Exploring subset models for justification logic

Eveline Lehmann and Thomas Studer
Institute of Computer Science, University of Bern, Switzerland
{eveline.lehmann,thomas.studer}@inf.unibe.ch

Abstract. Justification logic is a refinement of modal logic that includes explicit justifications for an agent's knowledge. So far, most semantics for justification logic interpret justifications symbolically, that is, as sets of formulas. We present a different and more flexible approach, called subset semantics, that models justifications as sets of possible worlds. We compare two variants of subset semantics and show that they are essentially different. Further, we discuss subset semantics in the context of hyperintensionality. Moreover we study a novel contraction operation on justifications and we introduce justifications with presumptions.

Keywords: justification logic, subset models, hyperintensionality, presumptions

1 Introduction

Justification logic is a refinement of modal logic with machinery for justification manipulation [7, 32]. Where modal logic uses a formula $\Box A$ to express that A *is known* or A *is provable*, justification logic uses a formula $t : A$ to express that A *is justified by reason* t or t *is a proof of* A. Justification logic supports operations on justifications. A prominent example is the application operator which represents modus ponens on the level of justifications: if t justifies A and s justifies $A \to B$, then $s \cdot t$ justifies B.

The first justification logic, the Logic of Proofs, has been introduced by Artemov [1, 2] to give a classical provability interpretation to $\mathsf{S4}$. Later it

College Publications, London 2021

turned out that the approach of representing justifications explicitly by terms is not only useful in proof theory but also in epistemic logic which led to the development of a great variety of different justification logics. There are justification counterparts of infinitely many modal logics other than S4 like Geach logics [7, 23, 25, 39], there are intuitionisitic and relevant justification logics [8, 35, 38], there are justification logics with common knoweldge [3, 19], there are dynamic epistemic justification logics [15, 18, 20, 21, 30, 37], there are fuzzy und probabilistic justification logics [24, 27, 28], and so on.

Several approaches to semantics for justification logic have been developed so far. Most of them interpret justification terms symbolically. In provability models [2, 31], terms are interpreted as (codes of) formal proofs in, e.g., Peano arithmetic. In order to obtain decidability results, Mkrtychev [36] introduced a class of models where terms are represented as sets of formulas. In Fitting models, there is a mapping from terms to sets of formulas for each possible world [22]. Also in modular models [5, 29], the logical type of a justification is a set of formulas.

Exceptions are [6, 13] where terms are modelled as sets of possible worlds. However, these papers do not consider the usual structure of justification terms. Furthermore, there are topological approaches to evidence [14, 16, 17] which, however, do not support an explicit representation of justificiations in their language.

Recently, Lehmann and Studer [33] introduced a different approach to semantics for justification logic. In their semantics, justification terms are represented as sets of possible worlds and operations on terms are modelled by operations on those sets of worlds. Then a formula $t : A$ holds if the interpretation of t, which is a set of worlds, is a subset of the set of worlds in which A holds. Hence the name *subset semantics*.

We first recall the basic definitions and results about subset models from [33]. In particular, we restate the two possibilities how subset models can deal with the application operation. Then, in Section 4, we give a detailed comparison between these two possibilities and show that they are essentially different. In Section 5, we discuss subset models in the context of hyperintensionality. Section 6 shows how subset models support a novel contraction operation on justification terms and Section 7 uses subset models to introduce justifications with presumptions. Finally we conclude the paper and discuss future work.

2 $\mathsf{L}^\star_{\mathsf{CS}}$-subset models

In this section we recall the basic syntax and semantics of subset models for justification logics and we mention some of its basic properties. This section is basically taken from [33] with additions from [34] where also all proofs can be found.

2.1 Syntax

Justification terms are built from countably many constants c_i and variables x_i and the special and unique constant $\mathsf{c}^\star$ according to the following grammar:

$$t ::= c_i \mid x_i \mid \mathsf{c}^\star \mid (t+t) \mid !t$$

The set of terms is denoted by Tm. The set of atomic terms, i.e. terms that do not contain any operator $+$ or ! is denoted by ATm. The operation $+$ is left-associative.

Formulas are built from countably many atomic propositions p_i, terms t and the symbol $\bot$ according to the following grammar:

$$F ::= p_i \mid \ \bot \ \mid F \to F \mid t : F$$

The set of atomic propositions is denoted by Prop and the set of all formulas is denoted by $\mathcal{L}_J$. The other classical Boolean connectives $\neg, \top, \wedge, \vee, \leftrightarrow$ are defined as usual.

Definition 1 ($\mathsf{c}^\star$**-term**). *A* $\mathsf{c}^\star$***-term*** *is defined inductively as follows:*

- $\mathsf{c}^\star$ *is a* $\mathsf{c}^\star$***-term***
- *if s and t are terms and c is a* $\mathsf{c}^\star$***-term*** *then $s+c$ and $c+t$ are* $\mathsf{c}^\star$***-terms***

So a $\mathsf{c}^\star$**-term** is either $\mathsf{c}^\star$ itself or a sum-term where $\mathsf{c}^\star$ occurs at least once.

We investigate a family of justification logics that differ in their axioms and how the axioms are justified. We have two sets of axioms, the first axioms are:

cl	all axioms of classical propositional logic;
jc*	$c : A \wedge c : (A \to B) \to c : B$ for all $\mathsf{c}^\star$**-terms** c;
j+	$s : A \vee t : A \to (s+t) : A$.

The set of these axioms is denoted by $\mathsf{L}^{\star}_{\alpha}$.
There is another set of axioms:

j4 $t : A \rightarrow !t : (t : A)$;
jd $t : \bot \rightarrow \bot$;
jt $t : A \rightarrow A$.

This set is denoted by $\mathsf{L}^{\star}_{\beta}$. It is easy to see that **jd** is a special case of **jt**. By $\mathsf{L}^{\star}$ we denote any logic that is composed from the whole set $\mathsf{L}^{\star}_{\alpha}$ and some subset of $\mathsf{L}^{\star}_{\beta}$. Moreover, a justification logic $\mathsf{L}^{\star}$ is defined by the set of axioms and its constant specification CS that determines which constant justifies which axiom. So the constant specification is a set

$$\mathsf{CS} \subseteq \{(c, A) \mid c \text{ is a constant and } A \text{ is an axiom of } \mathsf{L}^{\star}\}$$

If for a constant specification CS and a logic $\mathsf{L}^{\star}$ there exists for each axiom $A \in \mathsf{L}^{\star}$ a constant c s.t. $(c, A) \in \mathsf{CS}$ we say that CS is axiomatically appropriate w.r.t. $\mathsf{L}^{\star}$.

$\mathsf{L}^{\star}_{\mathsf{CS}}$ denotes the logic $\mathsf{L}^{\star}$ with the constant specification CS. To deduce formulas in $\mathsf{L}^{\star}_{\mathsf{CS}}$ we use a Hilbert system given by $\mathsf{L}^{\star}$ and the rules modus ponens:

$$\frac{A \qquad A \rightarrow B}{B}\ (\text{MP})$$

and axiom necessitation

$$\frac{}{\underbrace{!\ldots!}_{n}\, c : \underbrace{!\ldots!}_{n-1}\, c : \ \ldots : !!c : !c : c : A}\ (\text{AN!}) \quad \forall n \in \mathbb{N}, \text{ where } (c, A) \in \mathsf{CS}$$

2.2 Semantics

Definition 2 ($\mathsf{L}^{\star}_{\mathsf{CS}}$-subset models)**.** *Given some logic $\mathsf{L}^{\star}$ and some constant specification CS, then an $\mathsf{L}^{\star}_{\mathsf{CS}}$-subset model $\mathcal{M} = (W, W_0, V, E)$ is defined by:*

- *W is a set of objects called worlds.*
- *$W_0 \subseteq W$ and $W_0 \neq \emptyset$.*
- *$V : W \times \mathcal{L}_J \rightarrow \{0, 1\}$ such that for all $\omega \in W_0$, $t \in \mathsf{Tm}$, $F, G \in \mathcal{L}_J$:*

  - $V(\omega, \bot) = 0$;
  - $V(\omega, F \to G) = 1$ *iff* $V(\omega, F) = 0$ *or* $V(\omega, G) = 1$;
  - $V(\omega, t : F) = 1$ *iff* $E(\omega, t) \subseteq \{ \upsilon \in W \mid V(\upsilon, F) = 1 \}$.

- $E : W \times \mathsf{Tm} \to \mathcal{P}(W)$ *that meets the following conditions where we use*

$$[A] := \{\omega \in W \mid V(\omega, A) = 1\}. \tag{1}$$

For all $\omega \in W_0$, *and for all* $s, t \in \mathsf{Tm}$:

  - $E(\omega, s + t) \subseteq E(\omega, s) \cap E(\omega, t)$;
  - $E(\omega, \mathsf{c}^\star) \subseteq W_{MP}$ *where* W_{MP} *is the set of worlds that are deductively closed, see below;*
  - *if* ***jd*** $\in \mathsf{L}^\star$, *then* $\exists \upsilon \in W_0$ *with* $\upsilon \in E(\omega, t)$;
  - *if* ***jt*** $\in \mathsf{L}^\star$, *then* $\omega \in E(\omega, t)$;
  - *if* ***j4*** $\in \mathsf{L}^\star$, *then*

$$E(\omega, !t) \subseteq \{ \upsilon \in W \mid \forall F \in \mathcal{L}_J \ (V(\omega, t : F) = 1 \Rightarrow V(\upsilon, t : F) = 1) \} ;$$

  - *for all* $n \in \mathbb{N}$ *and for all* $(c, A) \in \mathsf{CS} : E(\omega, c) \subseteq [A]$ *and*

$$E(\omega, \underbrace{!\ldots!}_{n} c) \subseteq [\underbrace{!\ldots!}_{n-1} c : \ldots : !c : c : A].$$

The set W_{MP} *is formally defined as follows:*

$$W_{MP} := \{\omega \in W \mid \forall A, B \in \mathcal{L}_J \ ((V(\omega, A) = 1 \text{ and } V(\omega, A \to B) = 1) \text{ implies } V(\omega, B) = 1)\}.$$

So W_{MP} collects all the worlds where the valuation function is closed under modus ponens. W_0 is the set of *normal* worlds. The set $W \setminus W_0$ consists of the *non-normal* worlds. Moreover, using the notation introduced by (1), we can read the condition on V for formulas $t : F$ in normal worlds as:

$$V(\omega, t : F) = 1 \quad \text{iff} \quad E(\omega, t) \subseteq [F]$$

Since the valuation function V is defined on worlds and formulas, the definition of truth is pretty simple:

Definition 3 (Truth in $\mathsf{L}^\star_{\mathsf{CS}}$-subset models). *Let $\mathcal{M} = (W, W_0, V, E)$ be an $\mathsf{L}^\star_{\mathsf{CS}}$-subset model, $\omega \in W$ and $F \in \mathcal{L}_J$. We define the relation $\Vdash$ as follows:*

$$\mathcal{M}, \omega \Vdash F \quad \textit{iff} \quad V(\omega, F) = 1$$

Remark 4. *With the conditions on $E(\omega, \mathsf{c}^\star)$ and $E(\omega, s+t)$ we obtain the intended meaning of a $\mathsf{c}^\star$-**term** $s + \mathsf{c}^\star$, namely that we consider only deductively closed worlds of s. However, the set $E(\omega, s+\mathsf{c}^\star)$ does not have to be exactly the intersection of $E(\omega, s)$ with W_{MP} since we only have a subset-relation instead of equality. Hence $E(\omega, s+\mathsf{c}^\star) \neq E(\omega, \mathsf{c}^\star + s)$ in general. So even if in two $\mathsf{c}^\star$-**terms** the exactly same evidence sets occur, their order still matters. For the same reason $s + t : A \to t + s : A$ is not valid for any two distinct terms s and t.*

2.3 Soundness

Since non-normal worlds will not be sound even with respect to the axioms of classical logic, we only have soundness within W_0.

Lemma 5. *For any $\mathsf{L}^\star_{\mathsf{CS}}$-subset model $\mathcal{M} = (W, W_0, V, E)$, any $\omega \in W_0$ and any $\mathsf{c}^\star$-**term** c, we have that $E(\omega, c) \subseteq W_{MP}$.*

Proof. It is quite obvious and can be proven by induction on the structure of c.

- If $c = \mathsf{c}^\star$ the claim follows directly from the definition of E in W_0.
- Suppose $c = s + c'$ for a $\mathsf{c}^\star$-**term** c' and some term s.
  Then $E(\omega, c) \subseteq E(\omega, s) \cap E(\omega, c')$. By induction hypothesis we have $E(\omega, c') \subseteq W_{MP}$ and hence the claim follows. Case $c = c' + s$ is analogous. □

Theorem 6 (Soundness of $\mathsf{L}^\star_{\mathsf{CS}}$-subset models). *For any justification logic $\mathsf{L}^\star_{\mathsf{CS}}$ and any formula $F \in \mathcal{L}_J$:*

$$\mathsf{L}^\star_{\mathsf{CS}} \vdash F \quad \Rightarrow \quad \mathcal{M}, \omega \Vdash F \quad \textit{for all } \mathsf{L}^\star_{\mathsf{CS}}\textit{-subset models } \mathcal{M} \textit{ and all } \omega \in W_0$$

The **j**-axiom $s : (A \to B) \to (t : A \to s \cdot t : B)$ is not part of our logic. Using the $(\mathsf{c}^\star)$-axiom, we can define an application operation such that the $\mathbf{jc}^\star$-axiom is valid.

Definition 7 (Application). *We introduce a new abbreviation $\cdot$ on terms by:*

$$s \cdot t := s + t + \mathsf{c}^\star$$

Lemma 8 (The "**j**-axiom" follows). *For all $\mathcal{M} = (W, W_0, V, E)$, $\omega \in W_0$, $A, B \in \mathcal{L}_J$ and $s, t \in \mathsf{Tm}$:*

$$\mathcal{M}, \omega \Vdash s : (A \to B) \to (t : A \to s \cdot t : B)$$

2.4 Completeness

Since all presented logics are conservative extensions of classical propositional logic, we have the following consistency result.

Lemma 9 (Consistency of the logics). *All presented logics are consistent, that means that $\mathsf{L}^\star \nvdash \bot$ for all presented logics $\mathsf{L}^\star$.*

To prove completeness, we construct a canonical model for each logic $\mathsf{L}^\star_{\mathsf{CS}}$. A world in this model will simply be a set of formulas. If such a set is maximal $\mathsf{L}^\star_{\mathsf{CS}}$-consistent, then the corresponding world will be normal.

Definition 10 (Canonical Model). *For a given logic $\mathsf{L}^\star_{\mathsf{CS}}$ we define the canonical model $\mathcal{M}^C = (W^C, W^C_0, V^C, E^C)$ by:*

- $W^C = \mathcal{P}(\mathcal{L}_J)$;
- $W^C_0 = \left\{ \Gamma \in W^C \;\middle|\; \Gamma \text{ is maximal } \mathsf{L}^\star_{\mathsf{CS}}\text{-consistent set of formulas} \right\}$;
- $V^C : V^C(\Gamma, F) = 1 \quad \text{iff} \quad F \in \Gamma$;
- E^C : *With* $\Gamma/t := \{F \in \mathcal{L}_J \mid t : F \in \Gamma\}$ *and*

$$W^C_{MP} := \{\, \Gamma \in W^C \mid \forall A, B \in \mathcal{L}_J : \text{ if } A \to B \in \Gamma \text{ and } A \in \Gamma \text{ then } B \in \Gamma \,\}$$

we define :

$$E^C(\Gamma, t) = \left\{ \Delta \in W^C_{MP} \;\middle|\; \Delta \supseteq \Gamma/t \right\} \text{ if } t \text{ is a } \mathsf{c}^\star\textbf{-term};$$
$$E^C(\Gamma, t) = \left\{ \Delta \in W^C \;\middle|\; \Delta \supseteq \Gamma/t \right\} \text{ otherwise.}$$

Now we must show that the canonical model is indeed an $\mathsf{L}^\star_{\mathsf{CS}}$-subset model.

Lemma 11. *Given some logic* $\mathsf{L}^\star$ *and a constant specification* CS*, which is required to be axiomatically appropriate in case* (***jd***) $\in \mathsf{L}^\star$ *and* (***jt***) $\notin \mathsf{L}^\star$*, then the canonical model* $\mathcal{M}^C$ *is an* $\mathsf{L}^\star_{\mathsf{CS}}$*-subset model.*

Proof. In order to prove this, we have to show that $\mathcal{M}^C$ meets all the conditions we made for the valuation and evidence function and the constant specification i.e.:

1. $W_0^C \neq \emptyset$.
2. For all $\Gamma \in W_0^C$, $F, G \in \mathcal{L}_J$, and $t \in \mathsf{Tm}$:
   (a) $V^C(\Gamma, \bot) = 0$;
   (b) $V^C(\Gamma, F \to G) = 1 \quad$ iff $\quad V^C(\Gamma, F) = 0$ or $V^C(\Gamma, G) = 1$;
   (c) $V^C(\Gamma, t : F) = 1 \quad$ iff $\quad E^C(\Gamma, t) \subseteq [F]$.
3. For all $\Gamma \in W_0^C, F \in \mathcal{L}_J, s, t \in \mathsf{Tm}$:
   (a) $E^C(\Gamma, s + t) \subseteq E^C(\Gamma, s) \cap E^C(\Gamma, t)$;
   (b) $E^C(\Gamma, \mathsf{c}^\star) \subseteq W^C_{MP}$;
   (c) if **jd** in $\mathsf{L}^\star$: $\exists \Delta \in W_0^C$ s.t. $\Delta \in E^C(\Gamma, t)$;
   (d) if **jt** in $\mathsf{L}^\star$: $\Gamma \in E^C(\Gamma, t)$;
   (e) if **j4** in $\mathsf{L}^\star$:
   $$E^C(\Gamma, !t) \subseteq \Big\{\Delta \in W^C \mid \forall F \in \mathcal{L}_J \Big(V^C(\Gamma, t : F) = 1 \Rightarrow V^C(\Delta, t : F) = 1\Big)\Big\};$$
   (f) for all $(c, A) \in \mathsf{CS}$: $E^C(\Gamma, c) \subseteq [A]$ and
   $$E^C(\Gamma, \underbrace{!...!}_{n}\, c) \subseteq [\underbrace{!...!}_{n-1}\, c : .... : !c : c : A] \text{ for all } n \in \mathbb{N}.$$

So the proofs are here:

1. Since the empty set is proven to be $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent by Lemma 9 it can be extended by the Lindenbaum Lemma to a maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent set of formulas Γ with $\Gamma \in W_0^C$.

2. Suppose $\Gamma \in W_0^C$:

   (a) We claim $V^C(\Gamma, \bot) = 0$: Suppose the opposite, i.e. $V^C(\Gamma, \bot) = 1$ hence by the definition of V^C follows that $\bot \in \Gamma$. But this is a contradiction to the fact that Γ is consistent.

   (b) From left to right: Suppose $V^C(\Gamma, F \to G) = 1$, then by the definition of $V^C, F \to G \in \Gamma$. Since Γ is maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent this implies that $F \notin \Gamma$ or $G \in \Gamma$. Hence again by the definition of $V^C, V^C(\Gamma, F) = 0$ or $V^C(\Gamma, G) = 1$.
   From right to left: Suppose $V^C(\Gamma, F) = 0$ or $V^C(\Gamma, G) = 1$, then by the definition of V^C either $F \notin \Gamma$ or $G \in \Gamma$. Since $\Gamma \in W_0^C, \Gamma$ is maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent and hence in both cases $F \to G \in \Gamma$. But this means again by the definition of V^C that $V^C(\Gamma, F \to G) = 1$.

   (c) From left to right: Suppose $V^C(\Gamma, t : F) = 1$, then by Definition 10 $t : F \in \Gamma$. Hence with the definition of Γ/t we obtain $F \in \Gamma/t$. So for each $\Delta \in E^C(\Gamma, t), F \in \Delta$ (again by Definition 10). Hence for these Δ it follows by the definition of V^C that $V^C(\Delta, F) = 1$ and therefore $\Delta \in [F]$. Since this is true for all $\Delta \in E^C(\Gamma, t)$ we obtain $E^C(\Gamma, t) \subseteq [F]$.
   From right to left: The proof is by contraposition.
   Suppose $V^C(\Gamma, t : F) \neq 1$, by the definition of V^C $t : F \notin \Gamma$. We define a world Δ by $\Delta := \Gamma/t$. Since $\Delta \in \mathcal{P}(\mathcal{L}_J)$ we can be sure that Δ exists, i.e. $\Delta \in W^C$. Since $t : F \notin \Gamma$ it follows that $F \notin \Gamma/t$ and therefore $F \notin \Delta$. But obviously $\Delta \supseteq \Gamma/t$ hence $\Delta \in E^C(\Gamma, t)$ if t is not a $\mathsf{c}^{\star}$**-term**. So we conclude $E^C(\Gamma, t) \not\subseteq [F]$. It remains to show that in case t is a $\mathsf{c}^{\star}$**-term**, $\Delta := \Gamma/t \in W_{MP}^C$ since otherwise $\Delta \notin E^C(\Gamma, t)$. In fact this is the case. Since $\Gamma \in W_0^C$ we obtain that Γ is a maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent set of formulas and hence, whenever $t : A$, $t : (A \to B) \in \Gamma$ for a $\mathsf{c}^{\star}$**-term** t then by **jc*** we obtain $t : B \in \Gamma$. This means that whenever $A \in \Delta$ and $A \to B \in \Delta$ then $B \in \Delta$. Hence $\Delta = \Gamma/t$ is closed under modus ponens and therefore $\Delta \in W_{MP}^C$. So together with the former reasoning $\Delta \in E^C(\Gamma, t)$.

3. Suppose $\Gamma \in W_0^C$:

   (a) Given some $F \in \mathcal{L}_J, s, t \in \mathsf{Tm}$: To prove this, we start by an observation on the relation between the sets $\Gamma/(s+t)$ and Γ/s for $\Gamma \in W_0^C$. If $s : A \in \Gamma$ then since Γ is maximal $\mathsf{L}_{\mathsf{CS}}^{\star}$-consistent $s + t : A \in \Gamma$ hence $\Gamma/s \subseteq \Gamma/(s+t)$. With the same reasoning $\Gamma/t \subseteq \Gamma/(s+t)$. So if $\Delta \supseteq \Gamma/(s+t)$ then $\Delta \supseteq \Gamma/s$ and $\Delta \supseteq \Gamma/t$. Hence $E^C(\Gamma, s+t) \subseteq E^C(\Gamma, s)$ and $E^C(\Gamma, s+t) \subseteq E^C(\Gamma, t)$.[1] Therefore $E^C(\Gamma, s+t) \subseteq E^C(\Gamma, s) \cap E^C(\Gamma, t)$.

   (b) This follows directly from the fact that $\mathsf{c}^\star$ is a $\mathsf{c}^\star$**-term** and the definition of $E^C(\Gamma, t)$ for $\mathsf{c}^\star$**-terms**.

   (c) If **jd** in $\mathsf{L}^\star$, either CS is axiomatically appropriate or (**jt**) $\in \mathsf{L}^\star$ too.

      - CS is axiomatically appropriate.
        For any $\Gamma \in W_0^C$ we obtain $\neg(t :\bot) \in \Gamma$. Hence $\bot \notin \Gamma/t$. Suppose towards a contradiction that Γ/t is not $\mathsf{L}_{\mathsf{CS}}^{\star}$-consistent, i.e. there exist $A_1, \ldots, A_n \in \Gamma/t$ s.t.

        $$A_1, \ldots, A_n \vdash_{\mathsf{L}_{\mathsf{CS}}^{\star}} \bot \,. \qquad (2)$$

        This together with the construction of Γ/t leads to

        $$t : A_1, \ldots, t : A_n \in \Gamma.$$

        Since CS is axiomatically appropriate we can use (2) to infer $t : A_1, \ldots, t : A_n \vdash_{\mathsf{L}_{\mathsf{CS}}^{\star}} s(t) :\bot$, for some term $s(t)$ only based on t. Since Γ is assumed to be maximally consistent we can use (**jd**) and apply modus ponens to infer $\bot \in \Gamma$ which contradicts the assumption that Γ is consistent. Therefore Γ/t is $\mathsf{L}_{\mathsf{CS}}^{\star}$-consistent and can be expanded by the Lindenbaum Lemma to a maximal $\mathsf{L}_{\mathsf{CS}}^{\star}$-consistent set $\Delta \supseteq \Gamma/t$ with $\Delta \in W_0^C$ and $\Delta \in E^C(\Gamma, t)$.
      - (**jt**) $\in \mathsf{L}^\star$:
        The claim is a direct consequence of property (3d) (see next item).

[1] Please note if either s or t is a $\mathsf{c}^\star$**-term** this only holds due to $E^C(\Gamma, s+t)$ being constrained to W_{MP}^C by the fact that $s+t$ is a $\mathsf{c}^\star$**-term** too.

(d) Suppose for some $F \in \mathcal{L}_J, \Gamma \in W_0^C$ and $t \in \mathsf{Tm}$ that $F \in \Gamma/t$, i.e. $t : F \in \Gamma$, since Γ is maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent and $t : F \to F$ is an instance of the **jt**-axiom, we conclude that $F \in \Gamma$. Since F was arbitrary we obtain $\Gamma \supseteq \Gamma/t$ and hence $\Gamma \in E^C(\Gamma, t)$.

For the last step it is necessary to show that $\Gamma \in W_{MP}^C$ if t is a $\mathsf{c}^{\star}$**-term**. This claim follows directly from $\Gamma \in W_0^C$.

(e) Suppose for some Δ that $\Delta \in E^C(\Gamma, !t)$, hence $\Delta \supseteq \Gamma/!t$. Assume for some arbitrary $F \in \mathcal{L}_J$, $V^C(\Gamma, t : F) = 1$ i.e. by Definition 10 $t : F \in \Gamma$. Since $t : F \to !t : (t : F)$ is an instance of the **j4**-axiom and Γ is maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent we obtain $!t : (t : F) \in \Gamma$ and hence $t : F \in \Gamma/!t$. But then $t : F \in \Delta$ and by Definition 10 it follows that $V^C(\Delta, t : F) = 1$. Since F was an arbitrary formula and Δ an arbitrary world of $E^C(\Gamma, !t)$ we conclude that the condition holds.

(f) Suppose $(c, A) \in \mathsf{CS}$, then maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistency implies for all $\Gamma \in W_0^C$ that $c : A \in \Gamma$. Hence $A \in \Gamma/c$ and for all $\Delta \in E^C(\Gamma, c)$ we obtain $A \in \Delta$ and therefore $E^C(\Gamma, c) \subseteq [A]$.

Furthermore maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistency implies for all $\Gamma \in W_0^C$ by axiom necessitation that

$$\underbrace{!...!}_{n} c : ... :!c : c : A \in \Gamma.$$

Hence

$$\underbrace{!...!}_{n-1} c : ... :!c : c : A \in \Gamma/ \underbrace{!...!}_{n} c$$

and for all $\Delta \in E^C(\Gamma, \underbrace{!...!}_{n} c)$ we obtain

$$\underbrace{!...!}_{n-1} c : ... :!c : c : A \in \Delta$$

and therefore

$$E^C(\Gamma, \underbrace{!...!}_{n} c) \subseteq [\underbrace{!...!}_{n-1} c : ... :!c : c : A].$$

$\square$

Hence, given the mentioned conditions, the canonical model is an $\mathsf{L}^{\star}_{\mathsf{CS}}$-subset model and we are nearly done. The Truth Lemma follows very closely:

Lemma 12 (Truth Lemma)**.** *Let* $\mathcal{M}^C = (W^C, W^C_0, E^C, V^C)$ *be a canonical model. Then for any* $\Gamma \in W^C_0$*:*

$$\mathcal{M}^C, \Gamma \Vdash F \text{ if and only if } F \in \Gamma.$$

Hence each maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent set is represented by some world in the canonical model and thus completeness follows directly:

Theorem 13 (Completeness)**.** *Given some logic* $\mathsf{L}^{\star}$ *and a constant specification* CS*, which is required to be axiomatically appropriate in case* $(\boldsymbol{jd}) \in \mathsf{L}^{\star}$ *and* $(\boldsymbol{jt}) \notin \mathsf{L}^{\star}$*, then*

$$\mathcal{M}, \Gamma \Vdash F \quad \text{for all } \mathsf{L}^{\star}_{\mathsf{CS}}\text{-subset models } \mathcal{M} \text{ and for all } \Gamma \in W_0 \Longrightarrow \mathsf{L}^{\star}_{\mathsf{CS}} \vdash F.$$

Proof. The proof works with contraposition: Assume that $\mathsf{L}^{\star}_{\mathsf{CS}} \nvdash F$. Then $\{\neg F\}$ is $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent and by the Lindenbaum Lemma contained in some maximal $\mathsf{L}^{\star}_{\mathsf{CS}}$-consistent world Γ of the canonical model $\mathcal{M}^C$. Therefore $\mathcal{M}^C, \Gamma \nVdash F$. □

3 $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$-subset models

In this part we present an alternative definition of subset models for justification logic that directly interprets the application operator in the way traditional justification logic does. Hence we work with the standard language of justification logic and we consider the **j**-axiom instead of the axiom (**jc***). Again, this section is essentially taken from [33].

3.1 Syntax

In this section, justification terms are built from constants c_i and variables x_i according to the following grammar:

$$t ::= c_i \mid x_i \mid (t \cdot t) \mid (t + t) \mid !t$$

This set of terms is denoted by Tm^{A}. The operations $\cdot$ and $+$ are left-associative and ! binds stronger than anything else. Formulas are built from atomic propositions p_i, terms t and the symbol $\bot$ according to the following grammar:

$$F ::= p_i \mid \bot \mid F \to F \mid t : F$$

The set of formulas is denoted by $\mathcal{L}_J^{\mathsf{A}}$. Again we use the other logical connectives as abbreviations.

As in the first section, we investigate again a whole family of logics. They are arranged in two sets of axioms. The first set, denoted by $\mathsf{L}_\alpha^{\mathsf{A}}$ contains the following axioms:

cl	all axioms of classical propositional logic;
j	$s : (A \to B) \to (t : A \to s \cdot t : B)$;
j+	$s : A \lor t : A \to (s + t) : A$.

The other set of axioms is identical to $\mathsf{L}_\beta^\star$ (modulo the different language) and denoted by $\mathsf{L}_\beta^{\mathsf{A}}$. By L^{A} we denote any logic that is composed from the whole set $\mathsf{L}_\alpha^{\mathsf{A}}$ and some subset of $\mathsf{L}_\beta^{\mathsf{A}}$.

There are no differences between these logics and the ones of the former section except in case of application. Therefore we skip all the details already mentioned before.

CS and $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$ are defined as before except that the corresponding logic has changed as mentioned.

3.2 Semantics

Definition 14 ($\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$-subset models)**.** *Given some logic $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$, an $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$-subset model $\mathcal{M} = (W, W_0, V, E)$ is defined like an $\mathsf{L}_{\mathsf{CS}}^\star$-subset model where for*

$$E : W \times \mathsf{Tm}^{\mathsf{A}} \to \mathcal{P}(W)$$

the condition for $E(\omega, \mathsf{c}^\star)$ is replaced by the following condition for $\omega \in W_0$ and for terms of the form $s \cdot t$:

$$E(\omega, s \cdot t) \subseteq \mathfrak{W}_\omega(s, t)$$

where we use

$$\mathfrak{W}_\omega(s, t) := \{v \in W \mid \forall F \in \mathsf{APP}_\omega(s, t)(v \in [F])\}$$

with

$$\mathsf{APP}_\omega(s,t) := \{F \in \mathcal{L}_J^{\mathsf{A}} \mid \exists H \in \mathcal{L}_J^{\mathsf{A}} \text{ s.t. } E(\omega,s) \subseteq [H \to F] \text{ and } E(\omega,t) \subseteq [H]\}.$$

All other conditions are analogous to the ones of $\mathsf{L}_{\mathsf{CS}}^{\star}$*-subset models.*

The set $\mathsf{APP}_\omega(s,t)$ contains all formulas that are colloquially said derivable by applying modus ponens to a formula justified by s and a formula justified by t.

Truth in $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$-subset models is defined as before, i.e.

$$\mathcal{M},\omega \Vdash F \quad \text{iff} \quad V(\omega,F)=1.$$

Theorem 15 (Soundness of $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$-subset models)**.** *For any justification logic* L^{A}*, any constant specification* CS *and any formula* $F \in \mathcal{L}_J^{\mathsf{A}}$*:*

$$\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}} \vdash F \quad \Rightarrow \quad \mathcal{M},\omega \Vdash F \quad \text{for all } \mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}\text{-subset models } \mathcal{M} \text{ and all } \omega \in W_0.$$

Theorem 16 (Completeness of $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$-subset models)**.** *Given some logic* L^{A} *and a constant specification* CS*, which is required to be axiomatically appropriate in case* (***jd***) $\in \mathsf{L}^{\mathsf{A}}$ *and* (***jt***) $\notin \mathsf{L}^{\mathsf{A}}$*, then*

$$\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}} \vdash F \quad \Leftarrow \quad \mathcal{M},\omega \Vdash F \quad \text{for all } \mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}\text{-subset models } \mathcal{M} \text{ and all } \omega \in W_0.$$

The proof is analogous to the one of Theorem 13.

4 Comparing $\mathsf{L}_{\mathsf{CS}}^{\star}$ and $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$

In the previous chapters we introduced two different kinds of logics and semantics. Both of them use a subset relation to model justification. This leads to the question in which sense they differ. We start with presenting several lemmas that establish technical differences between the application operators used in $\mathsf{L}_{\mathsf{CS}}^{\star}$ and $\mathsf{L}_{\mathsf{CS}}^{\mathsf{A}}$. In the second part of this section, we discuss the conceptual difference between these two kinds of application.

Lemma 17 (Monotonicity of application in $\mathsf{L}_{\mathsf{CS}}^{\star}$)**.** *In* $\mathsf{L}_{\mathsf{CS}}^{\star}$ *the application operator is monotone, i.e.*

$$s : A \to s \cdot t : A, \text{ for all } s,t \in \mathsf{Tm}, A \in \mathcal{L}_J$$

Proof. This follows directly from axiom **j+** and Definition 7:

$$\begin{aligned} &\text{(axiom } \mathbf{j+}) && s : A \to s + t : A && (3)\\ &\text{(axiom } \mathbf{j+}) && s + t : A \to s + t + \mathsf{c}^\star : A && (4)\\ &\text{(Definition 7)} && s + t + \mathsf{c}^\star : A = s \cdot t : A && (5)\\ &\text{(3, 4, 5 and logical reasoning)} && s : A \to s \cdot t : A && \end{aligned}$$

□

In the corresponding semantics this fact holds because if $E(\omega, s) \subseteq [A]$, then any intersection of $E(\omega, s)$ with some other set will be a subset of $[A]$ as well.

This phenomenon illustrates the intended meaning of $\cdot$ in $\mathsf{c}^\star$-subset models: adding to a justification s the capacity of being able to apply modus ponens and combining other justifications with s does not reduce the power of the justification s to justify formulas.

Lemma 18 (Non-monotonicity of the application operator in $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$)**.** *In* $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$ *the application operator is not monotone.*

Proof. The proof is with a counterexample and by using soundness.

Given some atomic formula $A \neq \bot$ and atomic terms s and t. We consider the $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$-subset model $\mathcal{M} = (W, W_0, V, E)$ with $W = \{ \omega_1, \omega_2 \}$, $W_0 = \{ \omega_1 \}$, $V(\omega_1, A) = 1$ and for all other formulas the valuation is arbitrary but such that the conditions for V in W_0 are fulfilled, $V(\omega_2, X) = 0$ for all $X \in \mathcal{L}^{\mathsf{A}}_J$, $E(\omega_1, s) = \{ \omega_1 \}, E(\omega_1, t) = \{ \omega_2 \}, E(\omega_1, s \cdot t) = \{ \omega_2 \}$ and all other justifications are defined, s.t. they fulfil the conditions of E in worlds of W_0.

First we have to prove that $\mathcal{M}$ is an $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$-subset model, i.e. to show that all conditions made on V and E in Definition 14 for any $\omega \in W_0$ are fulfilled. For V this holds by definition of V in ω_1. So it remains to show that E behaves properly. Again, for all justifications except $E(\omega_1, s \cdot t)$ this holds by definition. So let us check whether

$$E(\omega_1, s \cdot t) \subseteq \mathfrak{W}_{\omega_1}(s, t). \tag{6}$$

Since $E(\omega_1, t) = \{\omega_2\}$, we obtain that the term t justifies nothing. Hence $\mathsf{APP}_{\omega_1}(s, t) = \emptyset$. Therefore we obtain that $\mathfrak{W}_{\omega_1}(s, t) = W$ and (6) is obvious.

Further we have $\mathcal{M}, \omega_1 \Vdash s : A$ but since $V(\omega_2, A) = 0$, there is

$$\{ \omega_2 \} = E(\omega_1, s \cdot t) \not\subseteq [A]$$

and hence $\mathcal{M}, \omega_1 \not\Vdash s \cdot t : A$.

Finally, with Theorem 15 we conclude that $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}} \not\vdash s : A \to s \cdot t : A$ □

So the meaning of the application operator in $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$ is different to the one in $\mathsf{L}^{\star}_{\mathsf{CS}}$. In $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$, a term of the form $s \cdot t$ only justifies formulas that can be obtained by modus ponens; whereas in $\mathsf{L}^{\star}_{\mathsf{CS}}$, the term $s \cdot t$ also justifies formulas that are already justified by its subterms s and t.

Another difference between the application in $\mathsf{L}^{\star}_{\mathsf{CS}}$-subset models and models for standard justification logic is that application does ignore which justification justifies the condition and which justifies the antecedent.

Lemma 19. *For all $\mathsf{L}^{\star}_{\mathsf{CS}}$-subset models $\mathcal{M} = (W, W_0, V, E)$ and all $\omega \in W_0$, $A, B \in \mathcal{L}_J$ and $s, t \in \mathsf{Tm}$, we have*

$$\mathcal{M}, \omega \Vdash s : (A \to B) \to (t : A \to t \cdot s : B).$$

Proof. The proof is analogous to the one of Lemma 8. □

However, application is not commutative.

Lemma 20. *The formula $s \cdot t : A \to t \cdot s : A$ is not valid in $\mathsf{L}^{\star}_{\mathsf{CS}}$-subset models.*

Proof. Since the evidence set of a sum-term $s + t$ is only required to be *some* subset of the intersection of $E(\omega, s)$ and $E(\omega, t)$, it is possible that $E(\omega, s \cdot t) \neq E(\omega, t \cdot s)$ and hence it is possible that only one of them is a subset of $[A]$. □

The previous lemmas show some technical differences between application in $\mathsf{L}^{\star}_{\mathsf{CS}}$ and $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$. There is, however, also an important conceptual difference concerning application between the two logics as the following remark shows.

Remark 21. *Consider the set of formulas*

$$\Gamma := \{t : A,\ s : (A \to B),\ s : (B \to C)\}.$$

Of course, we can derive

$$s \cdot (s \cdot t) : C \tag{7}$$

from Γ in both $\mathsf{L}^{\star}_{\mathsf{CS}}$ and $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$.

In $\mathsf{L}^{\star}_{\mathsf{CS}}$ *application is defined as an abbreviation according to Definition 7. Hence the term* $s \cdot (s \cdot t)$ *is an abbreviation for* $s + (s + t + \mathsf{c}^{\star}) + \mathsf{c}^{\star}$.

The derivation in $\mathsf{L}^{\star}_{\mathsf{CS}}$ *therefore works as follows, where we use* CR *as an abbreviation for classical reasoning:*

$s : (A \to B) \to s \cdot (s \cdot t) : (A \to B)$	*2 times* ***j+***
$t : A \to s \cdot (s \cdot t) : A$	*4 times* ***j+***
$s \cdot (s \cdot t) : (A \to B) \wedge s \cdot (s \cdot t) : A \to s \cdot (s \cdot t) : B$	***jc****
$s \cdot (s \cdot t) : B$	CR
$s : (B \to C) \to s \cdot (s \cdot t) : (B \to C)$	*2 times* ***j+***
$s \cdot (s \cdot t) : (B \to C) \wedge s \cdot (s \cdot t) : B \to s \cdot (s \cdot t) : C$	***jc****
$s \cdot (s \cdot t) : C$	CR

In $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$*:*

$s : (A \to B) \to (t : A \to s \cdot t : B)$	***j***
$s \cdot t : B$	CR
$s : (B \to C) \to (s \cdot t : B \to s \cdot (s \cdot t) : C)$	***j***
$s \cdot (s \cdot t) : C$	CR

However, for $\mathsf{L}^{\star}_{\mathsf{CS}}$ *we also find that*

$$\Gamma \vdash_{\mathsf{L}^{\star}_{\mathsf{CS}}} s \cdot t : C. \tag{8}$$

The proof is similar to the above $\mathsf{L}^{\star}_{\mathsf{CS}}$ *derivation but we use* $s \cdot t = s + t + \mathsf{c}^{\star}$ *instead of* $s \cdot (s \cdot t)$*. This derivation of* $s \cdot t : C$ *from* Γ *is neither possible in* $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$ *nor in traditional justification logics.*

The traditional application operation (and the one in $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$) corresponds to *one* application of modus ponens. Indeed, in order to derive C from A, $A \to B$, and $B \to C$ we need two applications of modus ponens, which results in the two occurrences of $\cdot$ in (7). Application in $\mathsf{L}^{\star}_{\mathsf{CS}}$, on the other hand, corresponds to *arbitrarily many* applications of modus ponens. Thus, the occurrence of $\cdot$ in (8) means that modus ponens has been applied once or several times in the derivation of C.[2]

[2] To be precise, also the case when modus ponens was not applied at all is included as Lemma 17 shows.

Maybe this difference is best explained in terms of a naive proof theoretic semantics where evidence terms represent multi-conclusion Hilbert-style proofs, i.e. sequences of formulas where each formula either is an axiom or follows by a rule application from formulas occurring earlier in the sequence. Then a formula $s : F$ holds if F is a formula occurring (anywhere) in the proof represented by s. In this proof-theoretic setting, the traditional application operation can be modeled by concatenating two proofs and then closing the result under one iteration of applying modus ponens, i.e. if A and $A \to B$ are in the result of the concatenation, then add B. The application of $\mathsf{L}^{\star}_{\mathsf{CS}}$ in contrast, can be modeled by concatenating two proofs and then taking the deductive closure of the result, i.e. applying modus ponens iteratively until a fixed point is reached.

Now this has some important consequences when it comes to applications. Traditional (and $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$) application terms (where one occurrence of $\cdot$ corresponds to one step of modus ponens) properly reflect an agent's reasoning process. Hence if a term s justifies an agent's belief in A, then the size of s is an adequate measure for the complexity of the agent's reasoning that led to believing A. Making use of this feature, Artemov and Kuznets [9, 10, 11] show that justification logic (as epistemic logic) avoids the logical omniscience problem. Since application in $\mathsf{L}^{\star}_{\mathsf{CS}}$ corresponds to deductive closure, the correspondence between the size of a term and the complexity of the reasoning process is lost and hence the logical omniscience problem will come back in $\mathsf{L}^{\star}_{\mathsf{CS}}$.

Another important application of justification logic is evidence tracking, used, e.g., in the analysis of the Red Barn example in [4]. For an analysis of this kind, the complexity of the reasoning process does not matter. Hence this (and similar) examples can be properly modeled not only in $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$ but also in $\mathsf{L}^{\star}_{\mathsf{CS}}$.

We summarize this comparision as follows. Application in $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$ corresponds to one step of internalized modus ponens (the traditional application operation in justification logic) whereas application in $\mathsf{L}^{\star}_{\mathsf{CS}}$ corresponds to taking the deductive closure. Thus we can say that $\mathsf{L}^{\mathsf{A}}_{\mathsf{CS}}$-application models small-step reasoning of an agent and $\mathsf{L}^{\star}_{\mathsf{CS}}$-application represents a big-step approach like combining datasets with deductive closure. The first is needed if complexities matter (like in the logical omniscience case); the latter is sufficient if one only has to keep track of what beliefs depend on (like in the Red Barn example).

5 Hyperintensionality

As mentioned by Artemov and Fitting [7] hyperintensionality is a key aspect of justification logic that makes a difference to standard modal logic. In modal logic, if A is logically equivalent to B, then $\Box A$ implies $\Box B$. This is not the case in justification logic where $A \leftrightarrow B$ and $s : A$ do not imply $s : B$. So justification logic is able to distinguish between logically equivalent contents. This is of importance when you think about propositions like '$0 = 0$' and Fermat's Last Theorem. Both have the same content but if some proof s is a justification for '$0 = 0$' it does not have to be a justification for Fermat's Last Theorem as well.

In the standard semantics of justification logic, where the interpretation of a justification term is a sequence or a set of formulas, hyperintensionality comes for free. In subset sematics this is not the case. Consider the formulas P, $P \vee P$, $P \wedge P$ and $(P \to \bot) \to \bot$. In a normal world all these formulas have the same truth value and if we have no impossible worlds at hand, each term that justifies one of them, will as well justify all the others. But in a hyperintensional context, like *the agent believes that...*, we have to be able to distinguish them.

A natural suggestion to achieve that is to introduce impossible worlds. In impossible worlds, axioms and tautologies do not have to evaluate to true, they might do so in some worlds but not in others. In an impossible world two equivalent propositions may be evaluated to different truth values, and hence a justification that contains such a world may support one proposition but not the other. The importance of impossible worlds for modelling hyperintensionality is worked out in detail by Jago [26].

6 Contraction

In belief revision contraction refers to the operation of 'removing' a sentence from a belief set. Justification logic is normally used to model how we obtain new justifications by combining older ones. In this sense it is quite strange to have a section on contraction, which, in fact, is a process where we lose information. Nevertheless, we tried to find a way of modelling contraction within models of justification logic.

We model contraction by the following two principles:

$$s : A \;\rightarrow\; s^{-B} : A \quad \text{for } A \neq B, \tag{9}$$

$$\neg(s^{-B} : B), \tag{10}$$

where s^{-B} denotes the justification s that loses the capacity to justify B. In other words: if justification s has the capacity to justify B, then the set of worlds within the interpretation of s is changed in s^{-B} such that it does no longer justify B. So (9) guarantees that s^{-B} only loses its power to justify B but apart from this, the justification keeps its power to justify all other formulas that s justifies. And (10) guarantees that B no longer is justified by s^{-B}.

To model these new features within subset models we have to adapt our syntax and semantics.

Definition 22 (The language $\mathcal{L}_J^C$)**.** *The language $\mathcal{L}_J^C$ is composed from terms and formulas such that:*

- *c_i and x_i are atomic terms for constants c_i and variables x_i. All atomic terms are terms.*
- *If s, t are terms and B is a formula, then $s + t$, $s \cdot t$, $!t$ and t^{-B} are terms too.*
- *$\bot$ and all atomic propositions are formulas.*
- *if t is a term and F, G are formulas, then $F \rightarrow G$ and $t : F$ are formulas too.*

We use Tm^C for the set of terms of the language $\mathcal{L}_J^C$.

Definition 23 (The logic L^C)**.** *Given a logic L^A we define the logic L^C by adding the following two new axioms to L^A:*

C1 $\quad s : A \;\rightarrow\; s^{-B} : A \quad$ *for $A \neq B$*
C2 $\quad \neg(s^{-B} : B)$

Definition 24 (Non-B world)**.** *Given a model $\mathcal{M} = (W, W_0, E, V)$ and a formula B, we say $\omega \in W$ is a* non-B world *if $V(\omega, B) = 0$. Further we say $\omega \in W$ is a* maximal non-B world *if*

$$V(\omega, A) = \begin{cases} 0 & \textit{if } A = B \\ 1 & \textit{otherwise.} \end{cases}$$

Obviously for any $B \in \mathcal{L}_J^C$, a maximal non-B world ω is not consistent with classical logic. Since we allow impossible worlds in our models, this is not a problem. In general, there is not a unique non-B world for some formula B, since worlds may have the same valuations but differ in their evidence function.

Definition 25 (L^C-subset model)**.** *A model $\mathcal{M} = (W, W_0, E, V)$ is called an $\mathcal{L}_J^C$-subset model if W, W_0, E, V are defined analogously to Definition 14 and*

- *for each formula B, there exists at least one maximal non-B world. For each formula B we pick one such maximal non-B-world and denote it with $\omega_{\overline{B}}$.*
- *E additionally satisfies for all $\omega \in W_0$ and all $s^{-B} \in \mathsf{Tm}^\mathsf{C}$:*

$$E(\omega, s^{-B}) = E(\omega, s) \cup \{\omega_{\overline{B}}\}.$$

Truth in an L^C-subset models is defined as before, i.e.

$$\mathcal{M}, \omega \Vdash F \quad \text{iff} \quad V(\omega, F) = 1.$$

Theorem 26 (Soundness)**.** *Given a logic L^C and a formula $F \in \mathcal{L}_J^C$*

$$\mathsf{L}^\mathsf{C} \vdash F \Rightarrow \mathcal{M}, \omega \Vdash F \text{ for all } \omega \in W_0 \text{ in all } \mathsf{L}^\mathsf{C}\text{-subset models } \mathcal{M}$$

Proof. The proof is analogous to the proof of Theorem 15. We only show the cases for the additional axioms.

- If F is an instance of **C1** then $F = s : A \to s^{-B} : A$ for $A \neq B$. Suppose $\mathcal{M}, \omega \Vdash s : A$ i.e. $E(\omega, s) \subseteq [A]$ for $\omega \in W_0$. Since $A \neq B$, we know $\omega_{\overline{B}} \in [A]$ and therefore $E(\omega, s^{-B}) = E(\omega, s) \cup \{\omega_{\overline{B}}\} \subseteq [A]$. We conclude $\mathcal{M}, \omega \Vdash s^{-B} : A$.
- If F is an instance of **C2** then $F = \neg(s^{-B} : B)$ for some $B \in \mathcal{L}_J^C$. Since $\omega_{\overline{B}} \notin [B]$, it is obvious that $E(\omega, s^{-B}) = E(\omega, s) \cup \{\omega_{\overline{B}}\} \not\subseteq [B]$. Hence $\mathcal{M}, \omega \nVdash s^{-B} : B$. □

Remember that we work in a hyperintensional context. Applying the contraction operator $\cdot^{-B}$ to a term s only removes B from the formulas justified by s. In particular, we may have that $s^{-B} : (A \wedge B)$ is true although

$s^{-B} : B$ must be false. This could be addressed by introducing some kind of selection function that chooses a (non-maximal) non-B set that satisfies certain closure conditions to define the interpretation of $\cdot^{-B}$. Of course then axiom **C1** needs to be changed accordingly.

Obviously, this part on contraction is very preliminary and a lot of work needs to be done in this context. In particular, we do not yet have a completeness result.

7 Justification with presumption

When we look around and explore the world, in every second we get new evidence that makes us understand the world we are living in. But usually we are not that open to new evidence that we consider all worlds as possible that do not contradict that specific evidence. In fact, we have some presumption about the world, of which we cannot always give explicit justifications, but which we simply added to our belief system at some point in life. We interpret new evidence in context to these presumptions. So we often reduce the set of worlds we consider to be the actual world in an evidence to those worlds that are consistent with our presumptions.

In standard justification logic this is usually not taken into account. With the constant specification we have some very special kind of presumptions, but only about the axioms we believe in. So far, there is no possibility to model that we believe B without indicating the explicit reason why we do so.[3]

The aim of justification logic with presumption is to model an agent's reasoning where not all of her beliefs are explicitly justified. We do this by allowing justification terms t_Γ where t is a usual term that stands for some evidence and Γ is a set of formulas that are believed without giving some explicit reason for believing them.

Definition 27 (The language $\mathcal{L}_J^P$)**.** *The language $\mathcal{L}_J^P$ is composed from terms and formulas such that*

- $0, 1, c_i, x_i$ *are atomic terms for constants c_i, variables x_i and the unique constants* $0, 1$*. All atomic terms are terms.*

[3]There are hybrid justification logics that feature both implicit and explicit knowledge [12]. There, however, the presumptions cannot be reflected on the level of terms.

- *If s, t are terms and Γ is a set of formulas, then $s + t$, $s \cdot t$, $!t$ and t_Γ are terms too.*
- *$\bot$ and all atomic propositions are formulas.*
- *if t is a term and F, G are formulas, then $F \to G$ and $t : F$ are formulas too.*

We use the abbreviation t_A to denote the term $t_{\{A\}}$.

Definition 28 (The logic L^{P})**.** *Given any logic L^{A} we define the logic L^{P} by adding the following axioms to L^{A}:*

P1	$\neg(0 : A)$	*for all $A \in \mathcal{L}_J^P$*
P2	$t_\Gamma : A$	*for all $A \in \Gamma$*
P3	$t : A \to t_\Gamma : A$	
P4	$t_A : B \to t + 1 : (A \to B)$	

and the new rule

$$\frac{A}{1 : A}\ \textit{(1-Nec)}$$

The idea behind **P1** is that 0 is like a blueprint of an evidence such that we can model the presumption without referring to a more detailed justification. So 0_A then is *the* evidence that A is true.

P2 claims that if we restrict some evidence to the worlds, where all our presumptions hold, then this is a justification that each of them holds.

The idea behind **P3** is that if we have a justification for something, restricting this justification to the worlds that correspond with our presumptions will not reduce its power to justify a specific formula. So adding information leads to a monotone process. This may look a bit strange, when we consider instances like $t : (\neg A) \to t_A : (\neg A)$ but in fact, this case just illustrates that we deal on one hand with justification that may justify untruthful formulas (as long as **jt** is not in our logic) and on the other hand with presumptions that may be wrong.

P4 relates the new type of justifications to the standard ones. If, under the condition that A is believed, t justifies B, then $t + 1$ justifies that A implies B without condition. Note that this only holds from left to right. In Lemma 31 we will show that the converse direction is not valid.

The justification 1 has a similar meaning as the justification $\mathsf{c}^\star$ in $\mathsf{c}^\star$-subset models; but instead of focusing on the deductively closed worlds, it focuses on the normal worlds. The rule **1-Nec** states that 1 is *the* justification that provable formulas hold. We have already used a similar justification for probabilistic evidence logic, which was introduced by Artemov [6] and adapted for subset models in [33].

Definition 29 (L^P-subset model). *A model $\mathcal{M} = (W, W_0, E, V)$ is called an L^P-subset model, if W is a set of worlds that contains a particular world $\omega_\emptyset$, W_0, V and E are defined analogously to Definition 14 with the following condition added on V and E for all $\omega \in W_0$, terms t, t_Γ of $\mathcal{L}_J^P$, and $A \in \mathcal{L}_J^P$:*

- $V(\omega_\emptyset, F) = 0, \quad \forall F \in \mathcal{L}_J^P$;
- $E(\omega, t_\Gamma) = (\bigcap_{A\in\Gamma}[A]) \cap E(\omega, t)$;
- $E(\omega, 0) = W$;
- $E(\omega, 1) = W_0$.

This new world $\omega_\emptyset$ is of course not an element of W_0 and models a world where nothing at all is true. We have to add it in order to be sure that 0 by itself does not justify anything.

Truth in an L^P-subset models is defined as before, i.e.

$$\mathcal{M}, \omega \Vdash F \quad \text{iff} \quad V(\omega, F) = 1.$$

Theorem 30 (Soundness). *Given a logic L^P and a formula F*

$$\mathsf{L}^\mathsf{P} \vdash F \quad \Rightarrow \quad \mathcal{M}, \omega \Vdash F \text{ for all } \omega \in W_0 \text{ of all } \mathsf{L}^\mathsf{P}\text{-subset models } \mathcal{M}$$

Proof. The proof is again by induction on the length of the derivation and analogous to the proof of Theorem 15. Since L^P-subset models only differ in the aspects of these new axioms, we only show this part of the proof here.

- If F is an instance of **P1** then $F = \neg(0 : A)$ for some formula A. Since $\omega_\emptyset \in W$ we obtain that $[A] \subsetneq W$ and hence $W = E(\omega, 0) \not\subseteq [A]$ for all formulas A and $\omega \in W_0$. So $V(\omega, 0 : A) = 0$ and hence we have $\mathcal{M}, \omega \nVdash 0 : A$ and therefore $\mathcal{M}, \omega \Vdash \neg(0 : A)$.

- If F is an instance of **P2** then $F = t_\Gamma : A$ for some justification t, some set of formulas Γ and some $A \in \Gamma$.
  Since $A \in \Gamma$ we obtain $\bigcap_{B \in \Gamma}[B] \subseteq [A]$ and then any further intersection on the left side is of course as well a subset of $[A]$. Therefore
  $$E(\omega, t_\Gamma) = \left(\bigcap_{B \in \Gamma} [B] \right) \cap E(\omega, t) \subseteq [A].$$
  Hence $V(\omega, t_\Gamma : A) = 1$ and finally $\mathcal{M}, \omega \Vdash t_\Gamma : A$.

- If F is an instance of **P3** then $F = t : A \to t_\Gamma : A$ for some term t, some $A \in \mathcal{L}_J^P$ and some set of formulas Γ.
  Suppose $\mathcal{M}, \omega \Vdash t : A$ then $E(\omega, t) \subseteq [A]$. Since
  $$E(\omega, t_\Gamma) = \left(\bigcap_{B \in \Gamma} [B] \right) \cap E(\omega, t) \subseteq E(\omega, t) \subseteq [A]$$
  we obtain $V(\omega, t_\Gamma : A) = 1$ and conclude $\mathcal{M}, \omega \Vdash t_\Gamma : A$.

- If F is an instance of **P4** then $F = t_A : B \to t + 1 : (A \to B)$ for some term t and formulas A, B.
  Suppose $\mathcal{M}, \omega \Vdash t_A : B$ then $V(\omega, t_A : B) = 1$ and therefore
  $$E(\omega, t) \cap [A] \subseteq [B] \tag{11}$$
  Take an arbitrary $v \in E(\omega, t + 1) \subseteq E(\omega, t) \cap W_0$. We have either $v \in [A]$ or $v \notin [A]$.
  - If $v \in [A]$ we obtain by $v \in E(w, t)$ and (11) that $v \in [B]$. The conditions on W_0 further allow us to conclude from $V(v, B) = 1$ that $V(v, A \to B) = 1$ and hence $v \in [A \to B]$.
  - If $v \notin [A]$ then we can directly deduce from $V(v, A) = 0$ and $v \in W_0$ that $V(v, A \to B) = 1$ and hence $v \in [A \to B]$.

  So both $v \in [A]$ and $v \notin [A]$ imply $v \in [A \to B]$. Therefore we find $E(\omega, t + 1) \subseteq [A \to B]$ and conclude $\mathcal{M}, \omega \Vdash t + 1 : (A \to B)$.

- If F is derived by **1-Nec** then $F = 1 : A$ for some L^P-derivable formula A. By induction hypothesis we then can assume that A is valid and hence $V(\omega, A) = 1$ for all $\omega \in W_0$. Therefore we have that $W_0 \subseteq [A]$ and hence $E(\omega, 1) = W_0 \subseteq [A]$ which means that $\mathcal{M}, \omega \Vdash 1 : A$. $\square$

The next lemma establishes that the converse direction of **P4** is not valid.

Lemma 31. *Let A and B be formulas and t be a term of $\mathcal{L}_J^P$. The formula*

$$t + 1 : (A \to B) \to t_A : B$$

is not valid in L^P.

Proof. To prove this we use a countermodel $\mathcal{M} = (W, W_0, V, E)$ where we define $W = \{\omega_1, \omega_2\}$, $W_0 = \{\omega_1\}$, V such that $[A] = W$, $[B] = \{\omega_1\}$ and $[A \to B] = \{\omega_1\}$ and $E(\omega_1, t) = W$. For all the other formulas and justifications the model is defined in a way that fulfils all conditions to be an L^P-subset model.

In ω_1 we have that $t + 1 : (A \to B)$ since

$$E(\omega_1, t + 1) \subseteq E(\omega_1, t) \cap E(\omega_1, 1) = W \cap W_0 \subseteq [A \to B].$$

The last step is because ω_1 is the only world in W_0 and $\omega_1 \in [A \to B]$. Therefore $\mathcal{M}, \omega_1 \Vdash t + 1 : (A \to B)$. But by Definition 29 we have that $E(\omega_1, t_A) = [A] \cap E(\omega_1, t) = W \cap W = W$ and $W \not\subseteq [B]$ since $\omega_2 \notin [B]$. $\square$

So far, we have no completeness result for L^P and we believe that further axioms are needed to obtain a system that is complete. We leave this to future work.

8 Conclusion

We discussed a novel semantics for justification logics, called subset semantics. Whereas most semantics interpret justifcations symbolically as sets of formulas, subset semantics interprets justifications as sets of possible worlds.

This provides a versatile mechanism to handle justifications. We have shown how to deal with hyperintensionality and we have presented a novel contraction operator on the level of justification terms. Furthermore, using subset models, we could introduce a framework for justifications with presumptions. These applications, however, are all in a preliminary stage and need further exploration. In particular, the question of completeness is open for L^C and L^P.

In future work, we also plan to investigate the full power of subset models in the context of belief change. Subset models also provide a promising approach to probabilistic justifications. Since terms are modelled as sets of possible worlds, we can apply a probability measure to get probabilistic justifications in a natural way.

Moreover, we plan to study Jago's [26] approach of ordering the worlds from possible ones over epistemically possible ones to impossible ones in the framework of subset models.

Another line of future work is to study the relation of subset semantics to Kripke semantics. We can treat each term t as a (non-normal) modality $\Box_t$ with its acessibility relation given by wR_tv iff $v \in E(w, t)$. Then the truth definition for normal worlds in subset models coincides with the Kripkean notion of truth of the modality $\Box_t$ with the accessibility relation R_t, i.e. $t : F$ is true at w if F is true at all v with wR_tv. The additional conditions for **jd** and **jt** in subset models then translate to the well-known seriality (w.r.t. normal worlds) and reflexivity conditions on R_t. The condition for **j4**, however, is more general due to the presence of non-normal worlds. It is an open question whether taking the union of the accessibility relations R_t over all terms yields a standard accessibility relation (corresponding to the understanding of $\Box$ as an existential quantifier over terms). A follow-up question then is whether one can retrieve a realization procedure from this connection of subset models and Kripke models.

Acknowledgements

We would like to thank Igor Sedlár for bringing our attention to the monotonicity property established in Lemma 17. We also would like to thank Sergei Artemov for asking about hyperintensionality in subset models. And finally we would like to thank the anonymous referee who found a mistake in the original completeness proof for $\mathsf{L}^{\star}_{\mathsf{CS}}$ and who did a great job giving us very helpful and detailed feedback with many highly valuable comments. They substantially improved the paper.

This work was supported by the Swiss National Science Foundation grant 200021_165549.

References

[1] Sergei Artemov *Operational modal logic* Technical Report MSI 95–29, Cornell University, December 1995.

[2] Sergei Artemov *Explicit Provability and Constructive Semantics,* Bulletin of Symbolic Logic, 7 (1), March 2001, pp. 1–36.

[3] Sergei Artemov *Justified Common Knowledge,* TCS, 357 (1–3), 2006, pp. 4–22.

[4] Sergei Artemov *The Logic of Justification,* The Review of Symbolic Logic, 1 (4), 2008, pp. 477–513.

[5] Sergei Artemov *The Ontology of Justifications in the Logical Setting,* Studia Logica, 100 (1–2), 2012, pp. 17–30.

[6] Sergei Artemov *On Aggregating Probabilistic Evidence,* in Sergei Artemov, Anil Nerode (eds.), LFCS 2016, pp. 27–42 Springer, 2016.

[7] Sergei Artemov, Melvin Fitting Justification Logic: Reasoning with Reasons, Cambridge University Press, Cambridge UK 2019.

[8] Sergei Artemov, Rosalie Iemhoff *The Basic Intuitionistic Logic of Proofs,* JSL, 72 (2), 2007, pp. 439–451.

[9] Sergei Artemov, Roman Kuznets *Logical omniscience via proof complexity,* In Zoltán Ésik (ed.), Computer Science Logic, 20th International Workshop, CSL 2006, 15th Annual Conference of the EACSL, Szeged, Hungary, September 25–29, 2006, Proceedings, pp. 135–149, Springer, 2006.

[10] Sergei Artemov, Roman Kuznets *Logical omniscience as a computational complexity problem* In Aviad Heifetz (ed.), Theoretical Aspects of Rationality and Knowledge, Proceedings of the Twelfth Conference (TARK 2009), pp. 14–23, ACM, 2009.

[11] Sergei Artemov, Roman Kuznets *Logical omniscience as infeasibility* Annals of Pure and Applied Logic, 165 (1), 2014, pp. 6–25.

[12] Sergei Artemov, Elena Nogina *Introducing Justification into Epistemic Logic,* Journal of Logic and Computation, 15 (6), 2005, pp. 1059–1073.

[13] Sergei Artemov, Elena Nogina *Topological Semantics of Justification Logic,* in Edward A. Hirsch, Alexander A. Razborov, Alexei Semenov, Anatol Slissenko (eds.), CSR 2008,, pp. 30–39 Springer, 2008.

[14] Alexandru Baltag, Nick Bezhanishvili, Aybüke Özgün, Sonja Smets *Justified Belief and the Topology of Evidence,* in WoLLIC, 2016.

[15] Alexandru Baltag, Bryan Renne, Sonja Smets *The Logic of Justified Belief Change, Soft Evidence and Defeasible Knowledge,* in Luke Ong, Ruy

de Queiroz (eds.), WoLLIC 2012, volume 7456 of *LNCS*, pp. 168–190 Springer, 2012.

[16] Johan van Benthem, David Fernández Duque, Eric Pacuit *Evidence Logic: A New Look at Neighborhood Structures,* in Advances in Modal Logic, 2012.

[17] Johan van Benthem, David Fernández Duque, Eric Pacuit *Evidence and plausibility in neighborhood structures,* CoRR, abs/1307.1277, 2014.

[18] Samuel Bucheli, Roman Kuznets, Bryan Renne, Joshua Sack, Thomas Studer *Justified Belief Change,* in Xabier Arrazola, María Ponte (eds.), LogKCA-10, pp. 135–155 University of the Basque Country Press, 2010.

[19] Samuel Bucheli, Roman Kuznets, Thomas Studer *Justifications for common knowledge,* Applied Non-Classical Logics, 21 (1), January–March 2011, pp. 35–60.

[20] Samuel Bucheli, Roman Kuznets, Thomas Studer *Partial Realization in Dynamic Justification Logic,* in Lev D. Beklemishev, Ruy de Queiroz (eds.), WoLLIC 2011, volume 6642 of *LNAI*, pp. 35–51 Springer, 2011.

[21] Samuel Bucheli, Roman Kuznets, Thomas Studer *Realizing public announcements by justifications,* Journal of Computer and System Sciences, 80 (6), 2014, pp. 1046–1066.

[22] Melvin Fitting *The logic of proofs, semantically,* APAL, 132 (1), 2005, pp. 1–25.

[23] Melvin Fitting *Modal logics, justification logics, and realization,* APAL, 167 (8), 2016, pp. 615–648.

[24] Meghdad Ghari *Pavelka-style fuzzy justification logics,* Logic Journal of the IGPL, 24 (5), 2016, pp. 743–772.

[25] Remo Goetschi, Roman Kuznets *Realization for justification logics via nested sequents: Modularity through embedding,* APAL, 163 (9), September 2012, pp. 1271–1298 Published online March 2012.

[26] Mark Jago The impossible: An essay on hyperintensionality, Oxford University Press, Oxford 2014.

[27] Ioannis Kokkinis, Petar Maksimović, Zoran Ognjanović, Thomas Studer *First steps towards probabilistic justification logic,* Logic Journal of IGPL, 23 (4), 2015, pp. 662–687.

[28] Ioannis Kokkinis, Zoran Ognjanović, Thomas Studer *Probabilistic justification logic,* in S. Artemov, A. Nerode (eds.), Proceedings of Logical Foundations of Computer Science LFCS'16 Springer, 2016.

[29] Roman Kuznets, Thomas Studer *Justifications, Ontology, and Conservativity,* in Thomas Bolander, Torben Braüner, Silvio Ghilardi, Lawrence Moss

(eds.), Advances in Modal Logic, Volume 9, pp. 437–458 College Publications, 2012.

[30] Roman Kuznets, Thomas Studer *Update as Evidence: Belief Expansion,* in Sergei [N.] Artemov, Anil Nerode (eds.), LFCS 2013, volume 7734 of *LNCS*, pp. 266–279 Springer, 2013.

[31] Roman Kuznets, Thomas Studer *Weak arithmetical interpretations for the Logic of Proofs,* Logic Journal of IGPL, 24 (3), 2016, pp. 424–440.

[32] Roman Kuznets, Thomas Studer Logics of Proofs and Justifications, College Publications, London 2019.

[33] Eveline Lehmann, Thomas Studer *Subset models for justification logic,* Logic, Language, Information and Computation (WoLLIC), 2019, pp. 433–449.

[34] Eveline Lehmann, Thomas Studer *Subset models for justification logic,* E-print 1902.02707, arXiv.org, 2019.

[35] Michel Marti, Thomas Studer *Intuitionistic modal logic made explicit* IfCoLog Journal of Logics and their Applications, 3, 2016, pp. 877–901.

[36] Alexey Mkrtychev *Models for the Logic of Proofs,* in Sergei Adian, Anil Nerode (eds.), LFCS'97, volume 1234 of *LNCS*, pp. 266–275 Springer, 1997.

[37] Bryan Renne *Public communication in justification logic,* Journal of Logic and Computation, 21 (6), 2011, pp. 1005–1034.

[38] Nenad Savić, Thomas Studer *Relevant Justification Logic,* IfCoLoG Journal of Logics and their Applications, 6 (2), 2019, pp. 397–412.

[39] Thomas Studer *Decidability for Some Justification Logics with Negative Introspection,* JSL, 78 (2), June 2013, pp. 388–402.