexicographic upgrade operator. Finally, we consider three propositions p, q, r such that $V(p) = \{w_1\}$, $V(q) = \{w_1, w_2\}$, $V(r) = W$, and $t(p) = \mathfrak{i}$, $t(r) = \mathfrak{b}$, and $t(q) = \mathbf{a}$. It is easy to see that $\mathcal{M}$ is a topic-sensitive model. In particular, the plausibility ordering at each stage can be obtained via a μ-revision from the one in the previous stage by some $\varphi \in \mathcal{L}$ such that $t(\varphi) \sqsubseteq \mathfrak{b} \sqcap \mathfrak{i}$.

To refute closure under valid equivalences for I, let the actual history be $h = (s_0, s_{13})$. We then have $\langle w_1, h\rangle \Vdash Ip$, since $\preceq_{s_{13}} = \preceq^p_{s_0}$ and $t(p) = \mathfrak{i} \sqsubseteq \mathfrak{b} \sqcap \mathfrak{i} = \mathfrak{i}$. However, note that $\langle w_1, h\rangle \nVdash I(p \wedge (r \vee \neg r))$, since $t(p \wedge (r \vee \neg r)) = t(p) \oplus t(r) = \mathfrak{b}$ and $\mathfrak{b} \not\sqsubseteq \mathfrak{b} \sqcap \mathfrak{i} = \mathfrak{i}$. So, even though p and $p \wedge (r \vee \neg r)$ are *logically* equivalent, the agent can imagine the former without imagining the latter as r is *off-topic* with respect to the

[23] The definitions of internally developed imaginative stages and intervened imaginative stages can be made topic-sensitive in a similar manner.



(a) $\langle S, \rightarrowtail, W, \{\preceq_s\}_{s \,\in\, S}, \mu\rangle$ (b) $\langle T, \oplus, t, \mathfrak{b}, \mathfrak{i}\rangle$

Figure 3: Counterexample 2; the plausibility ordering of each stage is given in the corresponding box in Fig. 3(a). Topic assignment is given by labelling the nodes in Fig. 3(b) with atomic formulae.

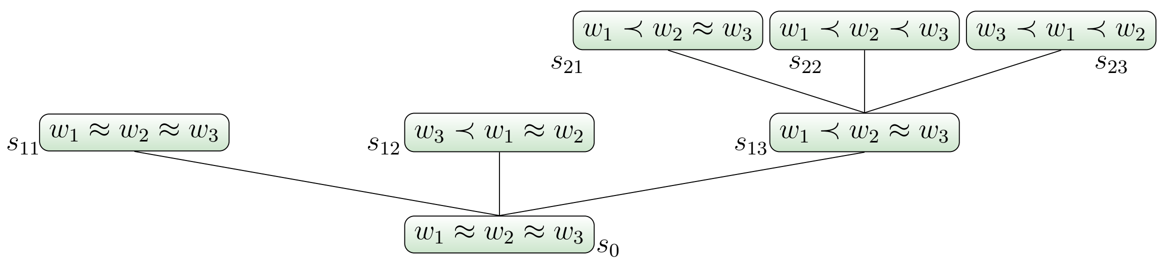
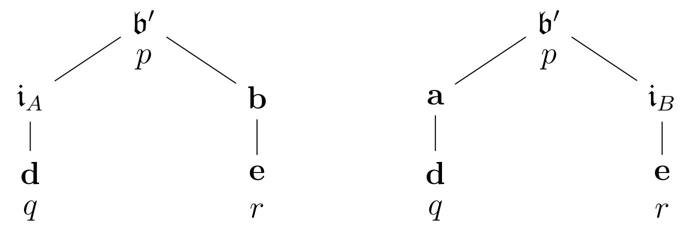


Figure 4: Structure $\langle S', \rightarrowtail', W', \{\preceq_s\}_{s \,\in\, S}, \mu'\rangle$. The plausibility ordering of each stage is given in the corresponding box.

overall topic of the imaginative episode. This is exactly what we would expect. As a counterexample for the omniscience rule for B, take $\varphi := q \vee \neg q$, and for closure under valid implications and equivalences for B, consider $\varphi := p$ and $\psi := p \wedge (q \vee \neg q)$. (As a counterexample for closure under valid implications and equivalences for $B_{@}$, consider $\varphi := r$ and $\psi := r \wedge (q \vee \neg q)$).

Tea-Parties and the Capital of France

Let us now stipulate that $r :=$ 'Paris is the capital of France'. In the model $\mathcal{M}$ given in Figure 3 and every world-history pair $\langle w, h\rangle$ of $\mathcal{M}$, we have that $\langle w, h\rangle \Vdash Br$ (since $t(r) = \mathfrak{b}$ and $|r|^h_{\mathcal{M}} = W$). However, $\langle w, h\rangle \nVdash Ir$ since $t(r) = \mathfrak{b} \not\sqsubseteq \mathfrak{b} \sqcap \mathfrak{i} = \mathfrak{i}$, i.e., r is not on-topic with the modelled imaginative episode. This shows that one can, for example, imagine a tea-party, without taking on board everything one believes.



(a) Topic lattice for A (b) Topic lattice for B

Figure 5: Topic components for Contexts A & B. Topic assignment is given by labelling the nodes with atomic formulae.

Tornadoes in Indonesia

As a last example, we return to the case of the tornadoes in Indonesia presented in Section 2.2. Consider the models $\mathcal{M}_A = \langle S', \rightarrowtail', W', \{\preceq_s\}_{s\in S}, \mu', T', \oplus', t', \mathfrak{b}', \mathfrak{i}_A, V'\rangle$ and $\mathcal{M}_B = \langle S', \rightarrowtail', W', \{\preceq_s\}_{s\in S}, \mu', T', \oplus', t', \mathfrak{b}', \mathfrak{i}_B, V'\rangle$, where $\langle S', \rightarrowtail', W', \{\preceq_s\}_{s\in S}, \mu'\rangle$ is as in Figure 4 (μ' is a lexicographic upgrade operator), with $V'(p) = \{w_1\}$ and $V'(q) = V'(r) = \{w_1, w_2\}$. The topic components $\langle T', \oplus', t', \mathfrak{b}', \mathfrak{i}_A\rangle$ and $\langle T', \oplus', t', \mathfrak{b}', \mathfrak{i}_B\rangle$ are as given in Figures 5(a) and 5(b), respectively. $\mathcal{M}_A$ and $\mathcal{M}_B$ are intended to model two distinct imaginative episodes of the same agent, where the distinction is solely due to the difference between the overall topics of the corresponding episodes. Thus, the only difference between the two models is the designated overall topics: $\mathfrak{i}_A$ and $\mathfrak{i}_B$. Now, let $p :=$ 'There is a tornado in Indonesia' be the input proposition, $q :=$ 'Booking a flight through the US rather than Indonesia is safer,' and $r :=$ 'My friend is in danger'. Then, $\mathcal{M}_A$ and $\mathcal{M}_B$ can be seen as models of **Context A** and **Context B** from p. 6, respectively. Suppose further that $\langle w_1, (s_0, s_{13}, s_{22})\rangle$ is the actual world-history pair. We then have that $\mathcal{M}_A, \langle w_1, (s_0, s_{13}, s_{22})\rangle \Vdash Bq \wedge Br$ and $\mathcal{M}_B, \langle w_1, (s_0, s_{13}, s_{22})\rangle \Vdash Bq \wedge Br$. However, $\mathcal{M}_A, \langle w_1, (s_0, s_{13}, s_{22})\rangle \Vdash Iq$ (since $\preceq_{22} = \preceq^q_{13}$, $\mathcal{M}_A, \langle w_1, (s_0, s_{13}, s_{22})\rangle \Vdash Bq$, and $t'(q) \sqsubset \mathfrak{b}' \sqcap \mathfrak{i}_A$), but $\mathcal{M}_B, \langle w_1, (s_0, s_{13}, s_{22})\rangle \nVdash Iq$ (since $t'(q) \not\sqsubset \mathfrak{b}' \sqcap \mathfrak{i}_B$). Similarly, we also have $\mathcal{M}_A, \langle w_1, (s_0, s_{13}, s_{22})\rangle \nVdash Ir$ and $\mathcal{M}_B, \langle w_1, (s_0, s_{13}, s_{22})\rangle \Vdash Ir$.

6 Conclusion: Logic of Pretense Imagination

We have developed a new formal model of pretense imagination. We have done so by using tools from dynamic epistemic logic, belief revision theory, as well as more recently introduced topic models. All these components together help us deal with issues concerning idealisations, irrelevant background beliefs, and the context-sensitivity of pretense imagination, as shown in Section 5.1. In this conclusion, we first explain how our topic-sensitive models can account for the central features of pretense imagination listed in Section 3. We then compare our formalism to three other recent formal approaches to imagination (Section 6.1). Finally, we reflect on the theoretical underpinnings of the notion of imagination that we have modelled and examine how it relates to other, philosophical, notions of imagination (Section 6.2).

We have focused exclusively on propositional imagination, so **PI** requires no comments. For **ESP**, recall that the imaginative episode $\mathcal{I} = (i_1, \ldots, i_n)$ is obtained from an actual history, where i_1 constitutes the explicit starting point of the imaginative episode. Relatedly, we have specified the actual beliefs of the agent outside of the pretense in such a way that it should be clear that the actual beliefs of the agent are kept fixed and only the pretend beliefs are revised. This gives us **QU**. **CHO** is accounted for as our models are rich enough to distinguish the operators I_i and I_a, where the latter is concerned with added content through intervention. **RAT** also follows straightforwardly from the fact that we construct the imaginative episodes from imaginative stages that are the result of simulated belief revisions, reliant on the same kind of revision mechanisms as actual belief revision (though, as mentioned in Section 4.1, our model allows for a variety of revision policies to be implemented). **ROI** is accounted for, on the one hand, because of the belief revision policy and, on the other hand, because of the topicality filter. The former makes it so that, given a particular input, focusing on the most plausible worlds holds fixed many (known) constitutive facts and causal laws (cf. Williamson 2007). The latter makes sure that the agent doesn't imagine random, off-topic things in an imaginative episode because imagination is restricted in important ways by the overall topic of the imaginative episode and the totality of topics the agent has grasped.

6.1 Other Logics of Imagination

We discuss four recent formal approaches to imagination – namely that of Berto (2018, 2021), Wansing (2017), Canavotto et al. (2020), and Casas-Roma et al. (2019) – in relation to our own models.

First of all, Berto (2018, 2021) presents a logic of imagination that is closely related to our own, based on variably strict modal operators and topicality models. On his account, imagination is explicitly *conditional*: one imagines φ *given that they imagine some* ψ. As we have argued throughout, Berto's logic is too coarse-grained to account for the differences between imaginative episodes that are due to the difference in their overall topics (see the discussion from Section 2.2). Our models can account for these differences, as shown in the previous section.

Secondly, Wansing's (2017) logic of imagination is rather static and aims to capture the *agentive* aspect of imagination which might be thought to be somewhat under-represented in our models.[24] To this end, he uses mechanisms from STIT ("seeing-to-it-that") logics – that, in this context, aim to represent an agent's direct voluntary control over their beliefs and imagination – combined with neighbourhood models. Unlike ours, Wansing's (2017) logic formalises only imagination, rather than belief and imagination together, and admits closure of imagination under valid equivalence.

Canavotto et al. (2020) combine Berto's and Wansing's work and take it one step further. Unlike Berto, they capture the agentive facet of imagination and distinguish its voluntary and involuntary aspects. Unlike Wansing, they use possible worlds semantics and conditional imagination operators to formalise imagination while also taking its topicality and relevance constraints into account. Thus, following Berto, they model what an agent imagines *given* a particular input. Our focus, on the other hand, has been on pretense imagination and, in particular, on *how* an agent develops an imaginative episode by following a certain revision policy, given their background beliefs and the overall topic of that imaginative episode. That is, our model focuses on the development of pretense imagination over time.

Finally, Casas-Roma et al. (2019) develop a formal model for imagination based also on the work of Nichols & Stich (2003) and Langland-Hassan (2016) using possible worlds semantics. They model the *act of imagination* as a dynamic process based on a relational structure, where the dynamics of imagination is formalised following their Imagination Algorithm that *creates new imaginary possible worlds* given an initial premise characterising the initial imaginary scenario. As Casas-Roma et al. work with complete possible worlds, they validate a number of principles that result in highly idealised agents. For example, their semantics validates the omniscience rules for I; closure under imagined implications; closure under valid implications for I; and closure under valid equivalences for I. Moreover, once you imagine something – i.e., there is a world that makes their $\langle \mathrm{Img} \rangle \varphi$ true for any φ – then, because they rely on complete possible worlds, it is so that for each ψ, you either imagine it or its negation. As we saw above, this is not the case in our models.

6.2 Philosophy of (Pretense) Imagination

In our discussion of pretense imagination, we've bracketed two issues: one on the relation between RAT and ROI and one on the possibility of pretense imagination being relevant to other cognitive endeavours. We briefly discuss these in turn.

As we noted above (footnote 13), when we talk about 'mundane' instances of pretense imagination (e.g., concerning pretend tea-parties), the distinction between the rationality constraint and the reality-orientedness of pretense imagination is relatively

[24]See also Olkhovikov & Wansing (2018, 2019).

straightforward. The former concerns the claim that agents deal with incoming information within the pretense in a way very similar to the way that they would if the relevant information were actually received (cf. Nichols & Stich 2003; Langland-Hassan 2012). The reality-orientedness, on the other hand, concerns the development of the situation in line with the (known) constitutive facts and causal laws (cf. Williamson 2007; Kind 2016). An interesting question is how RAT and ROI interact when the pretense imagination concerns more 'exotic' situations, for example, when one imagines being Luke Skywalker.[25] For example, what, if any, constitutive facts and causal laws does one hold fixed? Does it make sense to talk of such imaginings as being 'reality-oriented'?

It seems to us that there still is a relevant sense in which one could take these imaginings to be reality-oriented. Consider for example an imagining where one meets Luke Skywalker, who starts training them in using a lightsaber. Such imagining would be 'more reality-oriented' than an imagining where one meets Luke Skywalker, who then turns into a unicorn that is the king of France. Such examples suggest that even in more exotic cases of pretense imagination it makes sense to talk about the reality-orientedness of an imagining and to separate this from the RAT constraint. Of course, much remains to be said about such case (for example, what, if any, constitutive facts and causal laws does one hold fixed in such cases? What is the relation between these cases and the comments about unreliability when it comes to exotic imaginings made by Williamson (2007, p. 164)). We leave this for future work.

Secondly, there is the relation between pretense imagination as discussed here and other kinds of imagination discussed in the literature. We take the kind of imagination that plays a role in pretense and pretend play as a starting point, but there is nothing that suggests that this kind of imagination does not *also* feature in other cognitive activities. In fact, we think that the kind of imagination that we have dubbed 'pretense imagination' is quite ubiquitous. For example, it seems that pretense imagination is very similar to the kind of *reality-oriented mental simulations* that Berto (2021) discusses. Berto points out that this kind of imagination is constrained by topicality and minimal alteration (p. 2031), which is exactly what we suggest for pretense imagination. Similarly, the kind of imagination that is used to evaluate particular conditionals is very similar to the way we have described pretense imagination (cf. Byrne 2005; Gopnik & Walker 2013; Langland-Hassan 2016; Williamson 2020; Schoonen 2021). Finally, many authors who explain the epistemic usefulness of imagination by appeal to its recreative nature seem to have in mind a kind of imagination of which pretense imagination constitutes a part (cf. Kind 2016; Kind & Kung 2016; Williamson 2016; Berto 2021;

[25] Thanks to an anonymous reviewer for pushing us to think about this and for this particular example.

Schoonen 2021).[26] If these considerations are correct, then our discussion of pretense imagination, and, *a fortiori*, our model thereof, applies to many more instances of imagination than merely those relevant to pretense and pretend play.

Acknowledgements: Much of this work was conducted when the authors were employed by the ILLC, University of Amsterdam and affiliated with Arché Research Centre, University of St. Andrews through the project 'The Logic of Conceivability'. We thank our reviewers for exceptionally stimulating and helpful feedback. A version of this paper was presented at the Fiction and Imagination Workshop at the University of Turin in June 2019. Thanks to the audience for their valuable feedback. Special thanks to Francesco Berto, for his detailed comments on an earlier draft which allowed us greatly improve the paper. Also thanks to Ilaria Canavotto, who read a version of (parts of) this paper. Finally, thanks for the input of the Logic of Conceivability team: Francesco Berto, Peter Hawke, Karolina Krzyżanowska, and Anthi Solaki. This research is published within the project 'The Logic of Conceivability', funded by the European Research Council (ERC CoG), grant number 681404.

[26] For example, suggesting that imagination is a *recreative capacity* might involve the claim that imagination can simulate perceptual experiences *as well as* belief revision (see, for example, Currie & Ravenscroft 2002; Williamson 2007).

References

Alchourrón, C., Gärdenfors, P., & Makinson, D. (1985). On the Logic of Theory Change: Partial Meet Functions for Contraction and Revision. *Journal of Symbolic Logic*, *50*, 510–30.

Balcerak Jackson, M. (2016). On the Epistemic Value of Imagining, Supposing, and Conceiving. In A. Kind, & P. Kung (Eds.) *Knowledge Through Imagination*, (pp. 42–60). Oxford: Oxford University Press.

——— (2018). Justification by Imagination. In F. Macpherson, & F. Dorsch (Eds.) *Perceptual Imagination and Perceptual Memory*, chap. 10, (pp. 209–226). Oxford University Press.

Baltag, A., & Smets, S. (2006). Dynamic Belief Revision over Multi-Agent Plausibility Models. In G. Bonanno, W. van der Hoek, & M. Wooldridge (Eds.) *Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision (LOFT2006)*, (pp. 11–24). University of Liverpool.

Berto, F. (2017). Impossible Worlds and the Logic of Imagination. *Erkenntnis*, *82*(6), 1277–1297.

——— (2018). Aboutness in Imagination. *Philosophical Studies*, *175*, 1871–1886.

——— (2021). Taming the Runabout Imagination Ticket. *Synthese*, *198*, 2029–2043.

Berto, F., & Hawke, P. (2018). Knowability Relative to Information. *Mind*, *130*(517), 1–33.

Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal Logic*, vol. 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge: Cambridge University Press.

Bonanno, G. (2007). Axiomatic Characterization of the AGM Theory of Belief Revision in a Temporal Logic. *Artificial Intelligence*, *171*(2), 144–160.

——— (2012). Belief Change in Branching Time: AGM-consistency and Iterated Revision. *Journal of Philosophical Logic*, *41*(1), 201–236.

Boutilier, C. (1994). Unifying Default Reasoning and Belief Revision in a Modal Framework. *Artificial Intelligence*, *68*(1), 33–85.

——— (1996). Iterated Revision and Minimal Change of Conditional Beliefs. *Journal of Philosophical Logic*, *25*, 263–305.

Byrne, R. M. J. (2005). *The Rational Imagination*. London, England: The MIT Press.

Canavotto, I., Berto, F., & Giordani, A. (2020). Voluntary Imagination: A Fine-Grained Analysis. *Review of Symbolic Logic*. Forthcoming.

Casas-Roma, J., Huertas, A., & Rodríguez, M. E. (2019). The Logic of Imagination Acts: A Formal System for the Dynamics of Imaginary Worlds. *Erkenntnis*. Forthcoming.

Chellas, B. F. (1980). *Modal Logic: An Introduction*. Cambridge, MA.: Cambridge University Press.

Currie, G., & Ravenscroft, I. (2002). *Recreative Minds*. Oxford: Oxford University Press.

Fine, K. (1986). Analytic Implication. *Notre Dame J. Formal Logic*, *27*(2), 169–179.

——— (2016). Angellic Content. *Journal of Philosophical Logic*, *45*(2), 199–226.

Gopnik, A., & Walker, C. M. (2013). Considering Counterfactuals. The Relationship between Causal Learning and Pretend Play. *American Journal of Play*, *6*(1), 15–28.

Hawke, P. (2018). Theories of Aboutness. *Australasian Journal of Philosophy*, *96*(4), 697–723.

Hawke, P., Özgün, A., & Berto, F. (2020). The Fundamental Problem of Logical Omniscience. *Journal of Philosophical Logic*, *49*, 727–766.

Hintikka, J. (1962). *Knowledge and Belief*. Ithaca, NY.: Cornell University Press.

Holliday, W., & Icard, T. F. (2010). Moorean Phenomena in Epistemic Logic. In L. Beklemishev, V. Goranko, & V. Shehtman (Eds.) *Advances in Modal Logic*, vol. 8, (pp. 178–199). London: College Publications.

Kind, A. (2013). The Heterogeneity of the Imagination. *Erkenntnis*, *78*(1), 141–159.

——— (2016). Imagining Under Constraints. In A. Kind, & P. Kung (Eds.) *Knowledge Through Imagination*, chap. 6, (pp. 145–159). Oxford: Oxford University Press.

Kind, A., & Kung, P. (2016). Introduction: The Puzzle of Imaginative Use. In A. Kind, & P. Kung (Eds.) *Knowledge Through Imagination*, (pp. 1–37). Oxford University Press.

Langland-Hassan, P. (2012). Pretense, Imagination, and Belief: the Single Attitude Theory. *Philosophical Studies*, *159*, 155–179.

——— (2016). On Choosing What to Imagine. In A. Kind, & P. Kung (Eds.) *Knowledge Through Imaginaion*, chap. 2, (pp. 61–84). Oxford: Oxford University Press.

Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, *50*(1-3), 211–238.

Liu, F. (2008). *Changing for the Better: Preference Dynamic and Agent Diversity*. Ph.D. thesis, University of Amsterdam.

Nichols, S., & Stich, S. P. (2003). *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Oxford University Press.

Olkhovikov, G. K., & Wansing, H. (2018). An Axiomatic System and a Tableau Calculus for STIT Imagination Logic. *Journal of Philosophical Logic*, *47*(2), 259–279.

——— (2019). Simplified Tableaux for STIT Imagination Logic. *Journal of Philosophical Logic*, *48*, 981–1001.

Özgün, A., & Berto, F. (2020). Dynamic Hyperintensional Belief Revision. *The Review of Symbolic Logic*. Forthcoming.

Pacuit, E. (2017). *Neighbourhood Semantics for Modal Logic*. Dordrecht: Springer.

Ramsey, F. P. (1929). General Propositions and Causality. In R. Braithwaite (Ed.) *The Foundations of Mathematics and Other Logical Essays*, (pp. 237–255). Martino Publishing, 2013 ed.

Saint-Germier, P. (2021). Hyperintensionality in Imagination. In A. Giordani, & J. Malinowski (Eds.) *Logic in High Definition*, (pp. 77–115). Cham: Springer.

Schoonen, T. (2021). A Note on the Epistemological Value of Pretense Imagination. *Episteme*. Forthcoming.

Stalnaker, R. C. (1968). A theory of conditionals. In W. Harper, R. Stalnaker, & G. Pearce (Eds.) *Ifs: conditionals, belief, decision, chance, and time*, (pp. 41–55). Dordrecht, Holland: D. Reidel Publishing Company.

van Benthem, J. (2007). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, *17*(2), 129–155.

van Benthem, J., & Liu, F. (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, *17*(2), 157–182.

Wansing, H. (2017). Remarks on the logic of imagination. A step towards understanding doxastic control through imagination. *Synthese*, *194*(8), 2843–2861.

Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell Publishing.

——— (2016). Knowing by Imagining. In A. Kind, & P. Kung (Eds.) *Knowledge Through Imagination*, chap. 4, (pp. 113–123). Oxford: Oxford University Press.

——— (2020). *Suppose and Tell. The Semantics and Heuristics of Conditionals*. Oxford: Oxford University Press.

Yablo, S. (2014). *Aboutness*. Princeton, NJ.: Princeton University Press.