



HAL
open science

Arbitrary Radical Upgrades

Raul Fervari, Benjamin Icard

► **To cite this version:**

Raul Fervari, Benjamin Icard. Arbitrary Radical Upgrades. Proceedings of the International Workshop on Reconfigurable Transition Systems (ReactS'24, at Conference SEFM), Nov 2024, Aveiro, Portugal. hal-04733986

HAL Id: hal-04733986

<https://hal.science/hal-04733986v1>

Submitted on 13 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Arbitrary Radical Upgrades

Raul Fervari^{1,2} and Benjamin Icard³

¹ FAMAF, Universidad Nacional de Córdoba, Argentina

² Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

³ LIP6, Sorbonne University, CNRS, France

rfervari@unc.edu.ar

benjamin.icard@lip6.fr

Abstract. This paper presents and investigates ARUL, a variant of dynamic belief revision logic in which revision policies, in particular radical, or lexicographic, upgrades, can be arbitrary. We discuss the motivations of having this kind of soft arbitrary operator, concretely for refining the analysis of agentivity and modelling classical epistemic paradoxes. We introduce a sound and complete axiomatic system over models whose accessibility relation is a reflexive, transitive and locally connected pre-order, following an approach parallel to Arbitrary Public Announcement Logic (APAL) for proving completeness.

Keywords: Dynamic Logics · Plausibility · Radical Upgrades · Arbitrary Upgrades · Epistemic Paradoxes.

1 Introduction

Public Announcement Logic (PAL) [30,8,31] was tailored as an extension of Epistemic Logic [25] to reason about information change. Soon after its introduction, PAL was followed by logics for softer revision policies, such as the belief upgrades family of e.g. [16,11,12]. In the meantime, the expressivity of PAL increased substantially with the possibility to quantify over hard announcements, leading to Arbitrary Public Announcement Logic (APAL), introduced and axiomatized in [4]. The completeness argument was then fixed in [2], and improved in [5]. This paper proposes to make similar steps for dynamic belief revision theory by introducing ARUL, a logic for Arbitrary Radical Upgrades.

Arbitrary radical upgrades have been considered already in [28], not as syntactic primitive operators, but as defined notions to help analyze a Smullyan’s deception puzzle on surprise [33] (see also [27]). In [28], arbitrary radical upgrades are used meta-logically to model two versions of deception in Smullyan’s puzzle, the so-called deception by *commission* (i.e., false beliefs obtained by the existence of some radical upgrade) and deception by *omission* (i.e., false beliefs obtained by the lack of some radical upgrade). In this paper we aim to provide a formal and general treatment of these concepts, by defining and axiomatizing the logic ARUL.

The language of ARUL is based on dynamic epistemic languages divided into two parts. The static part elaborates on [34] by featuring modalities for “at least

as plausible as” situations and for an agent’s knowledge. In turn, the dynamic part of ARUL contains the standard radical upgrade modality on *specific* formulas of the language [19,16,12], but also incorporates an operator for arbitrary upgrades, authorizing quantification over *any* formulas of ARUL. The language of ARUL is interpreted over doxastic-plausibility models (see, e.g., [12,34]). Herein, the model’s plausibility relation considered is a reflexive, transitive and locally connected pre-order. To our knowledge, the extension of arbitrariness to radical upgrades proposed by ARUL is novel. Regarding technical results, ARUL connects to standard modal logic by using tools designed to axiomatize APAL [5]. Regarding expressivity, the agentive definitions and distinctions that ARUL helps delineate find applications into the philosophical analysis of various paradoxes, such as e.g. Gerbrandy’s formalization of the surprise examination paradox and variants, such as Fitch’s paradox and Moore’s paradox.

Outline of the paper. In Section 2, we introduce the dynamic logic of arbitrary radical upgrades (ARUL), describing the syntax and semantic interpretation of the language. Section 3 presents an axiom system for ARUL, akin to the axiomatization of APAL, that we prove to be sound and complete with respect to doxastic plausibility models. We present our completeness argument in Section 4, following the strategy adopted by [5] for APAL. In Section 5, we show the expressivity strengths of ARUL regarding action theory and related epistemic paradoxes. We conclude in Section 6 and set the basis for future logical extensions to other belief revision policies, and to more general results.

2 Preliminaries

We present the logic ARUL, an extension of classic dynamic belief revision logic [16,11,12] incorporating arbitrary radical upgrades. Unlike the aforementioned approaches, the syntax of ARUL contains a modal plausibility operator as in [34], instead of the a static modal operator for belief. Nevertheless, the belief operator is definable in our setting. In addition, ARUL features dynamic modal operators for plausibility upgrades with specific formulas, and a modality for upgrades with arbitrary formulas as a novelty.

Definition 1. *Let Prop be a countable set of propositional symbols. The set of all ARUL-formulas is given by the following Backus-Naur Form:*

$$\varphi, \psi ::= p \mid \neg\varphi \mid \varphi \vee \psi \mid [\leq]\varphi \mid \mathbf{K}\varphi \mid [\uparrow\varphi]\psi \mid [\uparrow]\psi,$$

where $p \in \text{Prop}$. On the one hand, the language contains the static modal formulas $[\leq]\varphi$ (standing for “ φ holds in all worlds the agent considers at least as plausible as the current one”), and $\mathbf{K}\varphi$ (for “the agent knows that φ ”). On the other hand, it contains dynamic formulas $[\uparrow\varphi]\psi$ (standing for “after a radical upgrade with formula φ , ψ holds”), and the arbitrary upgrade $[\uparrow]\psi$ (read as “after any radical upgrade takes place, ψ holds”). Finally, a formula φ is called epistemic, if it does not contain any occurrence of $[\uparrow\psi]$ or $[\uparrow]$.

The simple language of ARUL defined above enables us to define several other notions, as shown hereafter.

Definition 2. We define the possibility operator for knowledge as $\hat{\mathbf{K}}\varphi := \neg\mathbf{K}\neg\varphi$, and the existential modality for plausibility as $\langle\leq\rangle\varphi := \neg[\leq]\neg\varphi$ (standing for “there is at least a plausible world where φ holds”). Plain beliefs can be defined in terms of plausibility as $\mathbf{B}\varphi := \langle\leq\rangle[\leq]\varphi$ (see [34, Fact 1] for details), and read as “the agent believes that φ ”. We also define the existential radical upgrade as $\langle\uparrow\psi\rangle\varphi := \neg[\uparrow\psi]\neg\varphi$, and its arbitrary version as $\langle\uparrow\rangle\psi := \neg[\uparrow]\neg\psi$ (standing for “there exists a formula φ such that $[\uparrow\varphi]\psi$ holds”). Additional connectives (\top , \perp , \wedge , \rightarrow , \leftrightarrow) are defined as usual.

Formulas of ARUL are interpreted semantically by using (doxastic) plausibility models defined in [16,11,12,17]. Below we follow the presentation of [34].

Definition 3. A plausibility model for ARUL is a tuple $\mathfrak{S} = \langle\mathcal{S}, \leq, V\rangle$ where:

- \mathcal{S} is a countable non-empty set of “possible states” (or “worlds”),
- $\leq \subseteq \mathcal{S} \times \mathcal{S}$, the “plausibility relation” for the agent, is a locally connected pre-order, and
- $V : \text{Prop} \rightarrow \wp(\mathcal{S})$ is a standard “valuation map”, where $\wp(\mathcal{S})$ is the set of all subsets of \mathcal{S} .

The conventional reading of the plausibility order is that in case $s \leq t$ (for all $s, t \in \mathcal{S}$), the agent considers state t to be “at least as plausible as” state s . Given $s \in \mathcal{S}$, we call \mathfrak{S}, s a pointed model, with s being the “actual” state.

Definition 4. Let $\mathfrak{S} = \langle\mathcal{S}, \leq, V\rangle$ be a plausibility model. We define the epistemic accessibility relation \sim as $\sim := \leq \cup \leq^{-1}$. Notice that \sim is an equivalence relation (i.e., reflexive, symmetric and transitive relation), thus for each $s \in \mathcal{S}$ we define its equivalence class $[s]_{\sim} := \{t \mid s \sim t\}$.

Let $A \subseteq \mathcal{S}$. We define Max_{\leq}^A as the set of states in A that are maximal for the ordering \leq restricted to states in A , i.e.,

$$\text{Max}_{\leq}^A = \{u \in A \mid \text{for all } v \in A, v \leq u\}.$$

Recall that a *locally-connected pre-order* is a pre-order (i.e., it is reflexive and transitive), that is locally connected (for all $t, v \in [s]_{\sim}$ we have $t \sim v$). In e.g. [19,12,34], it is also required that the relation be *conversely well-founded* (i.e., it has a minimum element). However, as done in e.g. [16], we drop this condition as not being crucial in our logical presentation, and moreover, our completeness argument might no longer apply otherwise.

Definition 5. *The truth of a formula φ at the actual state s in the plausibility model \mathfrak{S} , denoted $\mathfrak{S}, s \models \varphi$, is defined inductively as follows:*

$$\begin{aligned}
\mathfrak{S}, s \models p & \quad \text{iff} \quad s \in V(p) \\
\mathfrak{S}, s \models \neg\varphi & \quad \text{iff} \quad \mathfrak{S}, s \not\models \varphi \\
\mathfrak{S}, s \models \varphi \vee \psi & \quad \text{iff} \quad \mathfrak{S}, s \models \varphi \text{ or } \mathfrak{S}, s \models \psi \\
\mathfrak{S}, s \models \leq\varphi & \quad \text{iff} \quad \text{for all } t \in \mathcal{S}, \text{ if } s \leq t \text{ then } \mathfrak{S}, t \models \varphi \\
\mathfrak{S}, s \models \mathbf{K}\varphi & \quad \text{iff} \quad \text{for all } t \in \mathcal{S}, \text{ if } s \sim t \text{ then } \mathfrak{S}, t \models \varphi \\
\mathfrak{S}, s \models [\uparrow\varphi]\psi & \quad \text{iff} \quad \mathfrak{S}^{[\uparrow\varphi]}, s \models \psi \\
\mathfrak{S}, s \models [\uparrow]\psi & \quad \text{iff} \quad \text{for all epistemic } \varphi, \mathfrak{S}, s \models [\uparrow\varphi]\psi,
\end{aligned}$$

where $\llbracket\psi\rrbracket^{\mathfrak{S}} := \{s \mid \mathfrak{S}, s \models \psi\}$, and $\mathfrak{S}^{[\uparrow\varphi]} = \langle \mathcal{S}, \leq^{[\uparrow\varphi]}, V \rangle$ is such that:

$$\leq^{[\uparrow\varphi]} := (\leq \cap (\mathcal{S} \times \llbracket\varphi\rrbracket^{\mathfrak{S}})) \cup (\leq \cap (\llbracket\neg\varphi\rrbracket^{\mathfrak{S}} \times \mathcal{S})) \cup (\sim \cap (\llbracket\neg\varphi\rrbracket^{\mathfrak{S}} \times \llbracket\varphi\rrbracket^{\mathfrak{S}})).$$

A formula φ is said to be satisfiable if $\mathfrak{S}, s \models \varphi$ for some model \mathfrak{S} and state s , and φ is valid in \mathfrak{S} if $\mathfrak{S} \models \varphi$ for all states s of \mathfrak{S} . Finally, φ is valid stricto sensu (notation: $\models \varphi$) if $\mathfrak{S} \models \varphi$ for all models \mathfrak{S} .

Following [16] and [12], the belief radical upgrade with respect to a formula φ defined above, written $[\uparrow\varphi]$, induces a mapping of the following kind:

$$[\uparrow\varphi] : \mathfrak{S} \mapsto \mathfrak{S}^{[\uparrow\varphi]}$$

Here, \mathfrak{S} is the initial plausibility model and $\mathfrak{S}^{[\uparrow\varphi]}$ is the transformed model obtained once the operation $[\uparrow\varphi]$ is performed on \mathfrak{S} . In this definition, \mathcal{S} and V remain unchanged. The special feature of $\mathfrak{S}^{[\uparrow\varphi]}$ is the plausibility order $\leq^{[\uparrow\varphi]}$. The reordering of states defined by $\leq^{[\uparrow\varphi]}$ ensures that the states where φ is true are promoted in plausibility. In the definition of $\leq^{[\uparrow\varphi]}$, the first part ($\leq \cap (\mathcal{S} \times \llbracket\varphi\rrbracket^{\mathfrak{S}})$) states that the relative ordering of worlds where φ is true is the same as in the original order \leq . The second part ($\leq \cap (\llbracket\neg\varphi\rrbracket^{\mathfrak{S}} \times \mathcal{S})$) states that the relative ordering of worlds where φ is false is the same as in the original order \leq . Finally, the third part ($\sim \cap (\llbracket\neg\varphi\rrbracket^{\mathfrak{S}} \times \llbracket\varphi\rrbracket^{\mathfrak{S}})$) states that the worlds where φ is true become equally or more plausible than worlds where φ is false in the locally connected component.

Remark 1. Notice that, by Definition 5, the radical upgrade operator $[\uparrow\psi]$ is self-dual, i.e., $[\uparrow\psi]\varphi \leftrightarrow \langle \uparrow\psi \rangle \varphi$ is valid, since $[\uparrow\psi]$ can always be executed (its semantics does not involve a pre-condition) and the mapping $[\uparrow\varphi] : \mathfrak{S} \mapsto \mathfrak{S}^{[\uparrow\varphi]}$ always yields one and only one model.

Finally, we state that plain beliefs can be recovered from plausibility.

Proposition 1 ([34]). *Let \mathfrak{S} be a plausibility model with s being one of its states, and let φ be an ARUL formula. Then,*

$$\mathfrak{S}, s \models \mathbf{B}\varphi \text{ iff for all } t \in \text{Max}_{\leq^{[s]}\sim}, \mathfrak{S}, t \models \varphi.$$

3 Axiom System

In this section, we present an axiom system for ARUL. This puts together the axiomatization for the $[\uparrow\varphi]$ -free fragment and reduction axioms to eliminate the occurrences of $[\uparrow\varphi]$ [11], plus axioms and rules for arbitrary announcements [4].

Our strategy closely follows the ideas introduced in [5] for APAL. In [4], it is noticed that single-agent APAL (over **S5** models) can be reduced into plain epistemic logic, while for $n > 1$ agents the expressive power of APAL is strictly greater than basic epistemic logic. For the case of ARUL, it has not been established whether this is also the case or not. Thus, to prove a completeness result for ARUL we will use the argument followed by [5] for multi-agent APAL.

Here we need to introduce the so-called *necessity forms* for ARUL. Necessity forms are crucial in axiomatizing the logic, as a mechanism for performing special restricted forms of substitution on certain formulas (see [24,5,1]).

Definition 6. *Let \sharp be a fresh propositional symbol. We define the set of necessity forms (whose members are notated $\eta(\sharp)$, $\eta'(\sharp)$, $\eta''(\sharp)$, etc.) of ARUL as:*

$$\eta(\sharp) ::= \sharp \mid \varphi \rightarrow \eta(\sharp) \mid \leq \eta(\sharp) \mid \mathbf{K}\eta(\sharp) \mid [\uparrow\varphi]\eta(\sharp),$$

where φ is a formula of ARUL.

Our strategy to provide a sound and complete axiom system for ARUL consists in three parts. Firstly, we provide standard modal axioms for the epistemic fragment of ARUL, i.e., the fragment whose only modalities are \leq and **K**, as in e.g. [34] (see also [18]). Secondly, we provide so-called *reduction axioms* for axiomatizing formulas containing the $[\uparrow\psi]$ modality. Reduction axioms enable us to eliminate occurrences of $[\uparrow\psi]$, as done in [34]. Finally, a block should be introduced for axiomatizing arbitrary upgrades $[\uparrow]$. To achieve this, we provide basic modal axioms and a modal necessitation rule using a necessity form, similar to the ones for arbitrary public announcements described in [5].

Definition 7. *The axiom system \mathcal{ARU} for ARUL is defined by the axioms and rules of Figure 1. Precisely, \mathcal{ARU} is the smallest set of formulas of ARUL that contains all the axioms of Figure 1 and it is closed by its rules.*

Let us briefly discuss the list of axioms and rules of Figure 1. The first block provides the propositional base of the system, as well as distribution axioms for each modality of ARUL. The second block axiomatizes the properties of knowledge (knowledge and factivity, and positive and negative introspection, respectively). In the third block, we introduce the axioms to capture the properties of the plausibility relation, as well as its connections with the knowledge modality **K**. Then, axiom (LC) establishes that the plausibility relation is locally connected, whereas (Int) characterizes the interaction between arbitrary and non-arbitrary radical upgrades. Finally, blocks are provided for the inferences rules, including Modus Ponens and Necessitation rules for each modality (using a necessity form for $[\uparrow]$), and for the reduction axioms of $[\uparrow\chi]$, including axiom (A5) which mimics every possible outcomes resulting from the application of a plausibility upgrade, and (A6) describing the composition of two upgrades.

Basic Axioms	
CPL	All tautologies from propositional logic
$K_{[\leq]}$	$[\leq](\varphi \rightarrow \psi) \rightarrow ([\leq]\varphi \rightarrow [\leq]\psi)$
K_K	$\mathbf{K}(\varphi \rightarrow \psi) \rightarrow (\mathbf{K}\varphi \rightarrow \mathbf{K}\psi)$
$K_{\uparrow\chi}$	$[\uparrow\chi](\varphi \rightarrow \psi) \rightarrow ([\uparrow\chi]\varphi \rightarrow [\uparrow\chi]\psi)$
K_{\uparrow}	$[\uparrow](\varphi \rightarrow \psi) \rightarrow ([\uparrow]\varphi \rightarrow [\uparrow]\psi)$
T_K	$\mathbf{K}\varphi \rightarrow \varphi$
4_K	$\mathbf{K}\varphi \rightarrow \mathbf{K}\mathbf{K}\varphi$
5_K	$\neg\mathbf{K}\varphi \rightarrow \mathbf{K}\neg\mathbf{K}\varphi$
$T_{[\leq]}$	$[\leq]\varphi \rightarrow \varphi$
$4_{[\leq]}$	$[\leq]\varphi \rightarrow [\leq][\leq]\varphi$
Inc	$\mathbf{K}\varphi \rightarrow [\leq]\varphi$
LC	$(\hat{\mathbf{K}}\varphi \wedge \hat{\mathbf{K}}\psi) \rightarrow \hat{\mathbf{K}}(\varphi \wedge \langle \leq \rangle \psi) \vee \hat{\mathbf{K}}(\psi \wedge \langle \leq \rangle \varphi)$
Int	$[\uparrow]\varphi \rightarrow [\uparrow\psi]\varphi$, for ψ epistemic
Inference Rules	
MP	from φ and $\varphi \rightarrow \psi$ infer ψ
$\text{Nec}_{[\leq]}$	from φ infer $[\leq]\varphi$
Nec_K	from φ infer $\mathbf{K}\varphi$
$\text{Nec}_{[\uparrow\chi]}$	from φ infer $[\uparrow\chi]\varphi$
$\text{Nec}_{[\uparrow]}$	from $\eta([\uparrow\psi]\varphi)$ infer $\eta([\uparrow]\varphi)$, for ψ epistemic
Reduction Axioms	
(A1)	$[\uparrow\chi]p \leftrightarrow p$
(A2)	$[\uparrow\chi]\neg\varphi \leftrightarrow \neg[\uparrow\chi]\varphi$
(A3)	$[\uparrow\chi](\varphi \vee \psi) \leftrightarrow [\uparrow\chi]\varphi \vee [\uparrow\chi]\psi$
(A4)	$[\uparrow\chi]\mathbf{K}\varphi \leftrightarrow \mathbf{K}[\uparrow\chi]\varphi$
(A5)	$[\uparrow\chi][\leq]\varphi \leftrightarrow [\leq](\chi \rightarrow [\uparrow\chi]\varphi) \wedge (\neg\chi \rightarrow [\leq][\uparrow\chi]\varphi) \wedge (\neg\chi \rightarrow \mathbf{K}(\chi \rightarrow [\uparrow\chi]\varphi))$
(A6)	$[\uparrow\chi][\uparrow\psi]\varphi \leftrightarrow [\uparrow([\uparrow\chi]\psi)]\varphi$

Fig. 1: Axiom system and inference rules \mathcal{ARU} for ARUL.

Lemma 1 (Soundness). *The axiomatization \mathcal{ARU} is sound, i.e., all the axioms are valid formulas, while inference rules preserve validity over formulas (on plausibility models).*

Proof. Soundness of most axioms and rules (or variants of them) has been shown in e.g. [19,12,34], and for arbitrary announcements in [3,5,2]. The reduction axiom (A6) deserves more attention, since upgrades deal with relation updates rather than worlds updates. We need to prove that: $s \leq^{[\uparrow\chi][\uparrow\psi]} t$ iff $s(\leq^{[\uparrow\chi]})^{[\uparrow\psi]} t$. Suppose $s \leq^{[\uparrow\chi][\uparrow\psi]} t$, then we have three alternatives:

1. $s \leq t$ and $\mathfrak{S}, t \models [\uparrow\chi]\psi$. Then, $s \leq^{[\uparrow\chi]} t$ and $\mathfrak{S}^{[\uparrow\chi]}, t \models \psi$.
2. $s \leq t$ and $\mathfrak{S}, t \models \neg[\uparrow\chi]\psi$. The latter is equivalent to state $\mathfrak{S}, t \models [\uparrow\chi]\neg\psi$, by (A2). Then, we get $s \leq^{[\uparrow\chi]} t$ and $\mathfrak{S}^{[\uparrow\chi]}, s \models \neg\psi$.

3. $s \sim t$, $\mathfrak{S}, s \models [\uparrow\chi]\psi$, and $\mathfrak{S}, t \models \neg[\uparrow\chi]\psi$. For a similar reason as above, we get $s \sim^{[\uparrow\chi]} t$, $\mathfrak{S}^{[\uparrow\chi]}, s \models \neg\psi$ and $\mathfrak{S}^{[\uparrow\chi]}, t \models \psi$.

The other direction of the implication is similar. Thus, we conclude $s(\leq^{[\uparrow\chi]})^{[\uparrow\psi]}t$.

We list now a number of definitions and properties that are useful to establish completeness of the axiom system \mathcal{ARU} . These properties are mostly inspired by their analogues from [5] (some proofs and properties are omitted for space reasons). We start by defining the notion of a *theory*, which is our main ingredient to build a canonical model.

Definition 8. *Let Γ be a set of formulas. We call Γ a theory if: 1) Γ contains \mathcal{ARU} , and 2) Γ is closed under MP and $\text{Nec}_{[\uparrow]}$.*

A theory Γ is consistent if $\perp \notin \Gamma$, and it is maximal if for all φ in \mathcal{ARUL} , we have either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$. We denote as MCT the set of all theories that are both consistent and maximal.

Notice that a theory is only required to be closed under rules MP and $\text{Nec}_{[\uparrow]}$, as these rules preserve truth while the others preserve only validity. Next, we present classical properties which ensure that a theory behaves as expected.

Lemma 2. *Let Γ be a MCT. The following properties hold:*

1. $\perp \notin \Gamma$,
2. $\varphi \in \Gamma$ iff $\neg\varphi \notin \Gamma$,
3. $\varphi \vee \psi \in \Gamma$ iff either $\varphi \in \Gamma$ or $\psi \in \Gamma$.

Definition 9. *Let Γ be a theory. Define:*

$$\begin{aligned} [\uparrow\psi]\Gamma &= \{\varphi \mid [\uparrow\psi]\varphi \in \Gamma\} & \Gamma \oplus \varphi &= \{\psi \mid \varphi \rightarrow \psi \in \Gamma\} \\ \mathbf{K}\Gamma &= \{\varphi \mid \mathbf{K}\varphi \in \Gamma\} & [\leq]\Gamma &= \{\varphi \mid [\leq]\varphi \in \Gamma\}. \end{aligned}$$

Lemma 3. *Let Γ be a theory, and let φ be a formula of \mathcal{ARUL} . Then, $[\uparrow\varphi]\Gamma$, $\Gamma \oplus \varphi$, $\mathbf{K}\Gamma$ and $[\leq]\Gamma$ are theories. Moreover, $\Gamma \cup \{\varphi\} \subseteq \Gamma \oplus \varphi$, and $\Gamma \oplus \varphi$ is consistent iff $\neg\varphi \notin \Gamma$.*

Lemma 4. *Each consistent theory can be extended to a MCT.*

Now, we provide some properties that are essential to guarantee that the canonical model is a proper plausibility model for \mathcal{ARUL} .

Lemma 5. *Let Γ, Δ, Π be MCTs. Then,*

1. $\mathbf{K}\Gamma \subseteq \Gamma$, $[\leq]\Gamma \subseteq \Gamma$, and $\mathbf{K}\Gamma \subseteq [\leq]\Gamma$,
2. if $[\leq]\Gamma \subseteq \Delta$ and $[\leq]\Delta \subseteq \Pi$ then $[\leq]\Gamma \subseteq \Pi$,
3. if $\mathbf{K}\Gamma \subseteq \Delta$ and $\mathbf{K}\Delta \subseteq \Pi$ then $\mathbf{K}\Gamma \subseteq \Pi$,
4. if $\mathbf{K}\Gamma \subseteq \Delta$ then $\mathbf{K}\Delta \subseteq \Gamma$.

4 Completeness

We can now introduce our completeness argument for ARUL, following the standard methodology via a *canonical model*. To do so, we borrow the developments of [5] for proving completeness of APAL.

Definition 10 (Canonical Model). *The canonical model for ARUL is defined as $\mathfrak{S}^c = \langle \mathcal{S}^c, \leq^c, V^c \rangle$, where:*

- $\mathcal{S}^c = \{\Gamma \mid \Gamma \text{ is a MCT}\}$,
- $\leq^c = \{(\Gamma, \Gamma') \mid [\leq] \Gamma \subseteq \Gamma'\}$,
- $V^c(p) = \{p \mid p \in \Gamma\}$, for all $p \in \text{Prop}$.

Clearly, \mathfrak{S}^c is a plausibility model (e.g. structural properties on \leq^c are guaranteed by Lemma 5). We now introduce some definitions and syntactic properties, that are useful to properly treat formulas in our completeness proof.

Definition 11. *The size of a formula φ of ARUL, is written $\text{size}(\varphi)$ and defined inductively as follows:*

$$\begin{aligned} \text{size}(p) &= 1 & \text{size}(\varphi \vee \psi) &= 1 + \max\{\text{size}(\varphi), \text{size}(\psi)\} \\ \text{size}(\delta\varphi) &= 1 + \text{size}(\varphi) & \text{size}([\uparrow\psi]\varphi) &= 5 \cdot \text{size}(\varphi) + \text{size}(\psi), \end{aligned}$$

with $\delta \in \{\neg, \mathbf{K}, [\leq], [\uparrow]\}$. The \uparrow -depth of φ , written $\text{depth}^\uparrow(\varphi)$, is defined as:

$$\begin{aligned} \text{depth}^\uparrow(p) &= 0 & \text{depth}^\uparrow(\varphi \vee \psi) &= \max\{\text{depth}^\uparrow(\varphi), \text{depth}^\uparrow(\psi)\} \\ \text{depth}^\uparrow(\gamma\varphi) &= \text{depth}^\uparrow(\varphi) & \text{depth}^\uparrow([\uparrow\psi]\varphi) &= \max\{\text{depth}^\uparrow(\varphi), \text{depth}^\uparrow(\psi)\} \\ \text{depth}^\uparrow([\uparrow]\varphi) &= 1 + \text{depth}^\uparrow(\varphi), \end{aligned}$$

with $\gamma \in \{\neg, \mathbf{K}, [\leq]\}$. We write $\varphi <_d^s \psi$ iff either $\text{depth}^\uparrow(\varphi) < \text{depth}^\uparrow(\psi)$, or $\text{depth}^\uparrow(\varphi) = \text{depth}^\uparrow(\psi)$ and $\text{size}(\varphi) < \text{size}(\psi)$.

Lemma 6. *The relation $<_d^s$ is a well-founded strict partial order over formulas.*

In Definition 11, we define $\text{size}(\varphi \vee \psi)$ as $1 + \max\{\text{size}(\varphi), \text{size}(\psi)\}$. By contrast, in [5], $\text{size}(\varphi \vee \psi)$ is defined as $1 + \text{size}(\varphi) + \text{size}(\psi)$. In addition, the ‘‘curious’’ factor of 3 in $\text{size}([\uparrow\psi]\varphi)$ becomes 5 here, the aim of which is to guarantee the application of the inductive hypothesis in the proof of Lemma 8. Below we prove the special property about the order $<_d^s$ that is not present in [5].

Lemma 7. *Let χ, ψ be ARUL formulas. Then,*

- 1) $[\leq](\neg\chi \vee [\uparrow\chi]\psi) <_d^s [\uparrow\chi][\leq]\psi$
- 2) $[\leq](\chi \vee [\leq][\uparrow\chi]\psi) <_d^s [\uparrow\chi][\leq]\psi$
- 3) $[\leq](\chi \vee \mathbf{K}(\chi \rightarrow [\uparrow\chi]\psi)) <_d^s [\uparrow\chi][\leq]\psi$.

Proof. In all cases, the \uparrow -depths coincide, so we need to check their sizes. Notice that by applying repeatedly Definition 11, we get:

$$\begin{aligned} \text{size}([\uparrow\chi][\leq]\psi) &= 5 \cdot \text{size}([\leq]\psi) + \text{size}(\chi) \\ &= 5 \cdot (1 + \text{size}(\psi)) + \text{size}(\chi) \\ &= 5 + 5 \cdot \text{size}(\psi) + \text{size}(\chi) \end{aligned}$$

Item 1 is shown below:

$$\begin{aligned}
\text{size}([\leq](\neg\chi \vee [\uparrow\chi]\psi)) &= 2 + \max\{1 + \text{size}(\chi), \text{size}([\uparrow\chi]\psi)\} \\
&= 2 + \max\{1 + \text{size}(\chi), 5 \cdot \text{size}(\psi) + \text{size}(\chi)\} \quad (\dagger) \\
&= 2 + 5 \cdot \text{size}(\psi) + \text{size}(\chi) < \text{size}([\uparrow\chi][\leq]\psi)
\end{aligned}$$

Step (\dagger) follows from the fact that the size is always at least 1. Using similar reasoning, we get $\text{size}([\leq](\chi \vee [\leq][\uparrow\chi]\psi)) = 3 + 5 \cdot \text{size}(\psi) + \text{size}(\chi) < \text{size}([\uparrow\chi][\leq]\psi)$, then item 2 follows. Finally, for item 3, we have:

$$\begin{aligned}
\text{size}([\leq](\chi \vee \mathbf{K}(\chi \rightarrow [\uparrow\chi]\psi))) &= 2 + \max\{\text{size}(\chi), \\
&\quad 2 + \max\{1 + \text{size}(\chi), 5 \cdot \text{size}(\psi) + \text{size}(\chi)\}\} \\
&= 4 + 5 \cdot \text{size}(\psi) + \text{size}(\chi) < \text{size}([\uparrow\chi][\leq]\psi)
\end{aligned}$$

With this property at hand, in addition to those introduced in [5], we can proceed with the crucial result in this section.

Lemma 8. *Let φ be a formula, let conditions P and H be defined as follows:*

$P(\varphi)$: *for all MCT Γ , we have $\varphi \in \Gamma$ iff $\mathfrak{S}^c, \Gamma \models \varphi$;*

$H(\varphi)$: *for all formulas ψ , if $\psi <_d^s \varphi$, then $P(\varphi)$.*

Then, if $H(\varphi)$ then $P(\varphi)$.

Proof. The proof is by structural induction on φ . One interesting case is $\varphi = [\uparrow\chi][\leq]\psi$. Assuming $H([\uparrow\chi][\leq]\psi)$, we suppose $[\uparrow\chi][\leq]\psi \in \Gamma$. By (A5), we get $[\leq](\chi \rightarrow [\uparrow\chi]\varphi) \wedge (\neg\chi \rightarrow [\leq][\uparrow\chi]\varphi) \wedge (\neg\chi \rightarrow \mathbf{K}(\chi \rightarrow [\uparrow\chi]\varphi)) \in \Gamma$. Then, we have $\{[\leq](\chi \rightarrow [\uparrow\chi]\varphi), [\leq](\neg\chi \rightarrow [\leq][\uparrow\chi]\varphi), [\leq](\neg\chi \rightarrow \mathbf{K}(\chi \rightarrow [\uparrow\chi]\varphi))\} \subseteq \Gamma$. By CPL, $\{[\leq](\neg\chi \vee [\uparrow\chi]\varphi), [\leq](\chi \vee [\leq][\uparrow\chi]\varphi), [\leq](\chi \vee \mathbf{K}(\chi \rightarrow [\uparrow\chi]\varphi))\} \subseteq \Gamma$. Thus, by Lemma 7, IH and (A5), we get $\mathfrak{S}^c, \Gamma \models [\uparrow\chi][\leq]\psi$. Hence, $P([\uparrow\chi][\leq]\psi)$ holds.

Lemma 9 (Truth Lemma). *Let φ be a formula of ARUL, and let Γ be a MCT. Then, $\varphi \in \Gamma$ iff $\mathfrak{S}^c, \Gamma \models \varphi$.*

Proof. By Lemmas 6 and 8.

Theorem 1. *The axiomatic system ARU from Figure 1 is sound and complete for ARUL over the class of plausibility models from Definition 3.*

Proof. Soundness follows by Lemma 1. For completeness, we need to show that $\models \varphi$ implies $\varphi \in \text{ARU}$. Aiming for a contradiction, suppose that $\models \varphi$ and that $\varphi \notin \text{ARU}$. By Lemmas 3 and 4, there exists a MCT Γ such that $\neg\varphi \in \Gamma$. By Lemma 9, $\mathfrak{S}^c, \Gamma \models \neg\varphi$, thus $\mathfrak{S}^c, \Gamma \not\models \varphi$. Then, $\not\models \varphi$, a contradiction. Therefore, ARU is complete for ARUL over the class of plausibility models from Definition 3.

5 Applications

This section shows that our framework fruitfully paves the way to address some central philosophical issues. We first discuss the *surprise deception paradox* investigated from a dynamic doxastic logic perspective in [28], — a work that inspired the definition of our arbitrary upgrades. Then, we discuss *Fitch's knowability paradox* [22], and *Moore's paradox* (see e.g. [29,25]) in the context of logic ARUL.

Application 1 (Surprise Deception Paradox). ARUL elaborates on the arbitrary radical upgrade introduced meta-logically in [28] using standard radical upgrades. The goal in [28] is to analyze a Smullyan’s puzzle on deception and surprise that can be summarized as follows: on the morning of April Fool’s Day, an agent announces to an addressee that he or she would be deceived later on that day, but apparently nothing happened. Since the addressee waited all day to be deceived by some action, he or she was actually deceived, but by the lack of an action, i.e., by omission. This deception left the addressee strongly surprised.

In [28], such a deception-based surprise is modelled along the lines of Gerbrandy’s formalization of the *three-day* Surprise Examination Paradox [23]. The difference is that exam days are replaced by types of deception, based on capturing the distinction between deception as a result of an action, or *deception by commission*, and deception by the lack of an action, or *deception by omission*.

The distinction between deception by commission versus by omission can be defined in ARUL using the plausibility operator $[\leq]$ to characterize deception, the existential radical upgrade $\langle \uparrow \rangle$ to capture the existence of an action, or *commission*, and its negation $\neg \langle \uparrow \rangle$ for the absence of such an action, or *omission*.

We define the fact that the addressee is deceived on a formula ψ in case ψ is true but the addressee judges $\neg\psi$ as more highly plausible, i.e., as $(\psi \wedge [\leq]\neg\psi)$.

Definition 12. *Let ψ be an arbitrary formula of ARUL, we define the fact that the addressee is deceived by commission on formula ψ (written \mathbf{d}_ψ^+), and that she is deceived by omission on ψ (written \mathbf{d}_ψ^-), respectively as:*

$$\mathbf{d}_\psi^+ := \langle \uparrow \rangle (\psi \wedge [\leq] \neg \psi); \quad \mathbf{d}_\psi^- := \neg \langle \uparrow \rangle (\psi \wedge [\leq] \neg \psi).$$

In [28], definitions of \mathbf{d}_ψ^+ and \mathbf{d}_ψ^- are used to express the fact that the agent’s deception results in the addressee’s surprise. Let \mathbf{d}^+ and \mathbf{d}^- be abbreviations for the complex formulas \mathbf{d}_ψ^+ and \mathbf{d}_ψ^- , respectively. Following [23], modulo minor adaptations to radical upgrades [10], the surprise aspect of the agent’s announcement can be encoded by the formula \mathbf{D} as follows:

$$\begin{aligned} \mathbf{D} := & ((\mathbf{d}^+ \wedge \neg \mathbf{d}^-) \wedge \neg [\leq] (\mathbf{d}^+ \wedge \neg \mathbf{d}^-)) \vee \\ & ((\neg \mathbf{d}^+ \wedge \mathbf{d}^-) \wedge [\uparrow \neg (\mathbf{d}^+ \wedge \neg \mathbf{d}^-)] \neg [\leq] (\neg \mathbf{d}^+ \wedge \mathbf{d}^-)) \vee \\ & ((\mathbf{d}^+ \wedge \mathbf{d}^-) \wedge [\uparrow \neg (\mathbf{d}^+ \wedge \neg \mathbf{d}^-)] [\uparrow \neg (\neg \mathbf{d}^+ \wedge \mathbf{d}^-)] \neg [\leq] (\mathbf{d}^+ \wedge \mathbf{d}^-)). \end{aligned}$$

The first disjunct of \mathbf{D} states that in case addressees are deceived only by commission, i.e. $(\mathbf{d}^+ \wedge \neg \mathbf{d}^-)$, they will be surprised because they deny this possibility as being plausible: $\neg [\leq] (\mathbf{d}^+ \wedge \neg \mathbf{d}^-)$. The second disjunct reflects an analogous situation in presence of deception by omission. The third disjunct of \mathbf{D} states that in case addressees are deceived both by commission and by omission, i.e. $(\mathbf{d}^+ \wedge \mathbf{d}^-)$, they will be surprised because after rejecting deception only by commission and only by omission: $[\uparrow \neg (\mathbf{d}^+ \wedge \neg \mathbf{d}^-)] [\uparrow \neg (\neg \mathbf{d}^+ \wedge \mathbf{d}^-)]$, they (still) deny option $(\mathbf{d}^+ \wedge \mathbf{d}^-)$ as being plausible: $\neg [\leq] (\mathbf{d}^+ \wedge \mathbf{d}^-)$.

Following [23,10], though, [28] shows that any radical upgrade with \mathbf{D} , or with a reinforced formula involving \mathbf{D} , will result in the agent failing to surprise

the addressee. But this failure is theoretical since in the puzzle at stake, the addressee is actually surprised as he fails to predict the type of deception he is preyed to, i.e. deception both by commission and by omission. The puzzle dynamics and interplay between deception and surprise can be modelled in ARUL. One noticeable difference is that arbitrary upgrades are defined externally in in [28]. ARUL goes a step further by internalizing these upgrades as primitive operators of the syntax for which sound and complete axiomatization is proven.

Application 2 (Fitch’s Paradox). Fitch’s knowability paradox [22,20] is rooted in the verificationist thesis according to which all truths are verifiable. Following van Benthem’s version in [15], this principle can be interpreted in terms of learnability and knowability as: “*what is true may come to be known*”. Accordingly, it can be expressed in APAL using the schema: $\{\text{for all } \varphi, \varphi \rightarrow \langle ! \rangle \mathbf{K}\varphi\}$, where φ is some given truth of the language, and $\langle ! \rangle$ is the arbitrary announcement modality. As shown in [15], paradox arises with the case of truths that are unknown by agents, expressed as $(\psi \wedge \neg \mathbf{K}\psi)$. It has been established in e.g. [6,7] that substituting $(\psi \wedge \neg \mathbf{K}\psi)$ to φ in the knowability principle with arbitrary announcements, leads to inconsistency and thus, to the necessary acceptance of the unrealistic omniscience conclusion that all truths are already known: $\varphi \rightarrow \mathbf{K}\varphi$.

It turns out that a variant of Fitch’s paradox with plausibility operator $\llbracket \leq \rrbracket$ and arbitrary upgrade $\langle \uparrow \rangle$ applies to ARUL. Let $\mathfrak{S} = \langle \mathcal{S}, \leq, V \rangle$ be a plausibility model with $s \in \mathcal{S}$. Aiming for a contradiction, suppose that the learnability principle transfers to the ARUL operators $\llbracket \leq \rrbracket$ and $\langle \uparrow \rangle$, stating that all truth ψ of ARUL will be considered plausible by the agent after some arbitrary upgrade: $\{\text{for all } \varphi, \varphi \rightarrow \langle \uparrow \rangle \llbracket \leq \rrbracket \varphi\}$. Suppose that a given truth ψ of ARUL is judged implausible by the agent: $(\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. By instantiating the learnability principle with this formula, substituting $(\psi \wedge \neg \llbracket \leq \rrbracket \psi)$ to φ , we have: $\mathfrak{S}, s \models (\psi \wedge \neg \llbracket \leq \rrbracket \psi) \rightarrow \langle \uparrow \rangle \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. By modus ponens: $\mathfrak{S}, s \models \langle \uparrow \rangle \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. By Definition 5, for some epistemic φ , $\mathfrak{S}, s \models \langle \uparrow \varphi \rangle \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$, iff $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. Observe that $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$ leads to contradiction since, by distributing $\llbracket \leq \rrbracket$ over \wedge , we have $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket \psi \wedge \llbracket \leq \rrbracket \neg \llbracket \leq \rrbracket \psi$. So, in particular $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket \psi$. But we also have $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket \neg \llbracket \leq \rrbracket \psi$. By instantiating axiom $\mathbf{T}_{\llbracket \leq \rrbracket}$, it holds that $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \llbracket \leq \rrbracket \neg \llbracket \leq \rrbracket \psi \rightarrow \neg \llbracket \leq \rrbracket \psi$. Then, $\mathfrak{S}^{\llbracket \uparrow \varphi \rrbracket}, s \models \neg \llbracket \leq \rrbracket \psi$, a contradiction.

Application 3 (Moore’s Paradox). Hidden behind Fitch’s paradox are so-called “*unsuccessful formulas*” [21], i.e. formulas φ that do not necessarily hold after an update with φ . As noticed in e.g. [26,14], this phenomenon is at the heart of Moore’s paradox [29], since announcing a true formula ψ that is not believed by the agent, i.e. a formula of the form $\psi \wedge \neg \mathbf{B}\psi$, leads to contradiction. This concerns **KD45** belief modalities, as well as **S5** knowledge modalities.

In ARUL, arbitrary upgrades $\llbracket \uparrow \rrbracket$ on formula $(\psi \wedge \neg \llbracket \leq \rrbracket \psi)$ give rise to such a Moorean phenomenon. Let \mathfrak{S} be a plausibility model with a state $s \in \mathcal{S}$. Let $(\psi \wedge \neg \llbracket \leq \rrbracket \psi)$ be the formula stating that ψ is true but not judged as plausible by the agent, and suppose that $\mathfrak{S}, s \models (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. As a way towards a contradiction, suppose that $\llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$ holds after any universal arbitrary radical upgrade, i.e., that $\mathfrak{S}, s \models \llbracket \uparrow \rrbracket \llbracket \leq \rrbracket (\psi \wedge \neg \llbracket \leq \rrbracket \psi)$. By Definition 5, it is the

case that $\mathfrak{S}, s \models [\uparrow][\leq](\psi \wedge \neg[\leq]\psi)$ if and only if $\mathfrak{S}, s \models [\uparrow\varphi][\leq](\psi \wedge \neg[\leq]\psi)$ for all epistemic $\varphi \in \text{ARUL}$. So this applies in particular for $\varphi \equiv (\psi \wedge \neg[\leq]\psi)$, i.e. $\mathfrak{S}, s \models [\uparrow(\psi \wedge \neg[\leq]\psi)][\leq](\psi \wedge \neg[\leq]\psi)$. A contradiction ensues since in presence of Moorean formulas, the formula $[\uparrow\varphi][\leq]\varphi$ is not valid (see e.g., [34]).

Accordingly, $(\psi \wedge \neg[\leq]\psi)$ is an unsuccessful formula of ARUL. Actually, there is no escaping the Moore’s paradox with the arbitrary variation of the learnability principle holding, as shown above with the instantiation of $(\psi \wedge \neg[\leq]\psi)$. As for APAL, this issue also concerns the more classical version of Moore’s paradox involving plain beliefs **B**, now with the version: $\{\text{for all } \varphi, \varphi \rightarrow \langle \uparrow \rangle \mathbf{B}\varphi\}$.

6 Conclusion

This paper introduces ARUL, a variant of dynamic belief revision logic that incorporates arbitrary radical upgrades, a concept not previously formalized in the literature. The syntax and semantics of ARUL extend the standard belief revision logics by allowing arbitrary quantification of radical upgrades. We present a sound and complete axiomatization for ARUL, operating over doxastic plausibility models characterized by a reflexive, transitive, and locally connected plausibility relation. Our completeness result relies on the proof for arbitrary announcements presented in [5]. Interestingly, ARUL complements extant logical frameworks by helping analyze classical epistemic paradoxes, such as the surprise examination paradox, Fitch’s paradox, and Moore’s paradox.

As pointed out in [5], the argument therein inspires a new realm of logics featuring information quantification, and our work should be seen as a first step towards the understanding of quantified upgrades in general. In this regard, many other dynamic belief revision policies exist that are weaker than radical upgrades [9,32]. A classical example is the “conservative upgrade” [16,11] with respect to φ , in which *only the best* φ -states increase in plausibility to reach the top of the plausibility ordering, leaving the rest of the ordering unchanged. Preliminary results obtained for ARUL could be extended to conservative upgrades. Also interesting would be to deal with arbitrary upgrades over different classes of models, including conversely well-founded orders as in [12,34]. Moreover, we would like to characterize the exact expressivity of ARUL, for instance to determine whether single-agent ARUL is already more expressive than its $[\uparrow]$ -free fragment or not. In fact, this would help adjudicate on the robustness of the puzzles analyzed with ARUL, showing whether Fitch’s and Moore’s paradoxes still hold in case of increased expressivity. Finally, this framework can be used to investigate agentic notions tied to commission versus omission and related notions (see, e.g. [13,35]). We leave those investigations for future work.

Acknowledgements We thank three anonymous reviewers for helpful comments and feedback, and Fernando R. Velázquez-Quesada for his advice. RF is supported by the IRP-SINFIN, the EU Grant 101008233 (MISSION), the CONICET project PIBAA-28720210100165CO and the ANPCyT project PICT-2021-00400. BI thanks the programs THEMIS (DOS0222794/00 and DOS0222795/00) and HYBRINFOX (ANR-21-ASIA-0003).

References

1. Ågotnes, T., Alechina, N., Galimullin, R.: Logics with group announcements and distributed knowledge: Completeness and expressive power. *Journal of Logic, Language and Information* **31**(2), 141–166 (2022)
2. Balbiani, P.: Putting right the wording and the proof of the truth lemma for *APAL*. *Journal of Applied Logic* **25**(1), 2–19 (2015)
3. Balbiani, P., Baltag, A., van Ditmarsch, H., Herzig, A., Hoshi, T., de Lima, T.: What can we achieve by arbitrary announcements?: A dynamic take on fitch’s knowability. In: Samet, D. (ed.) *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2007)*. pp. 42–51 (2007)
4. Balbiani, P., Baltag, A., van Ditmarsch, H., Herzig, A., Hoshi, T., de Lima, T.: ‘knowable’ as ‘known after an announcement’. *Review of Symbolic Logic* **1**(3), 305–334 (2008)
5. Balbiani, P., van Ditmarsch, H.: A simple proof of the completeness of *APAL*. *CoRR abs/1409.2612* (2014), <http://arxiv.org/abs/1409.2612>
6. Balbiani, P., Baltag, A., Van Ditmarsch, H., Herzig, A., Hoshi, T., De Lima, T.: What can we achieve by arbitrary announcements? a dynamic take on fitch’s knowability. In: *Proceedings of the 11th conference on Theoretical aspects of rationality and knowledge*. pp. 42–51 (2007)
7. Balbiani, P., Baltag, A., Van Ditmarsch, H., Herzig, A., Hoshi, T., De Lima, T.: Knowable’as ‘known after an announcement. *The Review of Symbolic Logic* **1**(3), 305–334 (2008)
8. Baltag, A., Moss, L.S., Solecki, S.: The logic of public announcements, common knowledge, and private suspicions. *Proceeding TARK ’98 Proceedings of the 7th conference on Theoretical aspects of rationality and knowledge* (1998)
9. Baltag, A., Rodenhäuser, B., Smets, S.: Doxastic attitudes as belief-revision policies. In: *Proceedings of the ESSLLI workshop on strategies for learning, belief revision and preference change* (2012)
10. Baltag, A., Smets, S.: Surprise?! an answer to the hangman, or how to avoid unexpected exams! In: *Logic and Interactive Rationality Seminar (LIRA)*, slides (2009)
11. Baltag, A., Smets, S.: Dynamic belief revision over multi-agent plausibility models. In: *Proceedings of LOFT*. vol. 6, pp. 11–24 (2006)
12. Baltag, A., Smets, S.: A qualitative theory of dynamic interactive belief revision. *Logic and the foundations of game and decision theory (LOFT 7)* **3**, 9–58 (2008)
13. Belnap, N.: Before refraining: Concepts for agency. *Erkenntnis* **34**(2), 137–169 (1991)
14. van Benthem, J.: *Logical dynamics of information and interaction*. Cambridge University Press (2011)
15. van Benthem, J.: What one may come to know. *Analysis* **64**(282), 95–105 (2004)
16. van Benthem, J.: Dynamic logic for belief revision. *Journal of applied non-classical logics* **17**(2), 129–155 (2007)
17. van Benthem, J., Smets, S.: Dynamic logics of belief change. In: *Handbook of Logics for Knowledge and Belief*, pp. 299–368. College Publications (2015)
18. Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*, Cambridge Tracts in Theoretical Computer Science, vol. 53. Cambridge University Press (2001)
19. Board, O.: Dynamic interactive epistemology. *Games and Economic Behavior* **49**(1), 49–80 (2004)
20. Brogaard, B., Salerno, J.: Fitch’s paradox of knowability (2002)

21. van Ditmarsch, H.P., Kooi, B.: The secret of my success. *Synthese* **153**(2), 339–339 (2006)
22. Fitch, F.B.: A logical analysis of some value concepts¹. *The journal of symbolic logic* **28**(2), 135–142 (1963)
23. Gerbrandy, J.: The surprise examination in dynamic epistemic logic. *Synthese* **155**(1), 21–33 (2007)
24. Goldblatt, R.: *Axiomatising the Logic of Computer Programming*. Springer-Verlag, Berlin, Heidelberg (1982)
25. Hintikka, J.: *Knowledge and Belief*. Cornell University Press, Ithaca N.Y. (1962)
26. Holliday, W.H., Icard III, T.F.: Moorean phenomena in epistemic logic. *Advances in Modal Logic*, Volume 8 **8**, 178–199 (2010)
27. Icard, B.: *Lying, deception and strategic omission: definition and evaluation*. Ph.D. thesis, Université Paris sciences et lettres (2019)
28. Icard, B.: The Surprise Deception Paradox (Jul 2024), <https://hal.science/hal-04647404>
29. Moore, G.E.: A reply to my critics. In: Schilpp, P.A. (ed.) *The philosophy of G. E. Moore*. Tudor Pub. Co. (1942)
30. Plaza, J.: Logics of public communications. In: Emrich, M.L., Pfeifer, M.S., Hadzikadic, M., Ras, Z.W. (eds.) *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*. pp. 201–216 (1989)
31. Plaza, J.: Logics of public communications. *Synthese* **158**(2), 165 (2007)
32. Rodenhäuser, B.: *A matter of trust: Dynamic attitudes in epistemic logic*. Universiteit van Amsterdam [Host] (2014)
33. Smullyan, R.M.: *What is the Name of this Book? The Riddle of Dracula and Other Logical Puzzles: Mysteries, Paradoxes, Gödel’s Discovery*. Prentice-Hall (1978)
34. Velázquez-Quesada, F.R.: Dynamic epistemic logic for implicit and explicit beliefs. *Journal of Logic, Language and Information* **23**(2), 107–140 (2014)
35. Xu, M.: Doing and refraining from refraining. *Journal of philosophical logic* **23**, 621–632 (1994)