Advances in Modal Logic
2024

Short Papers

Contents

First-order Probabilistic Logic with Sequence Variables and Unranked Symbols

Merium Hazem Anwar Labib Bishara

International Black Sea University
`22100636@ibsu.edu.ge`

Lia Kurtanidze

Georgian National University SEU
`l.kurtanidze@seu.edu.ge`

Mikheil Rukhaia

Institute of Applied Mathematics, Tbilisi State University
`mrukhaia@logic.at`

Lali Tibua

Institute of Applied Mathematics, Tbilisi State University
`ltibua@viam.tsu.ge`

Abstract

We present an unranked probabilistic logic, which is an extension of the first-order probability logic with sequence variables and flexible-arity (unranked) function and predicate symbols. The semantics of the logic is defined using Kripke worlds and the strong completeness theorem holds for it. Such a formalism is interesting as it provides very flexible and expressive platform to model various problems coming from real world applications.

Keywords: Unranked symbols, sequence variables, probabilistic primitives.

1 Introduction

Since the early days of Artificial Intelligence (AI) logical and probabilistic methods have been independently used in order to solve tasks that require some sort of "intelligence". Probability theory deals with the challenges posed by uncertainty, while logic is more often used for reasoning with perfect knowledge. Considerable efforts have been devoted to combining logical and probabilistic methods in a single framework, which influenced the development of several formalisms and programming tools. Among others, the most prominent ones include Independent Choice Logic (ICL) [22], PRISM [24], Markov

Logic Networks (MLN) [23], CLP(BN) [9], Bayesian Logic Programs [13], P-log [2], ProbLog [10], and Probabilistic Soft Logic (PSL) [1]. These languages and formalisms have been successfully applied to many domains. Some of the applications include web mining, natural language processing, robotics, transportation systems, communication networks, social networks, medicine, bio- and chemo-informatics, electronic games, and activity recognition.

All probabilistic logic formalisms studied so far are either propositional, or permit only individual variables, i.e., variables that can be instantiated by a single term. On the other hand, theories and systems that use not only individual variables but also sequence variables (these variables can be replaced by arbitrary finite, possibly empty, sequences of terms) have emerged. Recently, the usefulness of sequence variables and unranked symbols (function and/or predicate symbols without fixed arity) has been shown in several formalisms and illustrated in practical applications related to XML [6,7], schema transformation operations [5], knowledge representation [12,8], automated reasoning [21,4], rewriting [25], functional, functional-logic, and rule-based programming [17,16], just to name a few. There are systems for programming with sequence variables. Probably the most prominent one is Mathematica [26], with a powerful rule-based programming language that uses (essentially first order, equational) unranked matching with sequence variables [3]. The unranked term is a first-order term, where the same function symbol can occur in different places with different number of arguments. Unranked function symbols and sequence variables bring a great deal of expressiveness in this language, permitting writing a short, concise, readable code.

We develop a novel theory, where sequence variables, unranked terms and probabilistic primitives will be available together. Such a formalism is interesting from theoretical point of view as well as from practical one, since it provides very flexible and expressive platform to model various problems coming from real world applications. We define an unranked probabilistic first-order logic LFOP_u, which is an extension of the probabilistic first-order logic LFOP_1 [19,20,18] with sequence variables and unranked function and predicate symbols [15,11,14].

2 Syntax

We consider an alphabet $\mathcal{A}$ consisting of the following pairwise disjoint sets of symbols:

- the set $\mathcal{V}_i$ of individual variables, denoted by $x^i, y^i, z^i, \ldots$,
- the set $\mathcal{V}_s$ of sequence variables, denoted by $\overline{x}, \overline{y}, \overline{z}, \ldots$,
- the set $\mathcal{F}_i^r$ of fixed arity (ranked) individual function symbols, denoted by $f^r, g^r, \ldots$,
- the set $\mathcal{F}_i^u$ of flexible arity (unranked) individual function symbols, denoted by $f^u, g^u, \ldots$,
- the set $\mathcal{F}_s^r$ of fixed arity (ranked) sequence function symbols, denoted by

$\overline{f}^r, \overline{g}^r, \ldots$,

- the set $\mathcal{F}_s^u$ of flexible arity (unranked) sequence function symbols, denoted by $\overline{f}^u, \overline{g}^u, \ldots$,
- the set $\mathcal{P}^r$ of fixed arity (ranked) predicate symbols, denoted by $p^r, q^r, \ldots$,
- the set $\mathcal{P}^u$ of flexible arity (unranked) predicate symbols, denoted by $p^u, q^u, \ldots$,
- logical connectives and quantifiers $\neg, \wedge, \vee, \rightarrow, \exists, \forall$,
- a list of unary probability opertors $P_{[a,b]}$, for every $[a,b] \subseteq [0,1]$ and $a, b \in \mathbb{Q}$,
- auxiliary symbols: parentheses and the comma.

Each ranked symbol has a unique arity (rank) associated with it. If the rank of a function symbol is 0, then it is called a *constant*. Unranked symbols do not have a fixed arity. Each set of variables, function symbols, and predicate symbols is countably infinite. The letter $\mathcal{V}$ denotes the set $\mathcal{V}_i \cup \mathcal{V}_s$ and its elements are denoted by $x, y, z, \ldots$. The letter $\mathcal{P}$ denotes the set of all predicate symbols, i.e. $\mathcal{P} = \mathcal{P}^r \cup \mathcal{P}^u$. Respectively, the letters $\mathcal{F}_i = \mathcal{F}_i^r \cup \mathcal{F}_i^u$ and $\mathcal{F}_s = \mathcal{F}_s^r \cup \mathcal{F}_s^u$ denote the set of all individual and sequence function symbols. We use the letters f, g to denote elements of $\mathcal{F}_i$, the letters $\overline{f}, \overline{g}$ to denote elements of $\mathcal{F}_s$, and the letters p, q for the elements of $\mathcal{P}$. We might use these letters with or without the indices; the set to which they belong will be specified explicitly or will be clear from the context.

The intended meaning of the unary probability operators $P_{[a,b]}A$ is that "the probability of A is in interval $[a,b]$", where $[a,b]$ is an interval of rational numbers between $[0,1]$ [1].

Note that, defining the probability operator using intervals is a more general approach and the probability operators $P_{\geq s}$ (as defined in [18]) correspond to $P_{[s,1]}$. Analogously, the abbreviated operators $P_{\leq s}, P_{=s}, P_{<s}, P_{>s}$ can be represented as $P_{[0,s]}, P_{[s,s]}, P_{[0,s]} \wedge \neg P_{[s,s]}, P_{[s,1]} \wedge \neg P_{[s,s]}$, respectively.

The *terms* are defined as *individual* and *sequence* terms over $\mathcal{A}$ in the following inductive way:

$$
\begin{aligned}
t &::= x^i \mid f^r(t_1, \ldots, t_n) \mid f^u(\overline{s}_1, \ldots, \overline{s}_n) && \text{individual terms} \\
\overline{s} &::= t \mid \overline{x} \mid \overline{f}^r(t_1, \ldots, t_n) \mid \overline{f}^u(\overline{s}_1, \ldots, \overline{s}_n), \quad n \geq 0 && \text{sequence terms}
\end{aligned}
$$

An *atom* is a formula of the form $p^r(t_1, \ldots, t_n), n \geq 0$ and $p^u(\overline{s}_1, \ldots, \overline{s}_m), m \geq 0$, where $p^r \in \mathcal{P}^r$ is an n-ary predicate symbol, and $p^u \in \mathcal{P}^u$ is a flexible arity predicate symbol. *Formulas* are built in an usual inductive fashion from atoms, unary probability operators $P_{[a,b]}$, and logical connectives $\neg, \wedge, \vee, \rightarrow, \exists$, and $\forall$. Quantifications are allowed on both, individual as well as on sequence variables. We use the letters $A, B, \ldots$ to denote formulas.

Example 2.1 The famous lottery paradox, that for each ticket the winning

---
[1] The requirement of rational numbers is not essential, but for technical reasons only.

chance is very low, but there is a high chance that some tickets win, can be modelled as:

$$(\forall x)P_{[0,0.000001]}Win(x) \wedge P_{[0.999999,1]}(\exists x_1,\ldots,x_n)Win(x_1,\ldots,x_n)$$

where $Win \in \mathcal{P}^u$ is a flexible arity predicate symbol and $Win(x_1,\ldots,x_n)$ means that the tickets $x_1,\ldots,x_n$ win the lottery.

A *substitution* is a finite set of distinct variable bindings, where a variable binding is either an expression $x^i \mapsto t$ or $\overline{x} \mapsto \overline{s}$. Substitutions are denoted by σ, θ and the empty substitution is denoted by ϵ. The application of a substitution σ on a term t and a formula A is defined in the usual way (e.g. as given in [15]) and is denoted by $t\sigma$ and $A\sigma$, respectively.

3 Semantics

The semantics for LFOP_u is based on the well-known Kripke frames, where accessibility relations are replaced by probability measures.

An LFOP_u-model is a structure $\mathbf{M} = \langle W, D, I, Prob\rangle$, where:

- W is a nonempty set of worlds,
- D is a nonempty domain for every world $w \in W$ and it is a union of two disjoint sets D_i and D_s, where $D_i \neq \emptyset$,
- I is an interpretation that for each $w \in W$, associates $I(w)$ for every:
  · individual constant c to an element $I(w)(c) \in D_i$,
  · sequence constant $\overline{c}$ to an element $I(w)(\overline{c}) \in D^\infty$, [2]
  · n-ary individual function symbol $f^r \in \mathcal{F}_i^r$, with $n > 0$, to an n-ary function $I(w)(f^r)\colon D_i^n \to D_i$
  · flexible arity individual function symbol $f^u \in \mathcal{F}_i^u$ to a flexible arity function $I(w)(f^u)\colon D^\infty \to D_i$
  · n-ary sequence function symbol $\overline{f}^r \in \mathcal{F}_s^r$, with $n > 0$, to an n-ary multi-valued function $I(w)(\overline{f}^r)\colon D_i^n \to D^\infty$
  · flexible arity sequence function symbol $\overline{f}^u \in \mathcal{F}_s^u$ to a flexible arity multi-valued function $I(w)(\overline{f}^u)\colon D^\infty \to D^\infty$
  · n-ary predicate symbol $p^r \in \mathcal{P}^r$, with $n \geq 0$, to an n-ary predicate $I(w)(p^r) \subseteq D_i^n$
  · flexible arity predicate symbol $p^u \in \mathcal{P}^u$ to a flexible arity predicate $I(w)(p^u) \subseteq D^\infty$
- *Prob* is a probability assignment that for each $w \in W$ assigns a probability space $Prob(w) = \langle W(w), H(w), \mu(w)\rangle$, where:
  · $W(w) \subseteq W$ and $W(w) \neq \emptyset$,
  · $H(w)$ is an algebra over subsets of $W(w)$, defined as:
    $W(w) \in H(w)$, and
    if $\alpha, \beta \in H(w)$, then $W(w) \setminus \alpha \in H(w)$ and $\alpha \cup \beta \in H(w)$

[2] $D^\infty = \bigcup_{n\geq 0} D^n$, where D^n, $n \geq 1$, is a set of all n-tuples over D and $D^0 = \{\epsilon\}$, for ϵ denoting the empy tuple.

· $\mu(w)\colon H(w) \to [0,1]$ is a finitely additive probability measure, defined as:
$\mu(w)(W(w)) = 1$, and
if $\alpha \cap \beta = \emptyset$, then $\mu(w)(\alpha \cup \beta) = \mu(w)(\alpha) + \mu(w)(\beta)$.

An **M**-evaluation is a mapping e, which assigns to each individual variable x^i an element $e(x^i) \in D_i$ and to each sequence variable $\overline{x}$ an element $e(\overline{x}) \in D^\infty$. Further, $e[x^i \mapsto d]$, for $d \in D_i$, is an M-evaluation such that $e[x^i \mapsto d](x^i) = d$ and $e[x^i \mapsto d](y^i) = e(y^i)$ for each individual variable y^i different from x^i. Analogously, for any $d \in D^\infty$, $e[\overline{x} \mapsto d](\overline{x}) = d$ and $e[\overline{x} \mapsto d](\overline{y}) = e(\overline{y})$ for each sequence variable $\overline{y}$ different from $\overline{x}$.

The value of a term t in a world w with respect to an **M**-evaluation e, denoted by $\|I(w)(t)\|_e^{\mathbf{M}}$, is:

- $\|I(w)(x)\|_e^{\mathbf{M}} = e(x)$, for x being an individual or a sequence variable,
- $\|I(w)(f(t_1,\dots,t_n))\|_e^{\mathbf{M}} = I(w)(f)(\|t_1\|_e^{\mathbf{M}},\dots,\|t_n\|_e^{\mathbf{M}})$, for $f \in \{f^r, f^u\}$ and $t_1,\dots,t_n$ being individual terms,
- $\|I(w)(f(\overline{s}_1,\dots,\overline{s}_n))\|_e^{\mathbf{M}} = I(w)(f)(\|\overline{s}_1\|_e^{\mathbf{M}},\dots,\|\overline{s}_n\|_e^{\mathbf{M}})$, for $f \in \{\overline{f}^r, \overline{f}^u\}$ and $\overline{s}_1,\dots,\overline{s}_n$ being sequence terms.

The truth value of a formula A in a world w with respect to an **M**-evaluation e, denoted by $\|I(w)(A)\|_e^{\mathbf{M}}$, is:

- $\|I(w)(p^r(t_1,\dots,t_n))\|_e^{\mathbf{M}} = \mathit{True}$, if $\langle \|t_1\|_e^{\mathbf{M}},\dots,\|t_n\|_e^{\mathbf{M}} \rangle \in I(w)(p^r)$,
- $\|I(w)(p^u(\overline{s}_1,\dots,\overline{s}_n))\|_e^{\mathbf{M}} = \mathit{True}$, if $\langle \|\overline{s}_1\|_e^{\mathbf{M}},\dots,\|\overline{s}_n\|_e^{\mathbf{M}} \rangle \in I(w)(p^u)$,
- $\|I(w)(P_{[a,b]}A)\|_e^{\mathbf{M}} = \mathit{True}$, if $\mu(w)(\{u \mid u \in W(w) \text{ and } \|I(u)(A)\|_e^{\mathbf{M}} = \mathit{True}\}) \in [a,b]$,
- $\|I(w)(\neg A)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{False}$,
- $\|I(w)(A \wedge B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ and $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)(A \vee B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ or $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)(A \to B)\|_e^{\mathbf{M}} = \mathit{True}$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{False}$ or $\|I(w)(B)\|_e^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)((\forall x^i)A)\|_e^{\mathbf{M}} = \mathit{True}$, if for every $d \in D_i$, $\|I(w)(A)\|_{e[x^i \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)((\forall \overline{x})A)\|_e^{\mathbf{M}} = \mathit{True}$, if for every $d \in D^\infty$, $\|I(w)(A)\|_{e[\overline{x} \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)((\exists x^i)A)\|_e^{\mathbf{M}} = \mathit{True}$, if for some $d \in D_i$, $\|I(w)(A)\|_{e[x^i \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
- $\|I(w)((\exists \overline{x})A)\|_e^{\mathbf{M}} = \mathit{True}$, if for some $d \in D^\infty$, $\|I(w)(A)\|_{e[\overline{x} \mapsto d]}^{\mathbf{M}} = \mathit{True}$,
- in all other cases $I(w)(A) = \mathit{False}$

A formula A is *satisfied in a world* w, written as $\mathbf{M}, w \models A$, if $\|I(w)(A)\|_e^{\mathbf{M}} = \mathit{True}$ for every valuation e. A is *satisfiable in a model* **M**, if there exists $w \in W$ such that $w \models A$; and it is *valid in a model* **M**, written as $\mathbf{M} \models A$, if A is satisfied in every $w \in W$. Finally, A is a *valid formula* of LFOP$_u$, written as $\models A$, if it is valid in every model. We say that a formula A is a *semantic consequence* of a set of formulas T, written as $T \models A$, if in every model **M**, where $\mathbf{M} \models A_j$ for all $A_j \in T$, also $\mathbf{M} \models A$.

4 Axiom System

The axiom system for the unranked probabilistic logic LFOP_u contains the following axiom schemata:

(i) all instances of classical propositional logic axioms

(ii) $(\forall x)(A \rightarrow B) \rightarrow (A \rightarrow (\forall x)B)$, for an individual or a sequence variable x not being free in A

(iii) $(\forall x)A(x) \rightarrow A(t)$, for both x and t being either individual or sequence terms

(iv) $P_{[0,1]}A$

(v) $P_{[a,b]}A \rightarrow P_{[a_1,b_1]}A$, for $a_1 \leq a$ and $b_1 \geq b$

(vi) $(P_{[a,1]}A \wedge P_{[b,1]}B \wedge P_{[1,1]}(\neg A \vee \neg B)) \rightarrow P_{[\mathtt{min}(1,a+b),1]}(A \vee B)$

(vii) $(P_{[0,a]}A \wedge P_{[0,b]}B) \rightarrow P_{[0,\mathtt{min}(1,a+b)]}(A \vee B)$

and the inference rules:

$$\frac{A \quad A \rightarrow B}{B}\ \mathsf{MP} \qquad \frac{A}{P_{[1,1]}A}\ \mathsf{Nec} \qquad \frac{A(x^i)}{(\forall x^i)A}\ \mathsf{Gen_1} \qquad \frac{A(\overline{x})}{(\forall \overline{x})A}\ \mathsf{Gen_2}$$

$$\frac{A \rightarrow P_{[a-\frac{1}{a_1},b+\frac{1}{b_1}]}B, \text{ for every integer } a_1 \geq \frac{1}{a} \text{ and } b_1 \geq \frac{1}{1-b}}{A \rightarrow P_{[a,b]}B}\ \mathsf{Arch}$$

The notion of derivability of a formula A from a set of formulas T, denoted by $T \vdash A$, is defined in the usual sense.

Example 4.1 The set of formulas $T = \{\neg P_{[0,0]}A\} \cup \{P_{[0,\frac{1}{n}]}A \mid \text{ for all } n \in \mathbb{N}\}$ is inconsistent: $T \vdash P_{[0,0+\frac{1}{n}]}A$ for every integer $n > 0$, then by the Arch rule $T \vdash P_{[0,0]}A$; but, $T \vdash \neg P_{[0,0]}A$, thus $T \vdash \bot$.

It is easy to see that the LFOP_u logic preserves the properties of both, probabilistic and unranked predicate logics.

Theorem 4.2 (Soundness) *The axiom system (1)-(7) is sound with respect to the $LFOP_u$ semantics.*

There are usually two forms of completeness theorems: *the weak completeness* – a formula is consistent iff it is satisfiable; or *the strong completeness* – a set of formulas is consistent iff it is satisfiable. Clearly, the weak completeness follows from the strong one, but not vice versa. In classical logics these theorems are equivalent due to *the compactness theorem* – a set of formulas is satisfiable iff every finite subset of it is satisfiable. But, according to [20], in probabilistic logics compactness usually fails and even more, the strong completeness is not available in some of them. The LFOP_u logic inherits strong completeness from LFOP_1.

Theorem 4.3 (Strong completeness) *Let A be an $LFOP_u$ formula and T be a set of $LFOP_u$ formulas, then $T \models A$ iff $T \vdash A$*

Acknowledgments

This work was supported by Shota Rustaveli National Science Foundation of Georgia under the grant №FR-22-4254.

References

[1] Bach, S. H., M. Broecheler, B. Huang and L. Getoor, *Hinge-loss markov random fields and probabilistic soft logic*, Journal of Machine Learning Research **18** (2017), pp. 1–67.

[2] Baral, C., M. Gelfond and N. Rushton, *Probabilistic reasoning with answer sets*, Theory and Practice of Logic Programming **9** (2009), pp. 57–144.

[3] Buchberger, B., *Mathematica as a rewrite language*, in: T. Ida, A. Ohori and M. Takeich, editors, *Functional and Logic Programming - 2nd Fuji International Workshop, FLOPS 1996, Shonan Village, Japan, June 4-6, 1996. Proceedings*, Lecture Notes in Computer Science (1996), pp. 1–13.

[4] Buchberger, B., A. Craciun, T. Jebelean, L. Kovács, T. Kutsia, K. Nakagawa, F. Piroi, N. Popov, J. Robu, M. Rosenkranz and W. Windsteiger, *Theorema: Towards computer-aided mathematical theory exploration*, J. Applied Logic **4** (2006), pp. 470–504.

[5] Chasseur, E. and Y. Deville, *Logic program schemas, constraints, and semi-unification*, in: *International Workshop on Logic Programming Synthesis and Transformation*, Springer, 1997, pp. 69–89.

[6] Coelho, J. and M. Florido, *XCentric: logic programming for XML processing*, in: I. Fundulaki and N. Polyzotis, editors, *9th ACM International Workshop on Web Information and Data Management (WIDM 2007), Lisbon, Portugal, November 9, 2007* (2007), pp. 1–8.

[7] Coelho, J., M. Florido and T. Kutsia, *Collaborative schema construction using regular sequence types*, in: *Proceedings of the IEEE International Conference on Information Reuse and Integration, IRI 2009, 10-12 August 2009, Las Vegas, Nevada, USA* (2009), pp. 290–295.

[8] Common Logic Working Group, *Common Logic Working Group Documents: Common Logic Standard* (2007), http://common-logic.org/.

[9] Costa, V. S., D. Page and J. Cussens, *CLP(BN): Constraint Logic Programming for Probabilistic Knowledge*, in: *Probabilistic inductive logic programming*, Springer, 2008 pp. 156–188.

[10] De Raedt, L. and A. Kimmig, *Probabilistic (logic) programming concepts*, Machine Learning **100** (2015), pp. 5–47.

[11] Dundua, B., L. Kurtanidze and M. Rukhaia, *Unranked tableaux calculus for web related applications*, in: *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, IEEE, 2017, pp. 1181–1184.

[12] Genesereth, M. R., *Knowledge Interchange Format, draft proposed American National Standard (dpANS)*, Technical Report NCITS.T2/98-004, Stanford University (1998), available from http://logic.stanford.edu/kif/dpans.html.

[13] Kersting, K. and L. De Raedt, *Bayesian logic programming: theory and tool*, in: L. Getoor and B. Taskar, editors, *Introduction to Statistical Relational Learning*, MIT press, 2007 .

[14] Kurtanidze, L. and M. Rukhaia, *Skolemization in unranked logics*, Bulletin of TICMI **22** (2018), pp. 3–10.

[15] Kutsia, T. and B. Buchberger, *Predicate logic with sequence variables and sequence function symbols*, in: *International Conference on Mathematical Knowledge Management*, Springer, 2004, pp. 205–219.

[16] Marin, M. and T. Kutsia, *Foundations of the rule-based system ρLog*, Journal of Applied Non-Classical Logics **16** (2006), pp. 151–168.

[17] Marin, M. and D. Tepeneu, *Programming with sequence variables: The Sequentica package*, in: P. Mitic, P. Ramsden and J. Carne, editors, *Challenging the Boundaries*

of Symbolic Computation. Proceedings of 5th International Mathematica Symposium (2003), pp. 17–24.

[18] Ognjanović, Z. and A. Ilić-Stepić, *Logics with probability operators*, in: *Probabilistic Extensions of Various Logical Systems*, Springer, 2020 pp. 1–35.

[19] Ognjanovic, Z. and M. Raškovic, *Some first-order probability logics*, Theoretical Computer Science **247** (2000), pp. 191–212.

[20] Ognjanović, Z., M. Rašković and Z. Marković, "Probability logics: probability-based formalization of uncertain reasoning," Springer, 2016.

[21] Paulson, L. C., *Isabelle: The next 700 theorem provers*, in: *Logic and Computer Science*, Academic Press, 1990 pp. 361–386.

[22] Poole, D., *Abducing through negation as failure: Stable models within the independent choice logic*, The Journal of Logic Programming **44** (2000), pp. 5–35.

[23] Richardson, M. and P. Domingos, *Markov logic networks*, Machine learning **62** (2006), pp. 107–136.

[24] Sato, T. and Y. Kameya, *Parameter learning of logic programs for symbolic-statistical modeling*, Journal of Artificial Intelligence Research **15** (2001), pp. 391–454.

[25] Widera, M. and C. Beierle, *A term rewriting scheme for function symbols with variable arity*, Technical Report 280, Praktische Informatik VIII, FernUniversität Hagen, Germany (2001).

[26] Wolfram, S., "The Mathematica Book," Wolfram Media, 2003, 5th edition.

Simplicial Belief

Christian Cachin David Lehnherr [1] Thomas Studer

Institute of Computer Science, University of Bern, Bern, Switzerland

Abstract

Recently, much work has been carried out to study simplicial interpretations of modal logic. While notions of (distributed) knowledge have been well investigated in this context, it has been open how to model belief in simplicial models. We introduce polychromatic simplicial complexes, which naturally impose a plausibility relation on states. From this, we can define various notions of belief.

Keywords: Simplicial complex, epistemic logic, plausibility model, belief modality.

1 Introduction

Simplicial interpretations for modal logic are currently avidly researched; see, e.g., [3,7,8,10,12] due to their close connection with distributed computing [9]. At its core lies the epistemic interpretation of simplicial complexes of various kinds. Let $\mathcal{V}$ be a set of vertices. Each vertex corresponds to a local state of an agent, and we say that this vertex is of that agent's color. In the simplest case, a simplicial complex $(S, \mathcal{V})$ is a pair where S is a set of subsets of $\mathcal{V}$ that is closed under set inclusion. Vertices that belong to the same set must be of different colors, and maximal elements of S represent global states. An agent a cannot distinguish two global states if its local state is included in both. Hence, simplicial complexes offer sufficient structure for an epistemic interpretation. While (distributed) knowledge has been studied extensively in this context, it has been open, see [4], how to model belief on simplicial structures such that

(i) belief depends only on the topological structure of the simplicial complex;

(ii) the principle of knowledge-yields-belief holds.

In this brief announcement, we present polychromatic simplicial complexes, i.e., complexes that are not necessarily properly colored. We define a plausibility relation between the states based on the multiplicity of a color within a state. If the color of an agent a has a lower or equal multiplicity in a state X than in a state Y, then a considers X to be at least as plausible as Y. This relation is a wellfounded preorder, and hence, we can use the machinery of plausibility models [1,2] to define various notions of belief such as plausible belief

[1] Supported by the Swiss National Science Foundation (SNSF) under grant agreement Nr. 200020_219403 (Emerging Consensus).

and safe belief. Moreover, our structures also satisfy the knowledge-yields-belief principle.

Acknowledgements. We would like to thank the organizers and participants of the Dagstuhl meeting *Epistemic and Topological Reasoning in Distributed Systems* [4], especially the working group on representing epistemic attitudes via simplicial complexes.

2 Simplicial Knowledge

We quickly recall the standard definitions for distributed knowledge on simplicial complexes [7,8,12]. In the subsequent section, we will extend them to incorporate notions of belief.

Let Ag be the set of finitely many agents, and let Prop be a countable set of atomic propositions. We define the language of knowledge $\mathcal{L}_{\mathcal{K}}$ for $G \subseteq \mathsf{Ag}$ and $p \in \mathsf{Prop}$ inductively by the following grammar:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid [\sim]_G\phi.$$

The remaining Boolean connectives are defined as usual. In particular, we set $\bot := p \wedge \neg p$ for some fixed $p \in \mathsf{Prop}$. We write $\mathsf{alive}(G)$ for $\neg[\sim]_G\bot$ and $\mathsf{dead}(G)$ for $[\sim]_G\bot$.

Definition 2.1 Let $\mathcal{V}$ be a set of vertices. $C = (S, \mathcal{V})$ with $S \subseteq \mathsf{Pow}(\mathcal{V}) \setminus \{\emptyset\}$ is called a *simplicial complex* if

$$\text{for each } X \in S \text{ and each } \emptyset \neq Y \subseteq X, \text{ we have } Y \in S.$$

We call the elements of S *faces*. A face that is maximal under inclusion is called a *facet*. We denote the set of facets of C by $\mathcal{F}(C)$. A *coloring* is a mapping $\chi : \mathcal{V} \to \mathsf{Ag}$. A coloring is *proper* if it assigns a different agent to each vertex within a face. We use $\chi(U)$ for the set $\{\chi(u) \mid u \in U\}$.

Definition 2.2 Let $C = (S, \mathcal{V})$ be a simplicial complex. A *simplicial model* $\mathcal{C} = (C, \chi, W, \ell)$ is a quadruple where

(i) C is a simplicial complex;

(ii) $\chi : \mathcal{V} \to \mathsf{Ag}$ is a proper coloring;

(iii) $\mathcal{F}(C) \subseteq W \subseteq S$ is a set of worlds;

(iv) $\ell : W \to \mathsf{Pow}(\mathsf{Prop})$ is a valuation.

Given a simplicial model, a group of agents $G \subseteq \mathsf{Ag}$ cannot distinguish two worlds $X, Y \in W$, denoted by $X \sim_G Y$, if and only if $G \subseteq \chi(X \cap Y)$. We call $\sim_G$ the *epistemic indistinguishability relation.* If G contains only a single agent a, we write $X \sim_a Y$ and $[\sim]_a$ instead of $X \sim_{\{a\}} Y$ and $[\sim]_{\{a\}}$, respectively.

Definition 2.3 For a simplicial model $\mathcal{C} = (C, W, \chi, \ell)$, a world $X \in W$, and a formula $\phi \in \mathcal{L}_\mathcal{K}$, we define the relation $\mathcal{C}, X \Vdash \phi$ inductively by

$$\begin{array}{lll} \mathcal{C}, X \Vdash p & \text{iff} & p \in \ell(X) \\ \mathcal{C}, X \Vdash \neg\phi & \text{iff} & \mathcal{C}, X \nVdash \phi \\ \mathcal{C}, X \Vdash \phi \wedge \psi & \text{iff} & \mathcal{C}, X \Vdash \phi \text{ and } \mathcal{M}, X \Vdash \psi \\ \mathcal{C}, X \Vdash [\sim]_G\phi & \text{iff} & X \sim_G Y \text{ implies } \mathcal{C}, Y \Vdash \phi \quad \text{for all } Y \in W. \end{array}$$

We say that *agent a is alive in a world X* if $a \in \chi(X)$. The set of worlds in which a group $G \subseteq \mathsf{Ag}$ is alive is defined as

$$\mathsf{Alive}_\mathcal{C}(G) = \{X \in W \mid G \subseteq \chi(X)\}.$$

Lemma 2.4 *Let $\mathcal{C} = (C, \chi, W, \ell)$ be a simplicial model. For each $G \subseteq \mathsf{Ag}$, the relation $\sim_G$ is an equivalence relation on $\mathsf{Alive}_\mathcal{C}(G)$ and empty otherwise.*

3 Simplicial belief

We now drop the requirement that the coloring of a simplicial model must be proper. The resulting models are called polychromatic. We will define a wellfounded preorder on the states of a polychromatic model, which will serve as a plausibility relation [1,2]. This makes it possible to interpret various notions of belief on simplicial models.

It is straightforward to verify that Lemma 2.4 does not hold for polychromatic models because $\sim_G$ need not be transitive. Indeed, consider the set of vertices $\{0, 1, 2, 3\}$ and the complex consisting of the facets

$$X := \{0, 1\}, \quad Y := \{1, 2\}, \quad \text{and} \quad Z = \{2, 3\}$$

with a coloring χ that assigns the same agent a to every vertex. We find that $X \sim_a Y$ and $Y \sim_a Z$, but not $X \sim_a Z$.

In order to re-establish transitivity of $\sim_G$, we must require that for any three worlds $X, Y, Z \in W$ and any group of agents $G \subseteq \mathsf{Ag}$:

$$G \subseteq \chi(X \cap Y) \text{ and } G \subseteq \chi(Y \cap Z) \text{ implies } G \subseteq \chi(X \cap Z). \qquad (\star)$$

Definition 3.1 A *polychromatic model* is a simplicial model where:

(i) the coloring is not required to be proper;

(ii) condition $(\star)$ holds.

Definition 3.2 Let (C, χ, W, ℓ) be a polychromatic model. We define the *multiplicity* of $a \in \mathsf{Ag}$ in a world X by

$$m_a(X) = |\{v \in X \mid \chi(v) = a\}|$$

where $|\cdot|$ denotes the cardinality of a set. Note that if agent a is alive in a world X, then $m_a(X) \geq 1$.

For $X, Y \in W$ and $a \in \mathsf{Ag}$, we write

$$X \leq_a Y \quad \text{iff} \quad m_a(X) \leq_a m_a(Y).$$

The multiplicity of a color within a face induces for each agent a a well-founded relation $\leq_a$ on worlds. We call this the (a priori) plausibility relation. Notice that $\leq_a$ is a priori in the sense that it does not refer to the actual world, i.e., it does not account for possibility. We introduce a local plausibility relation

$$\trianglelefteq_a \ := \ \leq_a \cap \sim_a,$$

which captures the agent's plausibility relation at a given state. Further, we write $X \geq_a Y$ if and only if $m_X(a) \geq m_Y(a)$ and we use $\trianglerighteq_a$ and $\triangleleft_a$ in the obvious way. The following lemma shows that the indistinguishability relation can be given in terms of the local plausibility relation.

Lemma 3.3 $\sim_a = \trianglelefteq_a \cup \trianglerighteq_a$.

From the relation $\trianglerighteq_a$, we get a corresponding modal operator $[\trianglerighteq]_a$, which is referred to in the literature as safe belief [2]. Our language of knowledge and belief $\mathcal{L}_{\mathcal{KB}}$ extends $\mathcal{L}_{\mathcal{K}}$ by the modal operator $[\trianglerighteq]_a$ for each agent $a \in \mathsf{Ag}$. It is inductively defined by as follows:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid [\sim]_G\phi \mid [\trianglerighteq]_a\phi$$

where $p \in \mathsf{Prop}$. As usual, the dual of safe belief is defined as $\langle\trianglerighteq\rangle_a\varphi \equiv \neg[\trianglerighteq]_a\neg\varphi$.

Definition 3.4 For a polychromatic model $\mathcal{C} = (C, \chi, W, \ell)$, a world $X \in W$, and a formula $\phi \in \mathcal{L}_{\mathcal{KB}}$, we define the relation $\mathcal{C}, X \Vdash \phi$ inductively by

$$\begin{array}{lll}
\mathcal{C}, X \Vdash p & \text{iff} & p \in \ell(X) \\
\mathcal{C}, X \Vdash \neg\phi & \text{iff} & \mathcal{C}, X \nVdash \phi \\
\mathcal{C}, X \Vdash \phi \wedge \psi & \text{iff} & \mathcal{C}, X \Vdash \phi \text{ and } \mathcal{M}, X \Vdash \psi \\
\mathcal{C}, X \Vdash [\sim]_G\phi & \text{iff} & X \sim_G Y \quad \text{implies} \quad \mathcal{C}, Y \Vdash \phi \quad \text{for all } Y \in W \\
\mathcal{C}, X \Vdash [\trianglerighteq]_a\phi & \text{iff} & X \trianglerighteq_a Y \quad \text{implies} \quad \mathcal{C}, Y \Vdash \phi \quad \text{for all } Y \in W.
\end{array}$$

As usual with plausibility models, we can not only define safe belief but also other notions of belief.

Definition 3.5 Let $\mathcal{C} = (C, \chi, W, \ell)$ be a polychromatic model. For $X \in W$ we define

$$\mathsf{Min}_{\trianglelefteq_a}(X) = \{Y \in W \mid Y \sim_a X \text{ and } \nexists Z \in W. Z \triangleleft_a Y\}.$$

Since $\leq_a$ is wellfounded, we find that $\mathsf{Min}_{\trianglelefteq_a}(X) \neq \emptyset$ if agent a is alive in the world X.

We can now extend our language $\mathcal{L}_{\mathcal{KB}}$ with a new modality $\mathcal{B}_a$ for each agent a. We use the following truth definition.

Definition 3.6 For a polychromatic model $\mathcal{C} = (C, \chi, W, \ell)$, a world $X \in W$, and a formula $\phi \in \mathcal{L}_{\mathcal{KB}}$, we define

$$\mathcal{C}, X \Vdash \mathcal{B}_a\varphi \quad \text{iff} \quad Y \in \mathsf{Min}_{\trianglelefteq_a}(X) \text{ implies } \mathcal{C}, Y \Vdash \varphi \quad \text{for all } Y \in W.$$

The modality $\mathcal{B}_a$ models agent a's (most plausible) belief. It is well-known that $\mathcal{B}_a$ can be expressed in terms of the $[\trianglerighteq]_a$ modalitiy [2,11].

Lemma 3.7 *Let $\mathcal{C} = (C, \chi, W, \ell)$ be a polychromatic model, a an agent, and $X \in W$ such that a is alive in X. We find that*

$$\mathcal{C}, X \Vdash \mathcal{B}_a \phi \quad \text{if and only if} \quad \mathcal{C}, X \Vdash \langle \trianglerighteq \rangle_a [\trianglerighteq]_a \phi.$$

Our model satisfies the knowledge-yields-belief principle. In particular, we have the following lemma.

Lemma 3.8 *Let $\mathcal{C} = (C, \chi, W, \ell)$ be a polychromatic model and $X \in W$. For any agent a and any formula φ, we have*

$$\mathcal{C}, X \Vdash [\sim]_a \varphi \to [\trianglerighteq]_a \varphi \quad \text{and} \quad \mathcal{C}, X \Vdash [\trianglerighteq]_a \varphi \to \mathcal{B}_a \varphi.$$

4 Conclusion and future work

We presented the first interpretation of belief on a simplicial structure that depends only on the topological structure without requiring additional machinery like belief functions. Our approach consists of dropping the requirement that the coloring must be proper and using the multiplicity of color within a face as an inverted plausibility measure.

The study of polychromatic models is still in its infancy, and many basic properties need further investigation. For instance, simplicial models are proper, i.e. different worlds can be distinguished by at least one agent. Formally, Goubault et al. [6] express this as

$$\mathsf{alive}(G) \wedge \mathsf{dead}(G^c) \wedge \varphi \to [\sim]_G(\mathsf{dead}(G^c) \to \varphi)$$

being valid, where G^c stands for the complement of G. This no longer holds for polychromatic models.

Moreover, the analysis of polychromatic models is an important step towards simplicial models that are based on simplicial sets [5]. Informally, one could say that the actual vertex is repeated in such a model, and not just the color. In this case, the property ($\star$) trivially holds and must not be imposed as a restriction on the model.

References

[1] Baltag, A. and S. Smets, *Dynamic belief revision over multi-agent plausibility models*, Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision (LOFT) (2006), pp. 11–24.

[2] Baltag, A. and S. Smets, *A qualitative theory of dynamic interactive belief revision*, in: G. Bonanno, W. van der Hoek and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (LOFT7)* (2008), pp. 13–60.

[3] Cachin, C., D. Lehnherr and T. Studer, *Synergistic knowledge*, in: S. Dolev and B. Schieber, editors, *Stabilization, Safety, and Security of Distributed Systems - 25th International Symposium, SSS 2023, Jersey City, NJ, USA, October 2-4, 2023, Proceedings*, Lecture Notes in Computer Science **14310** (2023), pp. 552–567.
URL `https://doi.org/10.1007/978-3-031-44274-2_41`

[4] Castañeda, A., H. van Ditmarsch, R. Kuznets, Y. Moses and U. Schmid, *Epistemic and Topological Reasoning in Distributed Systems (Dagstuhl Seminar 23272)*, Dagstuhl Reports **13** (2024), pp. 34–65.
URL `https://drops.dagstuhl.de/entities/document/10.4230/DagRep.13.7.34`

[5] Friedman, G., *Survey Article: An elementary illustrated introduction to simplicial sets*, Rocky Mountain Journal of Mathematics **42** (2012), pp. 353–423.

[6] Goubault, E., R. Kniazev, J. Ledent and S. Rajsbaum, *Simplicial models for the epistemic logic of faulty agents*, CoRR **abs/2311.01351** (2023).
URL `https://doi.org/10.48550/arXiv.2311.01351`

[7] Goubault, É., J. Ledent and S. Rajsbaum, *A simplicial complex model for dynamic epistemic logic to study distributed task computability*, Inf. Comput. **278** (2021), p. 104597.
URL `https://doi.org/10.1016/j.ic.2020.104597`

[8] Goubault, É., J. Ledent and S. Rajsbaum, *A simplicial model for $KB4_n$: Epistemic logic with agents that may die*, in: P. Berenbrink and B. Monmege, editors, *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference)*, LIPIcs **219** (2022), pp. 33:1–33:20.
URL `https://doi.org/10.4230/LIPIcs.STACS.2022.33`

[9] Herlihy, M., D. N. Kozlov and S. Rajsbaum, "Distributed Computing Through Combinatorial Topology," Morgan Kaufmann, 2013.
URL `https://store.elsevier.com/product.jsp?isbn=9780124045781`

[10] Randrianomentsoa, R. F., H. van Ditmarsch and R. Kuznets, *Impure simplicial complexes: Complete axiomatization*, Logical Methods in Computer Science (to appear).

[11] Stalnaker, R., *On logics of knowledge and belief*, Philosophical Studies **128** (2006), pp. 169–199.

[12] van Ditmarsch, H., É. Goubault, J. Ledent and S. Rajsbaum, *Knowledge and simplicial complexes*, in: B. Lundgren and N. A. Nuñez Hernández, editors, *Philosophy of Computing* (2022), pp. 1–50.

Epistemic Positions: Towards a Formal Theory of Epistemic Injustice

Huimin Dong, Joris Hulstijn, and Réka Markovich

University of Luxembourg

1 Introduction

Automated decision support systems are more and more common. An important concern about such systems is bias in decision-making. In fact, automated decision support systems do sometimes suggest decisions that are unfair or unjustified for certain groups [2, 8]. How can we conceptualize such cases? In the field of social epistemology Miranda Fricker has introduced the notion of *epistemic injustice* [5]. There are two aspects: (i) *injustice*: there is a moral wrong or a legal right is being denied, and (ii) *epistemic*: the wrong is based on a difference in the knowledge or information that is available to some groups in society relative to other groups [5, 6, 11]. In this paper, we build on Fricker's concept of epistemic injustice and propose to formalize this notion as "the wrongful treatment of the right to be a knower" in the context of various irrational epistemic positions.

In a companion paper, we analyze the social mechanisms that may lead to epistemic injustice, based on two cases in automated government decision making [10]. Normally, a hearer is expected to trust the competence of the speaker unless there is evidence to the contrary. However, various forms of prejudice block these expected inferences, resulting in the speaker being unfairly denied access to knowledge or being disbelieved, misunderstood or ignored [9]. We present two cases of information systems that, when not developed and tested properly, prejudices that exist among developers or policymakers may end up in the system [10]. These cases illustrate that automated decision-making processes can be irrational and unjust.

In this paper, we focus on the logical aspects of such an analysis. We use a regular epistemic logic to characterize the structural properties of epistemic positions led by prejudice. The aim is to use logic, not to analyze idealized knowledge, but rather the deviant case, prejudice based on groups. Hence the contribution of this paper lies in the way and purpose modal logic is used and not technical advance. This paper is structured as follows. Section 2 provides background of epistemic injustice. Section 3 defines the epistemic logic. Section 4 provides a formal analysis. Section 5 concludes.

2 Epistemic Injustice

What is epistemic injustice? "Epistemic injustice refers to those forms of unfair treatment that relate to issues of knowledge, understanding, and participation in communicative practices." [11, preface]. We distinguish four specific types: (1) distributional injustice, (2) testimonial injustice, (3) hermeneutic injustice, and (4) content-focused injustice. Fricker [5] discusses in particular testimonial and hermeneutic epistemic injustice; distributional injustice is worked out in more detail later [6]. Content-focused epistemic injustice is introduced later by Dembroff and Whitcomb [3].

(1) Distributional injustice means that some groups unfairly do not have (access to) specific knowledge that they need to achieve some benefits or entitlements. This is unlike normal cases of advantage from education.

(2) Testimonial injustice relates to the relative trust in someone's capacity as a knower, for example, as a witness in a trial. An injustice of this kind can occur when someone is not believed or even ignored, because of group properties like their gender presentation, race, disability, or more generally, because of their identity [5].

(3) Hermeneutical injustice is related to how people understand their situation, and their ability to formulate statements about that situation. For example, before the 1970s, victims of sexual harassment had trouble describing in court the behavior of which they were the victim, because the concept had not yet been articulated. Legal procedures demanded physical evidence of abuse, which is hard to obtain.

(4) Finally, content-focused injustice is based on consensus in a group, or a common prejudice about the content of what is stated. For example, the Reagan administration deliberately rejected healthcare expert testimony on how to try and contain the spreading disease of HIV, because HIV was seen as the 'gay disease' [3].

Cases may display various types of epistemic injustice. When Marilyn Vos Savant, the person having the highest recorded IQ in the world, provided the *correct* answer to the Monty Hall problem in the column of Parade Magazine, tens of thousands of people (including many mathematicians and other academics) reacted by publicly rejecting *harshly* to what Vos Savant stated. Most of the reactions just considered her answer *unimaginable* to be correct, which falls into the category of content-focused injustice, but several reflected upon her being a woman as a reason for being wrong, which is a case of testimonial injustice. [1]

Essentially, cases of epistemic injustice display two aspects: *dependence* and *prejudice* (Figure 1). First, speaker a depends on hearer b to make a decision that will provide some benefit ψ. This decision ψ requires evidence φ, which must be provided by a. For example, b must decide whether to grant a subsidy. Formally, this decision depends on b accepting a's statement φ as true. Typically, statement φ is supported by written documents such as financial statements or tax returns. For instance, if a fails to provide proof of residence in the municipality where the subsidy is claimed, b will not believe that a is eligible for the subsidy and will not grant it. Second, hearer b, belonging to a privileged group D, may act unjustly due to prejudice against speaker a from group A (notation $A \leqslant D$). This prejudice leads to various *irrational epistemic positions* that influence the decision-making process. When we say "b doesn't believe a's statement φ," it may be due to several *wrong* reasons, primarily driven by prejudice. The four types of epistemic injustice mentioned can be categorized into distinct epistemic positions influenced by prejudice.

[1] See: https://priceonomics.com/the-time-everyone-corrected-the-worlds-smartest/

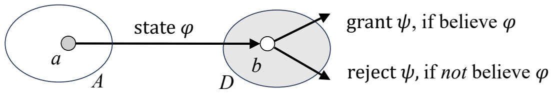


Fig. 1. Basic setting: speaker a from group A makes assertion φ to hearer b from group D. Speaker a *depends* on a decision by b to receive some benefit ψ. Decision to grant ψ requires that b believe φ. Members of group D are *prejudiced* against A.

3 Action-based Epistemic Logic

Epistemic injustice goes beyond the wrongful recognition of an individual's epistemic status, it also examines how this misrecognition can lead to unfair decision-making. Here, we conceptualize that as unsatisfactory decisions for a person making a request for action. In this section, we will explore both the epistemic and action-oriented elements of epistemic injustice by introducing an *action-based epistemic logic*.

The language $\mathscr{L}$ we use to address the types of epistemic injustice is the set of all formulas built recursively from atomic propositions $p \in Prop$ by negation $\neg\varphi$, conjunction $\varphi \wedge \psi$, individual knowledge modality $K_a\varphi$, individual belief modality $B_a\varphi$, common belief modality $C_G\varphi$, action modality $E_a\varphi$, and universal modality $\Box\varphi$, where $a \in \mathscr{I}$ is an element of a finite set of agents and $\emptyset \neq G \subseteq \mathscr{I}$.

The B-modality is a KD4-modality, K- and $\Box$-modalities are S5-modalities, and C_G-modality is a KD4-modality [4]. The action modality E_b is a T-operator, interpreted as agent b's *decisions*, from agency theory or STIT logic [1]. The dual $\hat{E}_b\psi$ is equal to $\neg E_b\neg\psi$. The following statements are theorems in our logic, where $a \in G$:

(KB) $K_a\varphi \to B_a\varphi$; (CB) $C_G\varphi \to B_a\varphi$.

Now we can define, for instance, the epistemic aspects of Fricker's notion of testimonial injustice, which involve the incorrect recognition of one's capacity as a knower, expressed as $K_a\varphi \wedge B_b\neg K_a\varphi$, and wrongly recognize agent a's credibility of knowledge as $K_a\varphi \wedge \neg B_bK_a\varphi$. In the following section we will provide a logical analysis of three of the four types of epistemic injustice. Hermeneutical injustice is left out, as we lack the formal tools to analyze the cognitive process of conceptualizing.

4 Towards a Formal Theory of Epistemic Injustice

As argued in [5, 6] *prejudices* – irrational unfounded *beliefs* about others based on group [7] – play a crucial role in epistemic injustice. Prejudice leads to the wrongful treatment of the right to be a knower. This section will present one assumption about the power of decision makers and three assumptions of *epistemic positions* involved in prejudice: a decision maker b, who, as hearer of a statement φ, doubts the credibility of the sender a and subsequently rejects a's request, especially when a is outside the privileged group D (see Figure 1).

Prejudice creates injustices due to the power of decision makers. We use φ to represent the evidence submitted by a and ψ for the requested decision to be made by b. So proposition $\Box(\psi \to \varphi)$ means that the evidence is a necessary condition for fulfilling the request. Here $E_b\psi \to B_b\varphi$ means that b believing evidence φ is a necessary condition for b to ensure ψ:

(RD) For all $b \in D$: $\Box(\psi \to \varphi) \to (E_b\psi \to B_b\varphi)$.

We introduce notation $A \leqslant D$ into the language $\mathscr{L}$ to represent prejudice: group A holds a disadvantaged *epistemic position* relative to group D. This symbol $\leqslant$ signifies a precondition of *morally wrong* beliefs between groups. So prejudice is a collective concept [7]. We have not introduced semantics for the expression $A \leqslant D$, however, it's worth noting that this is a possibility. For example, members of D may not trust members of A on certain topics (e.g. finance; law), where topics are represented as subsets of the set of proposition letters. Prejudice leads to various *irrational* beliefs underlying decisions. In our logic, prejudice $A \leqslant D$ involves three epistemic positions: [2]

(Ai) For all $a \in A$ and $b \in D$ with $A \leqslant D$: $B_b\varphi \rightarrow K_a\varphi$;

(Aii) For all $a \in A$ and $b \in D$ with $A \leqslant D$: $B_b\varphi \rightarrow B_bK_a\varphi$;

(Aiii) For all $b \in D$: $C_D\neg\varphi \rightarrow B_bE_bC_D\neg\varphi$.

Assumption (Ai) expresses that the beliefs of the privileged party take precedence over the knowledge of the disadvantaged party. When an individual lacks certain knowledge, this becomes a reason for the decision maker to disbelieve the information provided by this agent. As argued by Fricker, when a disadvantaged group lacks access to knowledge regarding a specific request or complaint—for instance, blue-collar workers lacking knowledge related to taxes—the request may be rejected based on the belief that the evidence is incorrect due to the agent's lack of tax knowledge, regardless of whether the evidence is factually true, as expressed by (Ai). The prejudice leads to this epistemic position $\neg K_a\varphi \rightarrow \neg B_b\varphi$.

Assumption (Aii) reflects a different epistemic position. The beliefs of the dominant group influence the perception about the knowledge of the disadvantaged group. This assumption captures another form of wrongful treatment of the right to be a knower. When a member b of the privileged group believes that an individual a lacks certain factual information, even when a possesses full knowledge, it results in the disbelief and rejection of a's claim. In the Robodebt case [8], blue-collar workers were well aware of their incomes, but the officers believed they were wrong because the data submitted by the workers did not match the predictions made by the IT system. The prejudice leads to this epistemic position $B_b\neg K_a\varphi \rightarrow \neg B_b\varphi$ from (Aii) and $\mathsf{D_B}$.

Assumption (Aiii) is about consensus in the privileged group. When a piece of information is part of the common ground for a group, every individual within that group believes it is impossible to revise such a common belief. While group prejudice may no longer be present, a new type of bias, known as confirmation bias, arises. This leads to the epistemic position illustrated by (Aiii). This morally wrong epistemic position still results in the wrongful treatment of the right to be a knower.

The reasoning behind distributional injustice, testimonial injustice, and content-focused injustice differs in three key aspects, respectively: distinctions in factual

---

[2] Given a structure $M = \langle W, \{R_a\}_{a\in\mathscr{I}}, \{S_a\}_{a\in\mathscr{I}}, \{\sim_a\}_{a\in\mathscr{I}}\rangle$ where R_a is an equivalence relation over W, S_a is a transitive and serial relation over W such that $S_a \subseteq R_a$, and $\sim_a$ is an equivalence relation over W. The truth conditions are defined as usual. $M, w \models K_a\varphi$ iff $R_a[w] \subseteq ||\varphi||$, $M, w \models B_a\varphi$ iff $S_a[w] \subseteq ||\varphi||$, $M, w \models C_G\varphi$ iff $D_G[w] \subseteq ||\varphi||$, $M, w \models E_a\varphi$ iff $\sim_a [w] \subseteq ||\varphi||$, $M, w \models \Box\varphi$ iff $W \subseteq ||\varphi||$, where $||\varphi|| = \{w \in W \mid M, w \models \varphi\}$. The frame conditions to validate (RD) and (Ai) – (Aiii) are as follows: (**RD**) $\forall wu \in W(wS_bu \rightarrow w \sim_b u)$; (**Ai**) $\forall wu \in W(wR_bu \rightarrow wS_au)$, if $A \leqslant D$, $a \in A$ and $b \in D$; (**Aii**) $\forall wuv \in W(wS_bu \wedge uR_av \rightarrow wS_bv)$, if $A \leqslant D$, $a \in A$ and $b \in D$; (**Aiii**) $\forall wuv \in W(wS_bu \wedge u \sim_b v \rightarrow wS_bv)$.

	Facts	Beliefs	Decisions	Inferential Elements
Distributional Injustice	$\Box(\psi \to \varphi)$ $\neg K_a \Box(\psi \to \varphi)$ $\neg K_a \varphi$	Weak: $\neg B_b \varphi$	$\neg E_b \psi$	Ai, $\neg K_a \varphi$ RD, $\Box(\psi \to \varphi)$, Weak
Testimonial Injustice	$\Box(\psi \to \varphi)$ $K_a \Box(\psi \to \varphi)$ $K_a \varphi$	Strong: $B_b \neg K_a \varphi$ Weak: $\neg B_b \varphi$	$\neg E_b \psi$	Aii, D_B, Strong RD, Weak
Content-focused Injustice	$\Box(\psi \to \varphi)$ $K_a \Box(\psi \to \varphi)$ $K_a \varphi$	Strong: $C_D \neg \varphi$ Weak$_1$: $B_b E_b C_D \neg \varphi$ Weak$_2$: $B_b E_b B_b \neg \varphi$ Weak$_3$: $\neg B_b \varphi$	$\neg E_b \psi$	Aiii, Strong CB, Weak$_1$ T_E, 4_B D_B, Weak$_2$ RD, Weak$_3$

Table 1

A classification of epistemic injustice (C1,C2,C4), where $a \in A$ and $b \in D$ with $A \leqslant D$.

information, beliefs of decision-makers about credibility of groups, and beliefs of decision-makers about credibility of content (Table 1).

Distributional injustice means that a, who is in a disadvantaged position, lacks knowledge of the evidence: $\neg K_a \varphi$. Given this fact and the assumption (Ai), it leads to *weak belief*: the decision maker b does not believe the submitted evidence: $\neg B_b \varphi$. This type of belief is considered *weak*, because it is derived from the other's lack of knowledge [12]. Following (RD), the decision doesn't fulfill the request: $\neg E_b \psi$.

The reasoning process for testimonial injustice follows a different path. Sender a does have knowledge of the evidence ($K_a \varphi$), but decision-maker b believes the sender lacks knowledge of the evidence ($B_b \neg K_a \varphi$). This is a *strong belief*, because it is assumed and not derived. It reflects *prejudice*, because it presupposes that *everyone* in a disadvantaged group A lacks knowledge about this topic , regardless of its actual veracity. From this strong belief, assumption Aii and axiom D, the weak belief $\neg B_b \varphi$ can be inferred. Ultimately, assumption RD leads to non-fulfillment of the request.

In content-focused injustice, confirmation bias occurs when every group member believes that a common consensus cannot be challenged, as illustrated by (Aiii). Specifically, when there is a common consensus $C_D \neg \varphi$ denying a piece of evidence, it results in all group members rejecting that piece of evidence, as shown in the last row of Table 1. On the one hand, as expressed in (Aiii), content-focused injustice is not agent-based. On the other hand, there are two notable cases of denial of the right to be a knower in content-focused injustice, which are worth to be discussed: (i) the individual case $B_b \neg \varphi$ and therefore $\neg B_b \varphi$, so the request is rejected, and also $B_b \neg K_a \varphi$, so the requester is denied in her right as a knower, and (ii) the group consensus case, $C_D \neg \varphi$ and therefore $\neg C_D \varphi$, so the request would be rejected by any official, but also $C_D \neg K_a \varphi$, so the requester is by consensus denied in the right as a knower.

5 Conclusions

This work introduces a formal framework for analyzing epistemic injustice through action-based epistemic logic. From a social epistemology viewpoint, we categorize epistemic injustice into four types of epistemic injustice: distributional, testimonial, hermeneutical and content-focused injustice [5, 3]. We characterizing the social setting in terms of *dependence* and *prejudice*, specifically decision rights and irrational beliefs. These irrational beliefs, as argued in [10], can be captured by various epistemic positions. Our modal logic characterizes three principles of epistemic positions influenced by prejudice in decision-making, showing how prejudice-driven logical relations lead from irrational beliefs to unjust decisions. This makes our modal logic valuable for analyzing the behaviors caused by epistemic injustice. In this paper, prejudice is characterized by epistemic positions. The current formalism lacks the tools to address either (C3) hermeneutical injustice or the interaction between group beliefs and group identities. So we leave all these to future work.

Acknowledgement

The authors gratefully acknowledge the valuable comments provided by the anonymous reviewers of AiML 2024. This work is supported by the *Fonds National de la Recherche* of Luxembourg through the project Deontic Logic for Epistemic Rights (OPEN O20/14776480).

References

[1] Belnap, N., M. Perloff and M. Xu, "Facing the future: agents and choices in our indeterminist world," Oxford University Press, 2001.

[2] Dam, C. v., *Ongekend onrecht - verslag van de parlementaire ondervraging kinderopvangtoeslag*, Tweede Kamer der Staten Generaal (2020).

[3] Dembroff, R. and D. Whitcomb, *Content-focused epistemic injustice*, in: T. S. Gendler, J. Hawthorne and J. Chung, editors, *Oxford Studies in Epistemology - Volume 7*, Oxford University Press, 2022 pp. 48–70.

[4] Fagin, R., J. Y. Halpern, Y. Moses and M. Vardi, "Reasoning about Knowledge," MIT Press, 1995.

[5] Fricker, M., "Epistemic injustice." Oxford University Press, Oxford, 2007.

[6] Fricker, M., *Epistemic justice as a condition of political freedom?*, Synthese **190** (2013), pp. 1317–1332.

[7] Hakli, R. and S. Negri, *Reasoning about collectively accepted group beliefs*, Journal of philosophical logic **40** (2011), pp. 531–555.

[8] Holmes, C., *Report of the royal commission into the robodebt scheme*, Report, Commonwealth of Australia (2023).

[9] Hulstijn, J., H. Dong and R. Markovich, *Analyzing epistemic injustice*, in: *LNGAI2024* (2024), pp. 1–14.

[10] Hulstijn, J., H. Dong and R. Markovich, *Epistemic injustice and government information systems: Lessons from two cases*, in: *EGOV2024* (2024), pp. 1–17.

[11] Kidd, I. J., J. Medina and G. Pohlhaus Jr., editors, "Handbook of Epistemic Injustice," Routledge, London, 2017, 1 edition.

[12] van Der Hoek, W., W. Jamroga and M. Wooldridge, *Towards a theory of intention revision*, Synthese **155** (2007), pp. 265–290.

Products of Horn Modal Logics

Aleksandr Gagarin [1]

National Research University Higher School of Economics, Moscow, Russia

Abstract

We study products of unimodal logics characterized by classes of Kripke frames defined by universal Horn formulas, classifying them with respect to the finite model property (FMP). Further, we show that products of modal logics defined only with variable-free axioms have the FMP. We also provide a partial result regarding products of logics with both Horn and variable-free axioms.

Keywords: Horn formula, product of modal logics, filtration

1 Introduction

A *Horn modal logic* is a modal logic characterized by the class of all Kripke frames satisfying several first-order *universal Horn clauses*. Unimodal Horn logics naturally fall into 4 types (cf. [7]); we refer to them as *transitive* (e.g., **K4**, **S4**), *reflexive-symmetric* (**K**, **KB**, **T**, **KTB**), *strong* (**K5**, **S5**), and *uniform* ($\mathbf{K} + \Diamond\Box p \to \Box^2 p$). Logics of the first two types are PSPACE-complete, and those of the other two have the polynomial model property and are coNP-complete [5] [7].

Some bimodal Horn logics are undecidable [7]. Apparently, there is no known decidability criterion for these.

The finite model property (FMP) is known for certain products of Horn logics, including $(\mathbf{K} + \Box p \to \Box^m p) \times \mathbf{S5}$ [3] and $(\mathbf{K} + \Box p \to \Box^m p) \times \mathbf{K}_m$ [8]. However, some other products (e.g., $\mathbf{K4} \times \mathbf{K4}$) are undecidable [4].

In this paper we classify all products of unimodal Horn logics with respect to the FMP. We deduce from [4] that a product of two *transitive* Horn logics is undecidable and does not have the FMP. By employing the *filtration via bisimulation* technique of [8], we establish that all other products of unimodal Horn logics have the FMP.

We also extend this result to products of Horn logics with additional variable-free axioms, provided the two Horn logics are not *uniform*. This includes the following special case: if λ_1 and λ_2 are variable-free, then $(\mathbf{K} + \lambda_1) \times (\mathbf{K} + \lambda_2)$ has the FMP and is decidable.

[1] I would like to thank my supervisor Valentin Shehtman for his constant support. This research was supported by the Academic Fund Program at HSE University (grant № 23-00-022).

2 Preliminaries

Basics. We consider the basic unimodal (**ML**) and bimodal ($\mathbf{ML}_2$) propositional languages; only normal logics are considered. We use the standard Kripke semantics; unimodal (bimodal) Kripke frame are called *frames (2-frames)* for short; the logic characterized by a class of (2-)frames $\mathcal{C}$ is denoted $\mathrm{Log}\,\mathcal{C}$. First-order formulas with a single binary predicate R are interpreted over frames; the logic characterized by the class of all frames modeling a first-order theory Γ is denoted $\mathbf{K}(\Gamma)$.

Products. The *product* of frames (W_1, R_1) and (W_2, R_2) is the 2-frame $(W_1 \times W_2, R_1', R_2')$, where: $R_1' := \{((x,z),(y,z)) : (x,y) \in R_1, z \in W_2\}$ and $R_2' := \{((z,x),(z,y)) : z \in W_1, (x,y) \in R_2\}$. The *product* of classes of frames $\mathcal{C}_1$ and $\mathcal{C}_2$ is the class of 2-frames $\mathcal{C}_1 \times \mathcal{C}_2 := \{\mathfrak{F}_1 \times \mathfrak{F}_2 : \mathfrak{F}_i \in \mathcal{C}_i\}$. The *product* of logics L_1 and L_2 is the bimodal logic $L_1 \times L_2 := \mathrm{Log}(\mathrm{Fr}\, L_1 \times \mathrm{Fr}\, L_2)$, where $\mathrm{Fr}\, L_i$ is the class of all frames for L_i.

The *commutator* of L_1 and L_2, denoted $[L_1, L_2]$, is the minimal bimodal logic containing the axioms $\Diamond_1\Box_2 p \to \Box_2\Diamond_1 p$ and $\Box_1\Box_2 p \leftrightarrow \Box_2\Box_1 p$ and extending $[\Box_1/\Box]L_1 \cup [\Box_2/\Box]L_2$.

Filtration. For a subformula-closed set $\Sigma \subseteq \mathbf{ML}$, a Σ-*filtration* of a Kripke model $(W, R, \mathfrak{V})$ is a Kripke model of the form $(W/{\sim}, S, \mathfrak{V}^{\sim})$ satisfying the following conditions:
(1) if $x \sim y$ and $\varphi \in \Sigma$, then $(\mathfrak{M}, x) \models \varphi \iff (\mathfrak{M}, y) \models \varphi$,
(2) $S \supseteq R^{\sim}$, where $R^{\sim} := \{([x],[y]) : xRy\}$,
(3) if $[x]S[y]$, $(\mathfrak{M}, x) \models \Box\varphi$, and $\Box\varphi \in \Sigma$, then $(\mathfrak{M}, y) \models \varphi$, and
(4) $\mathfrak{V}^{\sim}(p) := [\mathfrak{V}(p)]$.
A filtration satisfying $S = R^{\sim}$ is called a *minimal filtration.* A logic L *admits filtration with respect to a frame* $\mathfrak{F}$ if for any valuation $\mathfrak{V}$ on $\mathfrak{F}$ and any finite subformula-closed set $\Sigma \subseteq \mathbf{ML}$ there exists a finite Σ-filtration of $(\mathfrak{F}, \mathfrak{V})$ based on a frame for L. A logic L *admits fitlration* if it admits filtration with respect to each frame for L. Similar definitions apply to bimodal logics and Kripke models.

Trees. A frame (W, R) is a *tree* with a root $w \in W$ if for any $u \in W$ there exists a unique R-path from w to u.

Pseudo-finitness. A frame (W, R) is *s-pseudo-finite* if there exists an equivalence relation $\sim$ on W such that $|W/{\sim}| \le s$ and $\sim \circ R \circ \sim\, = R$. A frame is *pseudo-finite* if it is s-pseudo-finite for some s.

3 Horn Logics

Definition 3.1 A *Horn clause* is a first-order sentence of the following form:

$$\forall \overrightarrow{x} \left(\bigwedge_{s=1}^{m} x_{i_s} R x_{j_s} \to x_{i_0} R x_{j_0} \right).$$

A *Horn theory* is a set of Horn clauses. For a Horn theory Γ, the *Horn Γ-closure* of a frame $\mathfrak{F} = (W, R)$ is the frame $\mathfrak{F}^{\Gamma} := (W, R^{\Gamma})$, where R^{Γ} is the minimal relation containing R and satisfying $(W, R^\Gamma) \models \Gamma$. A *Horn logic* is a logic of the form $\mathbf{K}(\Gamma)$, where Γ is a Horn theory.

Definition 3.2 A *tree-clause* corresponding to a finite tree $(W, \widetilde{R})$ and nodes $u_0, v_0 \in W$ is the Horn clause $\forall \overrightarrow{x} \left(\bigwedge_{(u,v)\in\widetilde{R}} x_u R x_v \to x_{u_0} R x_{v_0} \right)$. Its *type* is the pair $(n, m) \in (\mathbb{Z}_{\geq 0})^2$ such that $w\widetilde{R}^n u_0$ and $w\widetilde{R}^m v_0$, where w is the least common ancestor of u_0 and v_0. A *tree-theory* is a set of tree-clauses.

Example 3.3 For any $n, m \geq 0$, the logic $\mathbf{K} + \Diamond^n \Box p \to \Box^m p$ is Horn, as its only axiom is a Sahlqvist modal equivalent of $\forall x, y, z\, (xR^n y \wedge xR^m z \to yRz)$. The latter is a tree-clause of type (n, m).

As shown in [6], tree-clauses have Sahlqvist modal equivalents and are the only (up to equivalence) Horn clauses having modal equivalents at all.

Proposition 3.4 *Every Horn logic coincides with* $\mathbf{K}(\Gamma)$ *for some (possibly infinite) tree-theory* Γ.

Lemma 3.5 ([3], [1]) *Let* Γ_1, Γ_2 *be Horn theories; set* $L_i := \mathbf{K}(\Gamma_i)$. *If* $\varphi \notin [L_1, L_2]$, *then there exist trees* $\mathfrak{T}_i$ *with roots* w_i *such that* $\left(\mathfrak{T}_1^{\Gamma_1} \times \mathfrak{T}_2^{\Gamma_2}, (w_1, w_2) \right) \not\models \varphi$. *In particular,* $[L_1, L_2] = L_1 \times L_2$.

Remark 3.6 It follows from Lemma 3.5 that the product $\mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$ coincides with $\mathrm{Log}\,\{\mathfrak{F}_1 \times \mathfrak{F}_2 : \mathfrak{F}_1 \models \Gamma_1, \mathfrak{F}_2 \models \Gamma_2\}$.

Remark 3.7 Proposition 3.4 and Lemma 3.5 also hold for polymodal Horn logics, with "n-trees" substituted for trees; cf. [3] [6].

4 Classification of Unimodal Horn Logics

We divide unimodal Horn logics into 4 classes, similarly to the classification used in [7].

Definition 4.1 A tree-theory Γ is —

- *reflexive-symmetric* if the type of each clause in Γ is $(0, 0)$, $(1, 0)$, or $(0, 1)$;
- *transitive* if all clauses in Γ have types of the form $(0, m)$, and at least one of them is with $m > 1$;
- *uniform* if all clauses in Γ have types of the form $(n, n + 1)$, and at least one of them is with $n > 0$;
- *strong* in all other cases.

Lemma 4.2 *If* Γ *is a strong tree-theory, then there exist integers* $d, s \geq 0$ *such that for any frame* $(W, R) \models \Gamma$ *and any* $u \in W$ *satisfying* $R^{-d}(u) \neq \emptyset$ *the restriction of* R *to* $\left\{ x \in R^{<\infty}(u) : R^d(x) \neq \emptyset \right\}$ *is* s*-pseudo-finite.*

We will need the following properties of logics of specific types:

Lemma 4.3 *Any reflexive-symmetric, transitive, or strong tree-theory is equivalent to some finite subtheory.*

Lemma 4.4 *If* Γ *is a transitive tree-theory, then* $\mathbf{K}(\Gamma)$ *admits filtration.*

Lemma 4.4 is proven by appropriately generalizing the proof of FMP for $\mathbf{K} + \Box p \to \Box^m p$ from [2].

5 Products without the FMP

Our main result is as follows.

Theorem 5.1 *Let Γ_1 and Γ_2 be tree-theories.*

1) If both Γ_1 and Γ_2 are transitive, then $\mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$ is undecidable and does not have the fmp.

2) If Γ_1 (or Γ_2) is not transitive, then $\mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$ has the fmp. [2]

First, we derive the negative part of the claim from the following fact.

Lemma 5.2 ([4]) *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be classes of transitive frames both containing frames of infinite depth. Then $\mathrm{Log}(\mathcal{C}_1 \times \mathcal{C}_2)$ is undecidable.*

Proof of Theorem 5.1(1) For $i \in \{1,2\}$ choose an integer $l_i > 0$ such that $\mathcal{C}_i := \{(W, R^{l_i}) : (W,R) \models \Gamma_i\}$ contains only transitive frames. Observe that Lemma 5.2 is applicable to $\mathcal{C}_1, \mathcal{C}_2$. The map $\mathbf{ML}_2 \to \mathbf{ML}_2$ replacing each occurrence of $\Box_i$ with $\Box_i^{l_i}$ is a reduction from $\mathrm{Log}(\mathcal{C}_1 \times \mathcal{C}_2)$ to $\mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$, hence the undecidability. By Lemma 4.3, $\mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$ is finitely axiomatizable, hence the lack of FMP. $\Box$

6 Products with the FMP

We outline a proof of Theorem 5.1(2) similar to the reasoning in [8].

Definition 6.1 *A relation $E \subseteq W_1 \times W_2$ is a* temporal bisimulation *between the frames (W_1, R_1) and (W_2, R_2) if $R_2 \circ E = E \circ R_1$ and $R_1 \circ E^{-1} = E^{-1} \circ R_2$. Two relations S_1 and S_2 on the same set* strongly commute *if $S_1 \circ S_2 = S_2 \circ S_1$ and $S_1 \circ S_2^{-1} = S_2^{-1} \circ S_1$. A frame (W,R)* admits temporal bisimulation *if for any finite $W/{\sim}$ there exists an equivalence relation of finite index $\approx$ strongly commuting with R such that $\approx \subseteq \sim$.*

Lemma 6.2 *Let Γ be a tree-theory and E a temporal bisimulation between $\mathfrak{F}$ and $\mathfrak{G}$; then E is also a temporal bisimulation between $\mathfrak{F}^\Gamma$ and $\mathfrak{G}^\Gamma$.*

Lemma 6.3 ("filtration via bisimulation") *Let Γ_1, Γ_2 be tree-theories and (W, R_1, R_2) a 2-frame such that:*

(1) $(W, R_i) \models \Gamma_i$ and $\mathbf{K}(\Gamma_i)$ admits filtration with respect to (W, R_i), for $i \in \{1,2\}$;

(2) R_1 and R_2 strongly commute; and

(3) (W, R_1) admits temporal bisimulation.

Then $[\mathbf{K}(\Gamma_1); \mathbf{K}(\Gamma_2)]$ admits filtration with respect to (W, R_1, R_2).

Proof outline Fix a valuation $\mathfrak{V}$ on W and a finite subformula-closed set $\Sigma \subseteq \mathbf{ML}_2$. It may be shown (using (1) only) that there exists an equivalence relation $\sim$ such that $(W/{\sim}, (R_1^\sim)^{\Gamma_1}, (R_2^\sim)^{\Gamma_2}, \mathfrak{V}^\sim)$ is a finite Σ-filtration of $(W, R_1, R_2, \mathfrak{V})$.

Let $\approx$ be the union of all relations containing in $\sim$ and strongly commuting with R_1. Since (W, R_1) admits temporal bisimulation, $\approx$ is an equivalence

[2] Decidability does not necessarily follow, as some uniform Horn logics are not finitely axiomatizable.

relation of finite index. We have $\approx \circ R_1 = R_1 \circ \approx$; thus $R_1^{\approx}$ and $R_2^{\approx}$ strongly commute; hence $(R_1^{\approx})^{\Gamma_1}$ and $(R_2^{\approx})^{\Gamma_2}$ strongly commute by Lemma 6.2; therefore $(W/{\approx}, (R_1^{\approx})^{\Gamma_1}, (R_2^{\approx})^{\Gamma_2})$ is a frame for $[\mathbf{K}(\Gamma_1); \mathbf{K}(\Gamma_2)]$. □

Definition 6.4 A frame is a *pseudo-tree of height* 1 if it is pseudo-finite. A frame (W, R) is a *pseudo-tree of height* $h > 1$ if W can be represented in the form $W_0 \cup \bigsqcup_{j \in J} W_j$ such that:
(1) $|W_0 \cap W_j| = 1$ for each $j \in J$;
(2) $R = R|_{W_0} \cup \bigsqcup_{j \in J} R|_{W_j}$;
(3) $(W_0, R|_{W_0})$ is a pseudo-tree of height $h - 1$; and
(4) for some s, all $(W_j, R|_{W_j})$ are s-pseudo-finite.

Lemma 6.5 *Every pseudo-tree admits temporal bisimulation.*

Proof outline By induction on height, in the same way as for trees of finite height in [8]. □

Lemma 6.6 *Let Γ be a strong tree-theory and $\mathfrak{T}$ a tree. Then $\mathfrak{T}^\Gamma$ coincides with the Horn Γ-closure of a pseudo-tree.*

Proof Follows from Lemma 4.2. □

Definition 6.7 The d-truncation of a pointed frame (W, R, w) is the frame $(W_d, R|_{W_d}, w)$, where $W_d := R^{\leq d}(w)$. Pointed frames (W, R, w) and (V, S, v), are *d-indistinguishable* if their d-truncations are isomorphic.

Lemma 6.8 *Let Γ be a reflexive-symmetric or uniform tree-theory, $d > 0$. Then every tree $\mathfrak{T}$ with root w has a subtree $\mathfrak{T}'$ of finite height such that $(\mathfrak{T}^\Gamma, w)$ and $(\mathfrak{T}'^\Gamma, w)$ are d-indistinguishable.*

Proof of Theorem 5.1(2) Set $L := \mathbf{K}(\Gamma_1) \times \mathbf{K}(\Gamma_2)$. Consider $\varphi \notin L$; let d be its modal depth. Choose trees $\mathfrak{T}_1, \mathfrak{T}_2$ as in Lemma 3.5. Now for each $i \in \{1, 2\}$, depending on the type of Γ_i, apply one of Lemmas 4.4, 6.6, or 6.8. It follows that there exist frames $\mathfrak{F}_i$ d-indistinguishable from (or coinciding with) $\mathfrak{T}_i^{\Gamma_i}$ such that Lemma 6.3 is applicable to $\mathfrak{F}_1 \times \mathfrak{F}_2$. □

7 Adding Variable-Free Axioms

Theorem 7.1 *Let Γ_1, Γ_2 be reflexive-symmetric, transitive, or strong tree-theories, and λ_1, λ_2 variable-free formulas.*

1) If both Γ_1 and Γ_2 are transitive, and $\mathbf{K}(\Gamma_i) + \lambda_i \nvdash \Box^n \bot$ for every $i \in \{1, 2\}$ and $n > 0$, then $(\mathbf{K}(\Gamma_1) + \lambda_1) \times (\mathbf{K}(\Gamma_2) + \lambda_2)$ is undecidable and lacks the fmp.

2) In all other cases, $(\mathbf{K}(\Gamma_1) + \lambda_1) \times (\mathbf{K}(\Gamma_2) + \lambda_2)$ is decidable and has the fmp.

Note that Theorem 7.1 does not cover *uniform* tree-theories; that case remains unclear. The proof is the same as for Theorem 5.1, except that we use the following lemma instead of Lemma 6.8.

Lemma 7.2 *Let Γ be a reflexive-symmetric tree-theory, λ a variable-free formula, $\mathfrak{T} = (W, R)$ a tree with root w such that $\mathfrak{T}^\Gamma \models \lambda$, and $d > 0$ an integer. Then there exists a pseudo-tree $\mathfrak{F}$ with a node v such that*

(1) $\mathfrak{F}^\Gamma \models \lambda$*, and*
(2) $(\mathfrak{T}^\Gamma, w)$ *and* $(\mathfrak{F}^\Gamma, v)$ *are* d*-indistinguishable.*

Proof outline By Lemma 4.3, we can assume that Γ is finite. One can derive that, for some finite $W/{\sim}$, the minimal filtration $(W/{\sim}, (R^\Gamma)^\sim)$ is a frame for both Γ and λ.

Set $x \approx y$ if (1) $x \sim y$ and (2) either $x = y$ or the least common ancestor of x and y is not in $R^{\leq d}(w)$. Note that $(W/{\approx}, R^\approx)$ is a pseudo-tree and that $(R^\approx)^\Gamma = (R^\Gamma)^\approx$. One can show that (W, R^Γ, w) and $(W/{\approx}, (R^\Gamma)^\approx, [w])$ are d-indistinguishable. □

References

[1] Gabbay, D., A. Kurucz, F. Wolter and M. Zakharyaschev, "Many-dimensional modal logics: theory and applications," Studies in Logic and the Foundations of Mathematics, Elsevier, 2003.

[2] Gabbay, D. M., *A general filtration method for modal logics*, Journal of Philosophical Logic **1** (1972), pp. 29–34.

[3] Gabbay, D. M. and V. Shehtman, *Products of modal logics, part 1*, Logic Journal of the IGPL **6** (1998), pp. 73–146.

[4] Gabelaia, D., A. Kurucz, F. Wolter and M. Zakharyaschev, *Products of 'transitive' modal logics*, Journal of Symbolic Logic **70** (2005), pp. 993–1021.

[5] Hemaspaandra, E. and H. Schnoor, *On the complexity of elementary modal logics*, in: *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science*, 2008, pp. 349–360.

[6] Kikot, S., *On modal definability of Horn formulas*, in: *Proceedings of Topology, Algebra and Categories in Logic*, 2011, pp. 175–178.

[7] Michaliszyn, J., J. Otop and E. Kieroński, *On the decidability of elementary modal logics*, ACM Trans. Comput. Logic **17** (2015).

[8] Shehtman, V., *Filtration via bisimulation*, in: *Advances in Modal Logic vol. 5*, 2005, pp. 289–308.

Modal Models in the Premodal Language

Daniel Gonzalez [1]

University of California, Berkeley

Abstract

We introduce a series of models of various modal logics built out of structures of ultrafilters on the algebra of premodal propositions. Given a pointed modal model, we give a general method for constructing a bisimilar pointed model using trees of indexed ultrafilters. At the end, we examine philosophical applications to accounts of metaphysical necessity.

Keywords: semantics of modal logic, stone duality, metaphysics.

1 Basic Definitions

Our premodal language $\mathcal{L}$ is built from a countable set of atomic sentence letters $\text{At} = \{p_n : n \in \mathbb{N}\}$, Boolean connectives $\neg, \wedge, \vee, \top, \bot$, and parenthesis $(,)$. We inductively define the set of all premodal formulas ϕ in the standard way, and define $\phi \to \psi$ as shorthand for $\neg\phi \vee \psi$ and $\phi \leftrightarrow \psi$ as shorthand for $(\phi \to \psi) \wedge (\psi \to \phi)$. The main algebra we will be considering is the Lindenbaum-Tarski algebra of propositional logic $\mathcal{B} = (\{[\phi]_\equiv : \phi \text{ a premodal formula}\}, \leq)$, where $\equiv$ is classical logical equivalence and $[\phi]_\equiv \leq [\psi]_\equiv$ iff $\phi \to \psi$ is a tautology. We define $1 = [\top]_\equiv$ and $0 = [\bot]_\equiv$, and lattice operations $\wedge, \vee$, and $\neg$ in the expected way.

An *ultrafilter* on $\mathcal{B}$ is a set $U \subseteq \mathcal{B}$ such that $1 \in U$, $0 \notin U$, $p \in U$ and $q \in U$ implies $p \wedge q \in U$, and either $p \in U$ or $\neg p \in U$ for all $p \in \mathcal{B}$. Clearly, we have if $p \in U$ and $p \leq q$, then $q \in U$. We let $\mathcal{U}$ be the set of all ultrafilters on $\mathcal{B}$.

Note that $|\mathcal{B}| = \aleph_0$ and $|\mathcal{U}| = 2^{\aleph_0}$.

In the modal language $\mathcal{L}_\Box$, we add a unary operator $\Box$ to the language, and define $\Diamond\phi$ as $\neg\Box\neg\phi$.

For a given pointed modal model $\mathcal{M}, w$, we define the *diagram* of $\mathcal{M}, w$, denoted $Diag(\mathcal{M}, w)$, as $\{[\phi]_\equiv : \phi \text{ is a premodal formula and } \mathcal{M}, w \models \phi\}$. Clearly, $Diag(\mathcal{M}, w)$ is an ultrafilter on $\mathcal{B}$. It will be important for finding bisimilar pointed modal models in the models we construct.

For a modal model $\mathcal{M}$ and logic L, we say that $\mathcal{M}$ has the *bisimulation property for L* if for every pointed model $\mathcal{N}, w$ such that L is valid on the frame of $\mathcal{N}$, there is a world $v \in \mathcal{M}$ and a bisimulation between $\mathcal{N}, w$ and

[1] daniel_gonzalez@berkeley.edu

$\mathcal{M}, v$. We say that $\mathcal{M}$ has the κ-restricted bisimulation property for L if the above condition holds for models $\mathcal{N}$ of size at most κ.

2 A Model of K45

We can construct models of K45, KD45, and S5 in a relatively straightforward manner that have the bisimulation property for their corresponding logics.

We define the model $\mathcal{M}_{\mathrm{K45}} = (W_{\mathrm{K45}}, R_{\mathrm{K45}}, V_{\mathrm{K45}})$ as follows:

- $W_{\mathrm{K45}} = \mathcal{U} \times \mathcal{P}(\mathcal{U})$
- For worlds $(U, S), (V, T) \in W_{\mathrm{K45}}$, we have $(U, S)\ R_{\mathrm{K45}}\ (V, T)$ iff $S = T$ and $V \in S$
- For $p_i \in \mathrm{At}$, we have $V_{\mathrm{K45}}(p_i) = \{(U, S) \in W_{\mathrm{K45}} : [p_i]_{\equiv} \in U\}$. We lift V_{K45} to a function from $\mathcal{L}_{\Box}$ to $\mathcal{P}(W_{\mathrm{K45}})$ in the standard way.

Theorem 2.1 *$\mathcal{M}_{K45}$ is a model of K45.*

Proof. Note that $\mathcal{M}_{\mathrm{K45}}$ is a Kripke model, so it satisfies K. Furthermore, since $=$ is transitive and Euclidean, the accessibility relation R_{K45} is transitive and Euclidean. Thus, $\mathcal{M}_{\mathrm{K45}}$ also satisfies 4 and 5. □

Theorem 2.2 *Suppose $\mathcal{N}, w$ is a pointed modal model whose accessibility relation is transitive and Euclidean. Then, there is some $(U, S) \in W_{K45}$ such that $\mathcal{M}_{K45}, (U, S)$ is bisimilar to $\mathcal{N}, w$.*

Proof. Our desired world (U, S) will be given by setting $U = Diag(\mathcal{N}, w)$ and $S = \{D \in \mathcal{U} : \exists v \in \mathcal{N}(wRv \text{ and } D = Diag(\mathcal{N}, v))$.

Our desired bisimulation E is given by $vE(D, S)$ iff $D = Diag(\mathcal{N}, v)$.

The atomic harmony condition is easy to check. For the zig and zag conditions, it suffices to show that the worlds that w can access are exactly the worlds that w's successors can access.

By transitivity, we know that w can access all the worlds its successors can access, and by Euclideanness, we know that if w can access a world, all its successors can access that worlds too. Thus, w can access a world v if and only if every successor of w can also access v. Since this is also the case on the side of $\mathcal{M}_{\mathrm{K45}}, (U, S)$, the zig and zag conditions follow. □

Corollary 2.3 *$\mathcal{M}_{K45}$ has the bisimulation property for K45.*

Note that the cardinality of W_{K45} is $2^{2^{\aleph_0}}$.

Models of KD45 and S5

We can easily modify the above construction to give models $\mathcal{M}_{\mathrm{KD45}}$ and $\mathcal{M}_{\mathrm{S5}}$ that have the bisimulation property for their corresponding logics. In particular, we set:

- $W_{\mathrm{KD45}} = \mathcal{U} \times (\mathcal{P}(\mathcal{U}) - \emptyset)$
- $W_{\mathrm{S5}} = \{(U, S) \in \mathcal{U} \times \mathcal{P}(\mathcal{U}) : U \in S\}$
- $R_{\mathrm{KD45}} = R_{\mathrm{S5}} = R_{\mathrm{K45}}$ and $V_{\mathrm{KD45}} = V_{\mathrm{S5}} = V_{\mathrm{K45}}$

3 Models of K

Unfortunately, it is not possible to make a single modal model with the bisimulation property for K. However, we can use a similar strategy to build models of K that have the κ-restricted bisimulation property for K for any desired cardinal κ.

First, we define a *κ-indexed ultrafilter tree* as a nonempty set $T \subseteq (\mathcal{U} \times \kappa)^{<\omega}$ such that for any node $b \in T$, every initial segment of b is also in T, and there exists a root (U, α) such that $b(0) = (U, \alpha)$ for every nonempty node $b \in T$. If (U, α) is the root of T, we denote U as $rt(T)$. We label as T^+ the set $\{(n+1, (U, \alpha)) : (n, (U, \alpha)) \in T\}$. Although T^+ is not a tree, it is important for defining our accessibility relation.

Now, we can define the model $\mathcal{M}_{\mathrm{K},\kappa}$ as follows:

- $W_{\mathrm{K},\kappa} = \{T : T \text{ is a } \kappa\text{-indexed ultrafilter tree}\}$
- $T_1\ R_{\mathrm{K},\kappa}\ T_2$ iff $T_2^+ \subseteq T_1$
- For $p_i \in \mathrm{At}$, we have $V_{\mathrm{K},\kappa}(p_i) = \{T \in W_{\mathrm{K},\kappa} : [p_i]_{\equiv} \in rt(T)\}$. We lift $V_{\mathrm{K},\kappa}$ to a function from $\mathcal{L}_\Box$ to $\mathcal{P}(W_{\mathrm{K},\kappa})$ in the standard way.

Theorem 3.1 *Let κ be an infinite cardinal. Then $\mathcal{M}_{K,\kappa}$ has the κ-restricted bisimulation property for K.*

Proof. Suppose $\mathcal{N}, w = (W, R, V), w$ is a pointed modal model of size at most κ. Let $\overrightarrow{\mathcal{N}} = (\overrightarrow{W}, \overrightarrow{R}, \overrightarrow{V})$ be the tree unravelling of $\mathcal{N}$ from w. We will construct a tree T such that $\mathcal{M}_{\mathrm{K},\kappa}, T$ is bisimilar to $\overrightarrow{\mathcal{N}}, \langle w \rangle$. It follows that $\mathcal{M}_{\mathrm{K},\kappa}, T$ is bisimilar to $\mathcal{N}, w$.

Fix an injection $f : W \to \kappa$. Now define a function $g : W \to \mathcal{U} \times \kappa$ so that $g(v) = (Diag(v), f(v))$ for all $v \in W$. We now define T_w as the following tree:

$$T_w = \{s : \exists h \in \overrightarrow{W}\ s = h \circ g\}$$

In other words, T_w looks exactly like the tree unravelling of $\mathcal{N}$ starting at w. Our bisimulation is now defined by carrying this construction out on all the remaining worlds in the submodel of $\mathcal{N}$ generated by w, and can be lifted to the tree unravelling of $\mathcal{N}$ from w in the obvious way.

Since $\langle w \rangle$ is always in $\overrightarrow{W}$, we always have that the root of T_w is $(Diag(w), f(w))$. Thus, it is easy to see that atomic harmony holds. It is also easy to check that zig and zag hold between the tree unravelling of $\mathcal{N}$ and the model $\mathcal{M}_{\mathrm{K},\kappa}$. Thus, we have that our relation is a bisimulation between the tree unravelling of $\mathcal{N}$ from w and the pointed model $\mathcal{M}_{\mathrm{K},\kappa}, T_w$. Thus, our model $\mathcal{M}_{\mathrm{K},\kappa}, T_w$ is bisimilar to $\mathcal{N}, w$. $\Box$

Intuitively, we can think of each indexed ultrafilter tree as encoding the structure of the indexed ultrafilter trees that are in its generated submodel. In fact, we could have just defined each indexed ultrafilter tree as an entire modal model instead of a single world. However, that is less interesting because our model $\mathcal{M}_{\mathrm{K},\kappa}$ is a single, giant modal model.

Note that the cardinality of $\mathcal{M}_{\mathrm{K},\kappa}$ is $\max\{2^{2^{\aleph_0}}, 2^\kappa\}$.

4 Further Comments on These Constructions

The construction of the model $\mathcal{M}_{\mathrm{K},\kappa}$ in particular is similar to tree unravelling. The model $\mathcal{M}_{\mathrm{K45}}$ is a very natural model, but it is hard to think of a way to generalize it that works. Initially, the author considered a simpler (unindexed) tree-based model like $\mathcal{M}_{\mathrm{K},\kappa}$. Unfortunately, an unindexed model does not contain enough bisimulation types; in particular, if a modal model contains two worlds satisfying the same premodal formulas and are accessed by the same worlds but can access different worlds, there is no bisimulation between this model and an unindexed model.

It is not a trivial fact that the logic K45 and its extensions have models that have the bisimulation property. It is impossible to make a model of K with the full bisimulation property, since otherwise we would be able to construct an injection from its powerset into itself by finding enough new bisimulation types.

5 Philosophical Applications

The above constructions should be particularly interesting for philosophers who want a strong, Ersatz account of metaphysical necessity. In philosophizing about what metaphysically possible worlds are, philosophers have sorted themselves into two main camps: modal realism and Ersatzism. In particular, Ersatzists endorse the view that a possible world is some ersatz object, such as a consistent set of propositions or maximal state of affairs. [1] It is tempting for a mathematically rigorous Ersatzist to identify the set of all possible worlds with the canonical model for a certain modal logic, such as S5, or some appropriate subset of the canonical model (with the atomic proposition letters interpreted appropriately). However, this view makes our account of metaphysical modality circular: we are defining metaphysical necessity in terms of possible worlds, and then defining possible worlds as sets of sentences that include the $\Box$ operator.

The advantage of the above models, especially the model $\mathcal{M}_{\mathrm{S5}}$, is that an Ersatz metaphysician can identify metaphysically possible worlds with a construction based on sets of premodal propositions. For many philosophers, identifying the set of metaphysically possible worlds with a construction such as $\mathcal{M}_{\mathrm{S5}}$ is ontologically parsimonious. For example, if a philosopher is a mathematical Platonist, she already has everything she needs (provided she can find an appropriate way to interpret the atomic proposition letters). Additionally, because of the bisimulation property, $\mathcal{M}_{\mathrm{S5}}$ can mimic any alternative account of metaphysically possible worlds a metaphysician would be interested in.

Some may object that the circularity problem does not completely disappear, and I admit that this is not yet a complete solution to a coherent Ersatz account of metaphysically possible worlds. But it is a step in the right direction.

It would be interesting to try to apply the constructions in this paper to other kinds of modality, such as time or belief. In any case, the constructions of such large, general models from the premodal language is interesting in itself, and may have more applications down the road.

References

[1] Lewis, D. K., "On the Plurality of Worlds," Wiley-Blackwell, Malden, Mass., 1986.

Canonical Extensions of Fuzzy Algebras

Alexander Kurz

Chapman University, Orange CA, USA

Apostolos Tzimoulis

Vrije Universiteit, Amsterdam, The Netherlands

1 Introduction

This note investigates the algebraic semantics of fuzzy modal logics from the point of view of quantale-enriched category theory. The set of truth-values (distances) is given by a complete lattice Ω with a (non-commutative) multiplication. Notions such as set, function, relation, downset, upset, (co)limit will be "internalized" in Ω. Algebras are equipped with a metric, functions are distance non-increasing, relations are fuzzy and (co)limits include not only meets and joins but unary modalities known as power and tensor.

Our main contribution is to show that the canonical extensions [3] known in order theory generalize to the setting of quantale-enriched category theory. [1]

2 Preliminaries

A **quantale** $(\Omega, \sqsubseteq, \bigsqcup, e, \cdot,)$ is a complete join semilattice $(\Omega, \sqsubseteq, \bigsqcup)$ and a monoid $(\Omega, e, \cdot)$ in which multiplication distributes over joins. We write top as $\top$ and bottom as $\bot$. Since Ω is complete it also has meets $\bigsqcap$. Multiplication has a left-residual $\triangleleft$ and the right-residual $\triangleright$ defined as $b \sqsubseteq a \triangleright c \Leftrightarrow a \cdot b \sqsubseteq c \Leftrightarrow a \sqsubseteq c \triangleleft b$.

Examples. (a) The *two-chain* $2 = \{0 \sqsubseteq 1\}$ is a commutative quantale.
(b) The *Lawvere quantale* $[0, \infty]$ is a subset of the extended real numbers [10]. It is ordered by $\geq$ with top $\top = 0$ and has $+$ as multiplication. The residual is truncated minus $a \triangleleft b = a \dot{-} b$.
(c) The *quantale of languages* $\mathcal{P}(\Sigma^*)$ over an "alphabet" Σ has as elements subsets of the set Σ^* of finite words over Σ . Multiplication is $L \cdot L' = \{vw \mid v \in L, w \in L'\}$ where vw denotes the concatenation of the words, residuals are $L \triangleright M = \{w \in \Sigma^* \mid \forall v \in L \,.\, vw \in M\}$ and $M \triangleleft L = \{w \in \Sigma^* \mid \forall v \in L \,.\, wv \in M\}$.

Quantale Spaces. An Ω-space [2] consists of a set X and a function $X(-,-) : X \times X \to \Omega$ such that $e \sqsubseteq X(x,x)$ for all $x \in X$ and $X(x,y) \cdot X(y,z) \sqsubseteq X(x,z)$ for all $x, y, z \in X$. Quantale spaces are ordered by $x \leq_X x' \Leftrightarrow e \sqsubseteq X(x,x')$.

[1] We provide the necessary category theoretic definitions but assume that the reader is familiar with the corresponding notions from order theory.

[2] In category theory, an Ω-space is known as a category enriched over Ω.

A **functor** $F : X \to Y$ between Ω-spaces is a function on the underlying sets satsifying $X(x, x') \sqsubseteq Y(Fx, Fx')$

If $\Omega = 2$, a quantale space is a preorder and a functor is an order-preserving map. If $\Omega = [0, \infty]$, a quantale space is a generalized metric space [10] and a functor is a distance non-increasing function. If $\Omega = \mathcal{P}(\Sigma^*)$, a quantale space is a generalized non-deterministic automata [1] and a functor is a "simulation".

Weighted Relations. A *weighted relation* $R : X \looparrowright Y$ between quantale spaces X and Y is a function $X \times Y \to \Omega$ satisfying

$$X(x', x) \cdot R(x, y) \sqsubseteq R(x', y) \qquad R(x, y) \cdot Y(y, y') \sqsubseteq R(x, y').$$

Composition of weighted relations $R : X \looparrowright Y$ and $S : Y \looparrowright Z$ is given by

$$(R \bullet S)(x, z) = \bigsqcup_{y \in Y} R(x, y) \cdot S(y, z)$$

and has residuals

$$(R \blacktriangleright T)(y, z) = \bigsqcap_{x \in X} R(x, y) \rhd T(x, z) \qquad (T \blacktriangleleft S)(x, y) = \bigsqcap_{z \in Z} T(x, z) \lhd S(y, z).$$

Weighted relations specialize to *weakening relations* [6] if $\Omega = 2$.

Weighted Downsets and Upsets. A weighted downset $\varphi \in \mathcal{D}X$ is a relation $X \looparrowright 1$ (where 1 is the one-element Ω-space) and a weighted upset $\psi \in \mathcal{U}Y$ is a relation $1 \looparrowright Y$.[3] $\mathcal{D}X$ and $\mathcal{U}A$ are Ω-space with

$$\mathcal{D}X(\varphi, \varphi') = \bigsqcap_{x \in X} (\varphi x \rhd \varphi' x) \qquad \mathcal{U}A(\psi, \psi') = \bigsqcap_{a \in A} (\psi a \lhd \psi' a).$$

Note that $\mathcal{U}A$ is ordered by "reverse inclusion".

Weighted Limits and Colimits. In preorders, the category theoretic notions of limit and colimit specialize to meets and joins. In Ω-spaces, in addition to meets and joins, we have a type of colimit known as tensor and a type of limit known as power.

Given an Ω-space B, $b \star r$ is called the **tensor** of $b \in B$ with $r \in \Omega$, and, dually, $b \uparrow r$ is the **power** of b with r, if

$$B(b \star r, c) = r \rhd B(b, c) \qquad B(c, b \uparrow r) = B(c, b) \lhd r.$$

Given $G : D \to B$, $\varphi \in \mathcal{D}D$ and $\psi \in \mathcal{U}D$, we can now define the **colimit of** G **weighted by** φ and the **limit of** G **weighted by** ψ via

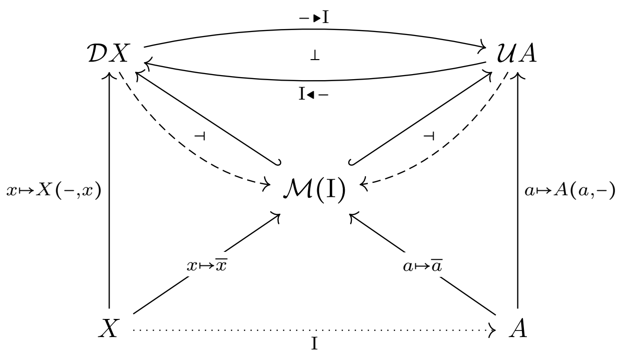
$$\mathrm{colim}_\varphi G = \bigsqcup_{d \in D} (Gd \star \varphi d) \qquad \lim_\psi G = \bigsqcap_{d \in D} (Gd \uparrow \psi d).$$

For example, if $B = \Omega = [0, \infty]$, then $b \star r = b + r$ and $b \uparrow r = b \dot{-} r$ [12].

[3] In CT, a weighted downset is known as a presheaf, a weighted upset as a co-presheaf.

3 MacNeille Completion

The MacNeille completion generalizes from the order to the quantale-enriched setting [11,13,8,5]. We adapt terminology from formal concept analysis [7].



Given a (weighted) relation $\mathrm{I} : X \nrightarrow A$, the MacNeille completion $\mathcal{M}(\mathrm{I})$ is defined as the set of pairs $\kappa = ([\![\kappa]\!], (\![\kappa]\!)) \in \mathcal{D}X \times \mathcal{U}A$ (often referred to as *concepts*) such that $[\![\kappa]\!] \blacktriangleright \mathrm{I} = (\![\kappa]\!)$ and $[\![\kappa]\!] = \mathrm{I} \blacktriangleleft (\![\kappa]\!)$. The MacNeille completion of an Ω-space C is the MacNeille completion of $C(-,-) : C \nrightarrow C$.

Remark.

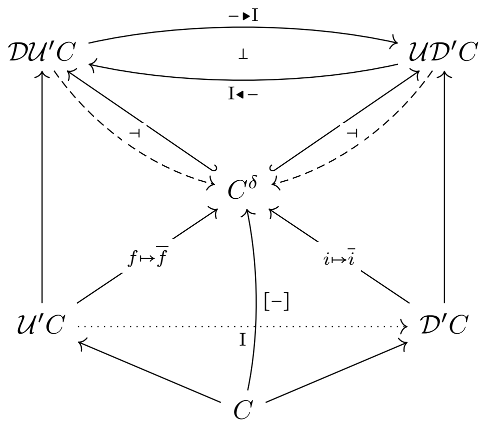
(i) $\mathcal{D}X$ is the free cocompletion of X and $\mathcal{U}A$ is the free completion of A. $\mathcal{M}(\mathrm{I})$ is a full reflective subcategory of $\mathcal{D}X$ and a full coreflective subcategory of $\mathcal{U}A$. Hence $\mathcal{M}(\mathrm{I})$ inherits colimits from $\mathcal{D}X$ and limits from $\mathcal{U}A$ and is, therefore, complete and complete.

(ii) If $I : C \nrightarrow C$ is the hom of a quantale space C, that is, if $X = A = C$ and $I(x,a) = C(x,a)$, then $X \to \mathcal{M}(\mathrm{I})$ and $A \to \mathcal{M}(\mathrm{I})$ are embeddings (fully faithful). Since the Yoneda embedding $x \mapsto X(-,x)$ preserves all limits and $a \mapsto A(-,a)$ preserves all colimits, the embedding $C \to \mathcal{M}(\mathrm{I})$ preserves all existing limits and colimits.

4 Canonical Extension

The canonical extension C^δ of a quantale space C is parameterised by a choice of full subcategories $\mathcal{U}'C \to \mathcal{U}C$ and $\mathcal{D}'C \to \mathcal{D}C$ that contain at least the principal upsets (respectively downsets). For the purpose of this abstract, the reader can choose $\mathcal{U}'C = \mathcal{U}C$ and $\mathcal{D}'C = \mathcal{D}C$. For applications to modal logics, one may want to require the (co)presheaves to preserve certain (co)limits. In the paradigmatic example, $\mathcal{U}'C$ consists of the "weighted" filters f and $\mathcal{D}'C$ of the "weighted" ideals i.

The canonical extension C^δ of a quantale space C is the MacNeille completion of the relation $\mathrm{I} : \mathcal{U}'C \nrightarrow \mathcal{D}'C$ given by $\mathrm{I}(f,i) = \bigsqcup_c f(c) \cdot i(c)$, that is, the set of fixed points of the adjunction given by $\varphi \blacktriangleright \mathrm{I} = \bigsqcap_f \varphi(f) \triangleright \mathrm{I}(f,-)$ and $\mathrm{I} \blacktriangleleft \psi = \bigsqcap_i \mathrm{I}(-,i) \triangleleft \psi(i)$.



For a survey of canonical extensions of lattices, we refer to [9]. In particular, Proposition 4 gives two equivalent formulations of compactness. The second, stating that for every filter f and ideal i with $\bigwedge f \le \bigvee i$ we have $f \cap i \neq \varnothing$ is the one we generalize to the quantale-enriched setting in our main theorem.

Theorem 4.1 *Let $f \in \mathcal{U}'C$ and $i \in \mathcal{D}'C$. Then C^δ is compact in the sense that*[4] $C^\delta(\lim_f[-], \operatorname{colim}_i[-]) = \mathrm{I}(f,i)$. *Moreover, every $(\varphi, \psi) \in C^\delta$ is the colimit of a limit of C and the limit of a colimit of C.*

Remark. In the applications we have in mind, C may be the Lindenbaum algebra of a fuzzy logic induced by Ω. In this setting, (monotone) modal operators correspond to functors on C, which will occupy us for the rest of the section.

We are interested in extending functors $G : C \to D$ between Ω-categories via the intermediate level to functors $G^\sigma, G^\pi : C^\delta \to D^\delta$ on canonical extensions. Given a functor $G : C \to D$ between Ω-spaces, we define by precomposition the functors

$$\mathcal{D}C \xleftarrow[G_r]{\qquad} \mathcal{D}D \qquad\qquad \mathcal{U}C \xleftarrow[G_l]{\qquad} \mathcal{U}D$$

In detail, given $f \in \mathcal{U}D$ and $i \in \mathcal{D}D$, we define $G_l(f)(c) = f(G(c))$ and $G_r(i)(c) = i(G(c))$.

We next show that if G preserves finite limits and colimits, then G_l and G_r restrict to filters and ideals.

Proposition 4.2 *Let $G : C \to D$. Let $f \in \mathcal{U}D$ preserve finite limits and let $i \in \mathcal{D}D$ preserve finite colimits. If G preserves finite limits, then so does $G_l(f)$ and if G preserves finite colimits, then so does $G_r(i)$.*

---

[4] $\lim_{\mathrm{f}}$ refers to the limit weighted by f and $\operatorname{colim}_{\mathrm{i}}$ to the colimit weighted by i.

We now extend G_l and G_r from the intermediate level to canonical extensions.

$$C^\delta \xleftarrow[G^r]{} D^\delta \qquad\qquad C^\delta \xleftarrow[G^l]{} D^\delta$$

Recall that for $\kappa \in D^\delta$, we have $[[\kappa]] \in \mathcal{D}\mathcal{U}'D$ and $([\kappa]) \in \mathcal{U}\mathcal{D}'D$. We define functors $G^l, G^r : D^\delta \to C^\delta$ as

$$G^l(\kappa) = \mathrm{colim}_{[[\kappa]]}\, \overline{G}_l \qquad\qquad G^r(\kappa) = \lim_{([\kappa])} \overline{G}_r.$$

The σ- and π-extensions on the intermediate level

$$\mathcal{D}C \underset{G_r}{\overset{G_\pi}{\rightleftarrows}} \mathcal{D}D \;(\bot) \qquad\qquad \mathcal{U}C \underset{G_l}{\overset{G_\sigma}{\rightleftarrows}} \mathcal{U}D \;(\top)$$

are defined as

$$G_\pi(i) = \mathrm{colim}_i\, D(-,G) \qquad\qquad G_\sigma(f) = \lim_f D(G,-).$$

Proposition 4.3 *We have $G_l(f) \bullet i = f \bullet G_\pi(i)$ and $f \bullet G_r(i) = G_\sigma(f) \bullet i$.*

Finally, we define functors $G^\pi, G^\sigma : C^\delta \to D^\delta$

$$G^\pi(\kappa) = \lim_{([\kappa])} \overline{G}_\pi \qquad\qquad G^\sigma(\kappa) = \mathrm{colim}_{[[\kappa]]} \overline{G}_\sigma$$

Theorem 4.4 (i) *If $G_l(f) \in \mathcal{U}'C$ and $G_\pi(i) \in \mathcal{D}'D$ for every $i \in \mathcal{D}'C$ and $f \in \mathcal{U}'D$, then $G^l \dashv G^\pi$.*

(ii) *If $G_\sigma(f) \in \mathcal{U}'D$ and $G_r(i) \in \mathcal{D}'C$ for every $i \in \mathcal{D}'D$ and $f \in \mathcal{U}'C$ then $G^\sigma \dashv G^r$.*

5 Conclusion and Future Work

We generalized the notion of canonical extension to quantale-enriched categories and developed a theory of σ- and π- extensions of functors between such categories.

This will be a stepping stone to developing logics of relational structures on quantale-enriched categories. Fuzzy modal logics have been developed and interpreted on sets with *many-valued* relations (fuzzy Kripke models) [4].

Such models can be generalized by substituting the set with a many-valued polarity, augmented with further fuzzy relations intended to interpret modal operations. These relational semantics on modal logics have been investigated in [2]. These logics include logical connectives and modalities. Even though the interpretation of formulas in these logic can be readily understood as fuzzy subsets, since these logics are not self-extensional, the interpretation of sequents needs to also be given. If sequents are interpreted simply via subsethood, then the expressiveness of the logic becomes the same as that over 2-valued polarities, i.e. lattices with operations, adding no expressivity despite a richer semantic environment. A way to augment the expressiveness of the language is to consider sequent relations $\vdash_r$, parametrized by truth-values $r \in \Omega$. In this case, quantale spaces will be the natural algebraic environment to interpret these logics: Logical connectives will be interpreted as weighted (co)limits (meets, joins, tensor, power), additional modal operators will be interpreted as endofunctors preserving certain (co)limits, and sequents will be interpreted via the defining relation of the quantale space. In particular, given a quantale space C and an interpretation of formulas $v(\cdot)$ one defines

$$C \vDash \varphi \vdash_r \psi \;\Leftrightarrow\; r \leq C(v(\varphi), v(\psi)).$$

Then, the canonical extensions developed here and extensions of functors will offer a modular and generic environment to prove completeness of *finitary* modal logics with respect to various relational semantics. The method hinges on extending functors $f : \mathbb{A} \to \mathbb{B}$ that preserve *finite* weighted (co)limits to maps on the canonical extensions $f^\sigma, f^\pi : \mathbb{A}^\delta \to \mathbb{B}^\delta$ that preserve arbitrary (co)limits and, hence, have adjoints and can be represented as fuzzy relations on the dual space of the quantal-enriched category.

References

[1] Betti, R. and S. Kasangian, *A quasi-universal realization of automata* (1982).
URL `http://hdl.handle.net/10077/6412`

[2] Conradie, W., A. De Domenico, K. Manoorkar, A. Palmigiano, M. Panettiere, D. Pinto Prieto and A. Tzimoulis, *Modal reduction principles across relational semantics*, Fuzzy Sets and Systems **481** (2024), p. 108892.

[3] Dunn, J. M., M. Gehrke and A. Palmigiano, *Canonical extensions and relational completeness of some substructural logics*, Journal of Symbolic Logic **70** (2005).

[4] Esteva, F., P. Garcia, L. Godo and R. Rodríguez, *A modal account of similarity-based reasoning*, International Journal of Approximate Reasoning **16** (1997), pp. 235–260.

[5] Fujii, S., *Completeness and injectivity*, Topology and its Applications **301** (2021).

[6] Galatos, N. and P. Jipsen, *The structure of generalized bi-algebras and weakening relation algebras*, Algebra universalis **81** (2020).

[7] Ganter, B. and R. Wille, "Formal Concept Analysis - Mathematical Foundations," Springer, 1999.

[8] Garner, R., *Topological functors as total categories*, Theory and Applications of Categories **29** (2014).

[9] Gehrke, M. and J. Vosmaer, *A view of canonical extension* (2010).
URL `https://arxiv.org/abs/1009.2803`

[10] Lawvere, F., *Metric spaces, generalized logic and closed categories*, Rendiconti del Seminario Matematico e Fisico di Milano **XLIII** (1973).

[11] Pavlovic, D., *Quantitative concept analysis*, in: *ICFCA 2012*.

[12] Rutten, J., *Weighted colimits and formal balls in generalized metric spaces*, Topology and its Applications **89** (1998), pp. 179–202.

[13] Willerton, S., *Tight spans, Isbell completions and semi-tropical modules*, Theory and Applications of Categories **28** (2013), pp. 696–732.
URL `http://www.tac.mta.ca/tac/volumes/28/22/28-22.pdf`

Nested Proof Theory for Quasi-Transitive Modal Logics

Sonia Marin

School of Computer Science, University of Birmingham
United Kingdom

Paaras Padhiar

School of Computer Science, University of Birmingham
United Kingdom

Abstract

Previous works by Goré, Postniece and Tiu have provided sound and cut-free complete proof systems for modal logics extended with path axioms using the formalism of nested sequent. Our aim is to provide (i) an internal cut-elimination procedure and (ii) alternative modular formulations for these systems. We present our methodology to achieve these two goals on a subclass of path axioms, namely quasi-transitivity axioms, and discuss how it could be extended further to quasi-symmetry axioms.

Keywords: Path axioms, Proof theory, Nested sequents, Cut-elimination, Modularity.

1 Introduction

The proof theory of modal logics has been explored thoroughly and many authors have contributed to the deep understanding gathered to this day. In particular, it has been remarked time and time again that in order to capture the validities of a modal logic, additional structure, often inspired by the semantics of the logic itself, is required within the proof-theoretical syntax. This led to the development of many formalisms extending Gentzen's sequent calculus, such as hypersequents [1], nested sequents [2,9], and labelled sequents [8].

It is not always clear however what sort of additional structure is precisely required to design the proof theory of a modal logic. For example, modal logic $\mathsf{S5}$ can be expressed using labelled or nested sequents, but can also be given a sound and complete system in the lighter hypersequent formalism, whereas such a result is conjectured not to be possible in ordinary sequent calculus [5].

Goré, Postniece and Tiu [4] have proposed a general algorithm to design nested sequent systems for modal logic K[1] extended with *path axioms*, of the

[1] Their work takes place in the context of tense logic where the language contains also adjoint

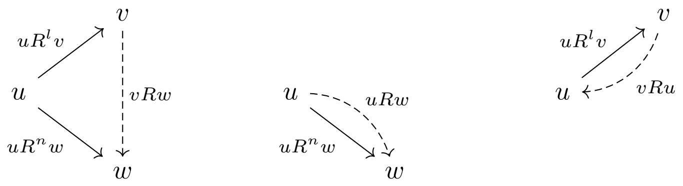


Fig. 1. Frame conditions for path, quasi-transitivity, and quasi-symmetry axioms

form $\Diamond^n a \supset \Box^l \Diamond a$. As a subclass of Scott-Lemmon axioms $\Diamond^n \Box^m a \supset \Box^l \Diamond^k a$ [6], they enjoy a well-behaved correspondence with the frame conditions displayed on the left of Fig. 1, i.e., if $uR^l v$ and $uR^n w$ then vRw. They are also known to be decidable logics, as shown by [4] using automata theoretic methods.

In this line of work, we set out to understand more precisely Goré et al.'s systems proof theoretically, in particular on a methodolgy to (i) equip them with an internal *cut-elimination* procedure and (ii) distill them into *modular* systems, i.e., such that each axiom corresponds to a (set of) rule(s) which can be freely mixed with others independently of the other axioms present.

We started with some restricted classes of path axioms. We call them *quasi-transitivity* when $l = 0$, giving $\mathsf{4^n} : \Diamond^n a \supset \Diamond a$ for $n \geq 1$, and *quasi-symmetry* when $n = 0$, giving $\mathsf{b^l} : a \supset \Box^l \Diamond a$ for $l \geq 1$. These correspond respectively to the frame conditions displayed in the middle and on the right of Fig. 1.

In this short paper we present our preliminary results regarding quasi-transitive modal logics and discuss the difficulty we faced so far to export the approach to quasi-symmetric ones.

2 Nested sequent for quasi-transitive logics

2.1 Nested sequents

A *nested sequent* is defined as $\Gamma ::= \varnothing \mid A, \Gamma \mid \Gamma, [\Gamma]$ which corresponds informally to a tree of sequents and can be expressed inductively in the modal language as: $\mathsf{form}(\varnothing) = \bot$, $\mathsf{form}(A, \Gamma) = A \vee \mathsf{form}(\Gamma)$, and $\mathsf{form}(\Gamma_1, [\Gamma_2]) = \mathsf{form}(\Gamma_1) \vee \Box\mathsf{form}(\Gamma_2)$.

A *context* is a nested sequent with one or several holes $\{\ \}$ which can take the place of a formula in the sequent (but does not occur inside a formula). This lets us write $\Gamma_1\{\Gamma_2\}$ when we replace the hole in $\Gamma_1\{\ \}$ by Γ_2.

Nested sequent system nK is composed of the rules id, $\wedge$, $\vee$, $\Box$ and $\Diamond_\mathsf{k}$ in Fig. 2 and is sound and complete wrt. K [2]. We can further extend this system with some of the modal propagation or modal structural rules in Fig. 2 to capture wider modal logics. The modal rules on the LHS ($\Diamond_\mathsf{k}$, $\Diamond_\mathsf{4}$, $\Diamond_\mathsf{b}$, as well as their structural versions) are taken from [2]. The ones on the RHS are the generalisations we are studying here, including $\Diamond_\mathsf{kn}$ which appears in [4].

A *proof* is then built from these rules as a tree in the same way as in

modalities, but we restrict our attention to the language with only $\Box$ and $\Diamond$.

$$\text{id}\ \frac{}{\Gamma\{a,\bar{a}\}} \qquad \wedge\ \frac{\Gamma\{A\} \quad \Gamma\{B\}}{\Gamma\{A \wedge B\}} \qquad \vee\ \frac{\Gamma\{A,B\}}{\Gamma\{A \vee B\}} \qquad \Box\ \frac{\Gamma\{[A]\}}{\Gamma\{\Box A\}} \qquad \text{cut}\ \frac{\Gamma\{A\} \quad \Gamma\{\bar{A}\}}{\Gamma\{\varnothing\}}$$

Modal propagation rules

$$\Diamond_{\mathsf{k}}\ \frac{\Gamma\{\Diamond A, [\Delta, A]\}}{\Gamma\{\Diamond A, [\Delta]\}} \qquad \Diamond_{\mathsf{kn}}\ \frac{\Gamma\{\Diamond A, [\Delta_1, [\ldots, [\Delta_n, A]\ldots]]\}}{\Gamma\{\Diamond A, [\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}\ n \geq 1$$

$$\Diamond_{\mathsf{4}}\ \frac{\Gamma\{\Diamond A, [\Delta, \Diamond A]\}}{\Gamma\{\Diamond A, [\Delta]\}} \qquad \Diamond_{\mathsf{4n}}\ \frac{\Gamma\{\Diamond A, [\Delta_1, [\ldots, [\Delta_{n-1}, \Diamond A]\ldots]]\}}{\Gamma\{\Diamond A, [\Delta_1, [\ldots, [\Delta_{n-1}]\ldots]]\}}\ n \geq 1$$

$$\Diamond_{\mathsf{b}}\ \frac{\Gamma\{A, [\Delta, \Diamond A]\}}{\Gamma\{[\Delta, \Diamond A]\}} \qquad \Diamond_{\mathsf{bn}}\ \frac{\Gamma\{A, [\Delta_1, [\ldots, [\Delta_n, \Diamond A]\ldots]]\}}{\Gamma\{[\Delta_1, [\ldots, [\Delta_n, \Diamond A]\ldots]]\}}\ n \geq 1$$

Modal structural rules

$$\boxtimes_{\mathsf{k}}\ \frac{\Gamma\{[\Delta], [\Sigma]\}}{\Gamma\{[\Delta, \Sigma]\}} \qquad \boxtimes_{\mathsf{kn}}\ \frac{\Gamma\{[\Sigma], [\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}{\Gamma\{[\Delta_1, [\ldots, [\Delta_n, \Sigma]\ldots]]\}}\ n \geq 1$$

$$\boxtimes_{\mathsf{b}}\ \frac{\Gamma\{[\Delta, [\Sigma]]\}}{\Gamma\{\Sigma, [\Delta]\}} \qquad \boxtimes_{\mathsf{bn}}\ \frac{\Gamma\{[\Delta_1, [\ldots, [\Delta_i, [\Sigma]]\ldots]]\}}{\Gamma\{\Sigma, [\Delta_1, [\ldots, [\Delta_i]\ldots]]\}}\ n \geq 1$$

Fig. 2. Nested sequent rules

ordinary sequent calculi. A rule $\frac{\Gamma_1}{\Gamma_2}$ is *admissible* in a nested sequent system N if whenever there is a proof of Γ_1 in N, there is a proof of Γ_2 in N.

2.2 Quasi-transitive modal logics

Let us look first at the transitive modal logic $\mathsf{K4}$, which is known to be sound complete with respect to $\mathsf{nK} + \Diamond_4$ (see Fig. 2). This can be proved, following Brünnler [2], via a cut-elimination argument, that is:

$$\begin{aligned} A \text{ is a theorem of } \mathsf{K}+\mathsf{4} &\iff A \text{ is provable in } \mathsf{nK} + \Diamond_4 + \mathsf{cut} \\ &\iff A \text{ is provable in } \mathsf{nK} + \Diamond_4 \end{aligned} \tag{1}$$

The cut-elimination proof itself is a bit involved as it requires the introduction of a complex generalisation of the cut rule, called a $\mathsf{4cut}$ in [2]

Meanwhile, Goré, Postniece and Tiu [4] prove a general soundness and completeness result for sets of path axioms, which we specialise to sets of quasi-transitivity axioms.

Let X be from now on be a subset of the positive natural numbers. We will write $\mathsf{K} + \mathsf{4}^{\mathsf{X}}$ to denote the modal logic K extended with (quasi-)transitivity axioms $\mathsf{4}^{\mathsf{n}} : \Diamond^n a \supset \Diamond a$ for each $n \in X$ and use the notations $\Diamond_{\mathsf{kX}} := \{\Diamond_{\mathsf{kn}} : n \in X\}$ and $\Diamond_{\mathsf{4X}} := \{\Diamond_{\mathsf{4n}} : n \in X\}$ (rules in Fig. 2).

The specialisation to quasi-transitivity allows us to simplify the notion of *completion*, namely, the set $\hat{X}$ can be defined inductively as:

$$X_0 := X \qquad X_{p+1} := X_p \cup \{i + j - 1 \mid i, j \in X_p\} \qquad \hat{X} := \bigcup_{p=0}^{\infty} X_p \tag{2}$$

This is a simplification of the completion given for a set of path axioms in [4], which is calculated from the algebra of paths for the propagation graph of a nested sequent. It allows Goré et al. to establish certain sets of $\Diamond_{\mathsf{kn}}$ rules to be sound and complete for certain path axioms.

Once simplified to quasi-transitive logics, their result can then be stated as:

$$A \text{ is a theorem of } \mathsf{K} + 4^{\mathsf{X}} \iff A \text{ is provable in } \mathsf{nK} + \Diamond_{\mathsf{k}\hat{\mathsf{X}}} \qquad (3)$$

However, the obtained systems are not *modular* as for two subsets of positive natural numbers X_1, X_2, the completion of $X_1 \cup X_2$ is not generally $\hat{X}_1 \cup \hat{X}_2$.

2.3 Main result and proof sketch

We generalise (1) and give an alternative proof of (3), via cut-elimination, to refine soundness and completeness for quasi-transitive modal logics:

Theorem 2.1 *The following are equivalent:*

(i) *A is a theorem of* $\mathsf{K} + 4^{\mathsf{X}}$

(ii) *A is provable in* $\mathsf{nK} + \Diamond_{\mathsf{kX}} + \mathsf{cut}$

(iii) *A is provable in* $\mathsf{nK} + \Diamond_{\mathsf{k}\hat{\mathsf{X}}}$

(iv) *A is provable in* $\mathsf{nK} + \Diamond_{\mathsf{4X}}$

We give a sketch of the proof here; the full version can be found in [7].

For the direction (i) $\Rightarrow$ (ii), knowing that the axioms and rules of K are derivable using $\mathsf{nK} + \mathsf{cut}$, we only need to show that for any $n \geq 1$, axiom 4^{n} is derivable using the corresponding rule $\Diamond_{\mathsf{kn}}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{[\ldots[\bar{a}, a]\ldots], \Diamond a}\,\mathsf{id}}{[\ldots[\bar{a}]\ldots], \Diamond a}\,\Diamond_{\mathsf{kn}}}{\Box^n \bar{a}, \Diamond a}\,\Box \;\; n \text{ times}}{\Box^n \bar{a} \vee \Diamond a}\,\vee}{\Diamond^n a \supset \Diamond a}$$

The direction (ii) $\Rightarrow$ (iii) is a cut-elimination proof. The result can be deduced via the proof translations from [4]; we provide here an internal proof which moreover lets us pinpoint precisely where the need for completion arises. By using the rules $\Diamond_{\mathsf{kn}}$ designed by [4], rather than the generalisation $\Diamond_{\mathsf{4n}}$ of the rule from [2], the cut-reduction case for the quasi-transitivity rules becomes much simpler, without the need to consider a $\mathsf{4cut}$-style rule. The cut-elimination argument follows as usual and most of the cases are identical to [2]. We are left to consider a cut of the form:

$$\dfrac{\dfrac{\Gamma\{[A], [\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}{\Gamma\{\Box A, [\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}\,\Box \quad \dfrac{\Gamma\{\Diamond\bar{A}, [\Delta_1, [\ldots, [\Delta_n, \bar{A}]\ldots]]\}}{\Gamma\{\Diamond\bar{A}, [\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}\,\Diamond_{\mathsf{kn}}}{\Gamma\{[\Delta_1, [\ldots, [\Delta_n]\ldots]]\}}\,\mathsf{cut}$$

which can be reduced to a cut on smaller formulas A and $\bar{A}$

$$\dfrac{\boxtimes_{\mathsf{kn}}\dfrac{\Gamma\{[A],[\Delta_1,[\ldots,[\Delta_n]\ldots]]\}}{\Gamma\{[\Delta_1,[\ldots,[\Delta_n,A]\ldots]]\}} \quad \Gamma\{[\Delta_1,[\ldots,[\Delta_n,\bar{A}]\ldots]]\}}{\Gamma\{[\Delta_1,[\ldots,[\Delta_n]\ldots]]\}}\ \mathsf{cut}$$

where the right premiss is obtained by applying the induction hypothesis on height to a cut on $\Box A$ and $\Diamond\bar{A}$ and the left premiss, is obtained by admissibility of the modal structural rules, namely:

Lemma 2.2 *For each $n \in X$, $\boxtimes_{\mathsf{kn}}$ is admissible in $\mathsf{nK} + \Diamond_{\mathsf{k}\hat{\mathsf{X}}}$.*

In the proof of this lemma, the requirement for completion becomes apparent.

The direction (iii) $\Rightarrow$ (iv) is where we achieve modularity. Indeed, using the $\Diamond_{\mathsf{4n}}$ rule which *propagates* formulas $\Diamond A$ (similar to the $\Diamond_{\mathsf{4}}$ rule from [2]), rather than the $\Diamond_{\mathsf{kn}}$ rule, allows us to drop the requirement of completion.

We need to show that for $n \in \hat{X}$, the rules $\Diamond_{\mathsf{kn}}$ and $\Diamond_{\mathsf{4n}}$ are derivable in $\mathsf{nK} + \Diamond_{\mathsf{4X}}$ by induction on the definition of $\hat{X}$. As a matter of example, if $n \in X_{p+1}$ and $n = l + m - 1$ for some $l, m \in X_p$, by induction hypothesis $\Diamond_{\mathsf{4l}}$ and $\Diamond_{\mathsf{km}}$ are derivable, hence $\Diamond_{\mathsf{kn}}$ can be shown derivable:

$$\Diamond_{\mathsf{k(l+m)}}\dfrac{\Gamma\{\Diamond A,[\Delta_1,[\ldots,[\Delta_{l+m},A]\ldots]]\}}{\Gamma\{\Diamond A,[\Delta_1,[\ldots,[\Delta_{l+m}]\ldots]]\}}$$

$$\equiv \quad \Diamond_{\mathsf{4l}}\dfrac{\Diamond_{\mathsf{km}}\dfrac{\Gamma\{\Diamond A,[\Delta_1,[\ldots,[\Delta_{l+m},A]\ldots]]\}}{\Gamma\{\Diamond A,[\Delta_1,[\ldots,[\Delta_{l-1},\Diamond A,[\ldots,[\Delta_{l+m}]]\ldots]\ldots]]\}}}{\Gamma\{\Diamond A,[\Delta_1,[\ldots,[\Delta_{l+m}]\ldots]]\}}$$

Finally, the direction (iv) $\Rightarrow$ (i) is stating the soundness of rules in $\mathsf{nK}+\Diamond_{\mathsf{4X}}$. The soundness of the rules in the system nK are proved in [2]. For a rule $\Diamond_{\mathsf{4n}}\frac{\Gamma_1}{\Gamma_2}$, we can similarly show that $\mathsf{form}(\Gamma_1) \supset \mathsf{form}(\Gamma_2)$ is a theorem of $\mathsf{K} + \mathsf{4}^{\mathsf{X}}$.

Note that the nested sequent systems we considered in this section, for quasi-transitive logics, are exclusively *propagation rule* based. The *structural rules* are used in the process of the cut-elimination proof but do not need to be explicitly added to the systems.

3 Towards nested sequents for quasi-symmetric logics

The completion we gave in (2) is a simplification of the completion given for a set of path axioms in [4], which is calculated by looking at the algebra of paths for a propagation graph of a nested sequent.

When specialising it again for the quasi-symmetric axioms $\mathsf{b}^{\mathsf{l}} : a \supset \Box^l \Diamond a$, we get the rules $\Diamond_{\mathsf{bn}}$ in Fig. 2. However, we have so far been unable to replicate the methodology developed in the previous section for this set of rules. In particular, the approach is unsuccessful when attempting to prove $\boxtimes_{\mathsf{bl}}$ admissible.

The strategy, similar to Lemma 2.2, would be to prove this through an induction on the height of the proof. That is, for example when $i = 2$, we would like to transform the derivation

$$\boxtimes_{b2}\dfrac{\Diamond_{b2}\dfrac{\Gamma\{[\Delta_1,[\Delta_2,A,[\Sigma_1,[\Sigma_2,\Diamond A]]]]\}}{\Gamma\{[\Delta_1,[\Delta_2,[\Sigma_1,[\Sigma_2,\Diamond A]]]]\}}}{\Gamma\{\Sigma_1,[\Sigma_2,\Diamond A],[\Delta_1,[\Delta_2]]\}}$$

into a derivation of this shape

$$\dfrac{\boxtimes_{b2}\dfrac{\Gamma\{[\Delta_1,[\Delta_2,A,[\Sigma_1,[\Sigma_2,\Diamond A]]]]\}}{\Gamma\{\Sigma_1,[\Sigma_2,\Diamond A],[\Delta_1,[\Delta_2,A]]\}}}{\Gamma\{\Sigma_1,[\Sigma_2,\Diamond A],[\Delta_1,[\Delta_2]]\}}$$

Unfortunately, the step indicated as a dashed line does not seem to correspond to any of the rules considered in this paper. It is not clear at this point whether the completion in [4] is enough to perform this step.

An alternative approach, inspired by Brünnler and Straßburger's [3], would be to renounce modal propagation rules and try to obtain a cut-elimination result for a system based exclusively on modal structural rules. These different avenues are the subject of ongoing work.

4 Concluding remarks

In this work, we proposed a proof-theoretic study of nested sequent systems for path axioms [4]. We showed its benefits on the special case of quasi-transitivity axioms and how it is challenged by quasi-symmetry axioms. The interest of the approach is not strictly to provide new nested sequent systems to a restricted family of logics for which we already know sound and complete systems in many different formalisms, including nested sequents. It is rather to dive deeper into some of these existing systems to understand the reason behind non-modularity and to ease the requirement of completion for path axioms in general.

References

[1] Avron, A., *The method of hypersequents in the proof theory of propositional non-classical logics*, in: *European logic colloquium*, Clarendon Press, 1996 pp. 1–32.

[2] Brünnler, K., *Deep sequent systems for modal logic*, Archive for Mathematical Logic **48** (2009), pp. 551–577.

[3] Brünnler, K. and L. Straßburger, *Modular sequent systems for modal logic*, in: M. Giese and A. Waaler, editors, *Automated Reasoning with Analytic Tableaux and Related Methods, TABLEAUX'09*, LNCS **5607** (2009), pp. 152–166.

[4] Goré, R., L. Postniece and A. Tiu, *On the correspondence between display postulates and deep inference in nested sequent calculi for tense logics*, LMCS **7** (2011), pp. 1–38.

[5] Lellmann, B., *Axioms vs hypersequent rules with context restrictions: Theory and applications*, in: *International Joint Conference on Automated Reasoning IJCAR*, Lecture Notes in Computer Science **8562** (2014), pp. 307–321.

[6] Lemmon, E. J. and D. S. Scott, "An Introduction to Modal Logic," Blackwell, 1977.

[7] Marin, S. and P. Padhiar, *Nested sequents for quasi-transitive modal logics* (2024). URL `https://arxiv.org/abs/2406.09300`

[8] Negri, S., *Proof analysis in modal logics*, Journal of Philosophical Logic **34** (2005), pp. 507–544.

[9] Poggiolesi, F., "Gentzen calculi for modal propositional logic," Trends in Logic **32**, Springer Science & Business Media, 2010.

The Exact Correspondence Between Intuitionistic and Modal Logic

Jan von Plato [1]

University of Helsinki

Abstract

A normalizing system of classical natural deduction for **S4** is given. It is shown that steps of indirect proof can be eliminated from derivations of formulas $T(A)$ translated from intuitionistic logic to **S4**. For the converse translation, the modal operators and their rules are simply deleted, to obtain a derivation of A in intuitionistic logic.

Keywords: modal translation, S4, normalization.

1 The question

Gödel published in 1933 a one-page paper in which he defined the modal system now known as **S4**, then gave a translation from intuitionistic propositional logic to **S4** such that translations $T(A)$ of theorems A of the former turn into theorems of the latter:

1.	$\Box(A) \supset A$
2.	$\Box(A \supset B) \,\&\, \Box(A) \supset \Box(B)$
3.	$\Box(A) \supset \Box\Box(A)$

Table 1. Axioms of S4.

1.	$A \vee B$	$\Box(A) \vee \Box(B)$
2.	$A \supset B$	$\Box(A) \supset \Box(B)$
3.	$A \,\&\, B$	$\Box(A) \,\&\, \Box(B)$
4.	$\neg A$	$\sim\Box(A)$

Table 2. Gödel's modal translation.

The rule of inference is that if A has been proved, $\Box(A)$ can be inferred. Gödel's article does not add the $\Box$-operator in the translation of conjunctions, but he does it elsewhere (cf. [4]). The translation shown gives a useful uniformity to the structure of translated formulas.

Gödel conjectured in 1933 that the correspondence he had established between the theorems of intuitionistic propositional logic and **S4** is exact: If the translation of a formula A is a theorem of **S4**, A is a theorem of intuitionistic logic. This was proved by semantic means in 1948 (cf. [5] and Troelstra's introduction to Gödel's short paper in the *Collected Works,* vol. 1, [2]).

[1] Full paper to appear in *The Bulletin of Symbolic Logic.* Research of the author financed by the European Research Council Advanced Grant GODELIANA (grant agreement No 787758).

It was found out recently that Gödel had found a complicated syntactic proof of his conjecture in 1941, in a book he wrote in 1940–42, *Results on Foundations* (cf. [3]). The proof uses a propositional version of what is today called Barr's theorem to first constructivize theorems of classical **S4**. The details and circumstances of this result are found in [6].

When natural deduction is used, the modal translation extends easily to derivations, with some added steps of introduction and elimination of $\Box$, with no instance of indirect proof visible anywhere. Thus, the proper perspective on the matter is to show that indirect proof in **S4** is conservative over the class of provable translated formulas. That is the content of our Main Lemma proved below. (Other proofs are only indicative, for brevity.)

2 Natural deduction for classical and modal logic

In [9], normalization is proved for classical natural deduction, with no *ad hoc* restrictions, global proof transformations, or similar tricks. Also natural deduction for the classical modal logic **S4** was covered, as an extension of a proof of normalization for intuitionistic **S4** in [8]. It will be useful to have the elimination rules in their general form, analogous to $\vee E$ and $\exists E$ in which the major premiss (the one with the connective) stands on the inference line and the derivations of the minor premiss or premisses to its right (cf. [7]). This arrangement is seen in the $\Box E$ rule in table 4 below.

In classical natural deduction, the subformula property of normal derivations includes also negations of subformulas. In rules that close assumptions, the assumptions can occur any number of times. If an assumption occurs 0 times, it is **vacuously** discharged. The system of intuitionistic natural deduction **NI** is obtained as that special case of the classical system **NK** in which the two assumptions $\neg A$ on top of the subderivations in rule $\neg E$ are absent, i.e., vacuously discharged:

$$\dfrac{\begin{matrix} \overset{1}{A} & \overset{1}{A} \\ \vdots & \vdots \\ C & \neg C \end{matrix}}{\neg A}{\scriptstyle \neg I,1} \qquad \dfrac{\begin{matrix} \overset{1}{\neg A} & \overset{1}{\neg A} \\ \vdots & \vdots \\ C & \neg C \end{matrix}}{A}{\scriptstyle \neg E,1} \qquad \dfrac{\begin{matrix} \vdots & \vdots \\ C & \neg C \end{matrix}}{A}{\scriptstyle \neg E}$$

Table 3. Rules for primitive negation, ex falso.

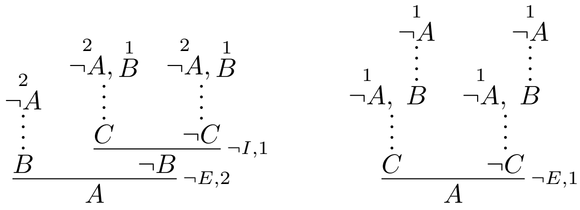
Definition. *A derivation in* **NK** *is* **normal** *if all major premisses of E-rules are assumptions.*

Here, contrary to [9], negation is a primitive notion. It can then happen that a major premiss of an elimination is derived by rule $\neg E$. That rule has no major premiss, but there are some obvious convertibilities with the two negation rules.

Theorem. Normalization for NK. *Derivations in* **NK** *normalize.*

The cases to consider, beyond the proof of normalization for **NI** (with general elimination rules), are convertibilities when the major premiss of one of $\&E, \vee E$, or $\supset E$ is derived by $\neg E$, and secondly cases with successive instances of rules $\neg I$ and $\neg E$. For brevity, only the latter is shown here:

Rule $\neg I$ followed by rule $\neg E$: Consider an uppermost such occurrence. The derivation and its transformation are:



As can be expected, the conversion will not make indirect proof disappear, even if there is the rule pair $\neg I, \neg E$. The composition formula B can lead to a new convertibility, but it is in a strictly smaller subderivation and therefore such cases get removed in a bounded number of steps.

There are, in addition to the above obvious cases, three further reductions (omitted here) that we take as belonging to a normal derivation:

Lemma. Reduction of indirect proof. *Successive applications of the negation rules reduce to at most one such application.*

Corollary. Indirect proof reduced to one last instance in a derivation. *Indirect proof can be permuted down with respect to all of the propositional rules.*

It is readily seen, by routine transformations, that rule $\neg E$ permutes down with respect to all the introduction and elimination rules for $\&, \vee$, and $\supset$. Moreover, as shown in the above lemma, consecutive instances of rule $\neg E$ collapse into just one such instance.

The rules of the system of modal logic **S4** are, with rule $\Box E$ formulated as a general elimination rule:

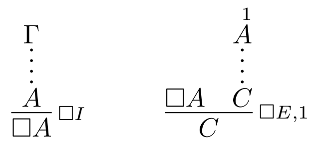


Table 4. Natural modal rules.

Rule $\Box I$ has the restriction that the open assumptions Γ on which the premiss A depends must be modal formulas, and the effect is that indirect proof cannot be permuted below $\Box I$. The standard $\Box E$ rule is obtained when $C = A$. The intuitionistic and classical modal systems will be denoted by **NIS4** and **NKS4**.

Theorem. Normalization for NKS4. *Derivations in* **NKS4** *normalize.*

The only new case to consider is when the major premiss of $\Box E$ has been derived by $\neg E$. The derivation and its transformation are:

$$\dfrac{\dfrac{\begin{matrix}\overset{1}{\neg\Box A}\\ \vdots\\ B\end{matrix}\qquad\begin{matrix}\overset{1}{\neg\Box A}\\ \vdots\\ \neg B\end{matrix}}{\Box A}{\scriptstyle\,\neg E,1}\qquad\begin{matrix}\overset{2}{A}\\ \vdots\\ C\end{matrix}}{C}{\scriptstyle\,\Box E,2}
\qquad
\dfrac{\begin{matrix}\dfrac{\dfrac{\overset{2}{\Box A}\quad\begin{matrix}\overset{1}{A}\\ \vdots\\ C\end{matrix}}{C}{\scriptstyle\,\Box E,1}\qquad\overset{5}{\neg C}}{\neg\Box A}{\scriptstyle\,\neg I,2}\\ \vdots\\ B\end{matrix}\qquad\begin{matrix}\dfrac{\dfrac{\overset{4}{\Box A}\quad\begin{matrix}\overset{3}{A}\\ \vdots\\ C\end{matrix}}{C}{\scriptstyle\,\Box E,3}\qquad\overset{5}{\neg C}}{\neg\Box A}{\scriptstyle\,\neg I,4}\\ \vdots\\ \neg B\end{matrix}}{C}{\scriptstyle\,\neg E,5}$$

Whereas rule $\neg E$ permutes down with respect to rule $\Box E$, the condition in rule $\Box I$, entirely analogous to the eigenvariable condition of rule $\forall I$, rules out such permutations.

Corollary. CM-lemma. *If in a normal derivation of a theorem in* **NKS4** *the last rule is* $\neg E$*, the derivation of its right premiss is degenerate.*

By the degeneracy, $C = A$, and with the right premiss left unwritten, the rule is:

$$\dfrac{\begin{matrix}\overset{1}{\neg A}\\ \vdots\\ A\end{matrix}}{A}{\scriptstyle\,CM,1}$$

Here *CM* stands for *Consequentia Mirabilis,* a rule that corresponds to the axiom $(\neg A \supset A) \supset A$. The CM-lemma holds in particular for **NK**.

The modal translation, as in Table 2, is best illustrated by an example:

$$T((A \supset \neg B) \supset (B \supset \neg A)) = \Box(\Box A \supset \Box\neg\Box B) \supset \Box(\Box B \supset \Box\neg\Box A)$$

Let formulas with one of $\&, \vee, \supset, \neg$ as the main connective be called **I**-*formulas* and ones with $\Box$ as the main connective **S**-*formulas.* The example shows how the immediate subformulas of translated formulas come in a succession **I**, **S**, **I**, **S**... until translations $T(P)$ of atomic formulas are reached. The curious subformula structure dictates that **NI**-rules must alternate with $\Box$-rules in normal derivations.

Main Lemma. Conservativity of NKS4 over NIS4 for T-formulas. *If the translation* $T(C)$ *of an intuitionistic formula* C *is derivable in* **NKS4***, it is already derivable in* **NIS4**.

Proof. Consider an uppermost step of $\neg E$ in a normal derivation in **NKS4**:

$$\dfrac{\begin{matrix}\overset{1}{\neg A}\\ \vdots\\ B\end{matrix}\qquad\begin{matrix}\overset{1}{\neg A}\\ \vdots\\ \neg B\end{matrix}}{A}{\scriptstyle\,\neg E,1}$$

In the derivation of the right premiss $\neg B$ of rule $\neg E$, assume the assumption $\neg A$ is not vacuous, the vacuous case postponed to the end. It is not a premiss in any

I-rule: Rule $\Box I$ is excluded by its restriction and the other I-rules would lead to formulas that don't have the alternating **S-I** subformula structure. Therefore $\neg A$ is a premiss in rule $\neg I$. The other premiss is not $\neg\neg A$ because then the conclusion of $\neg E$ would be a negative formula for which no indirect proof is needed. Therefore the other premiss is A, derived from some assumption C, and the conclusion of $\neg I$ is $\neg C$, as in

$$\dfrac{\begin{matrix}\overset{1}{\neg A}\\ \vdots \\ B\end{matrix} \qquad \begin{matrix}\dfrac{\begin{matrix}C\\ \vdots\\ A\end{matrix} \quad \overset{1}{\neg A}}{\neg C}{\scriptstyle \neg I}\\ \vdots\\ \neg B\end{matrix}}{A}{\scriptstyle \neg E,1}$$

$\neg C$ cannot be a minor premiss in rule $\supset E$, because the immediate subformulas of implications must be **S**-formulas. $\neg C$ is not a premiss in rule $\neg I$ as that would give two successive instances of $\neg I$. Therefore we must have $\neg B = \neg A$, then also $B = A$ and the derivation is:

$$\dfrac{\begin{matrix}\overset{1}{\neg A}\\ \vdots\\ A\end{matrix} \quad \overset{1}{\neg A}}{A}{\scriptstyle \neg E,1}$$

If A is the conclusion of the whole derivation, it is an **I**-formula, and the same if it is a premiss in rule $\Box I$. Then, if A is a negation, the step of indirect inference is not needed. Therefore A is equal to one of $\Box D \,\&\, \Box E, \Box D \vee \Box E$, or $\Box D \supset \Box E$, and A is derived by $\& I, \vee I$, or $\supset I$. In each of these rules, the premisses contain one of $\Box D, \Box E$, but in the presence of the negative assumption $\neg A$, these cannot have been derived by rule $\Box I$. Therefore the assumption $\neg A$ is not used in the derivation of the premiss A and the step of $\neg E$ can be left out to obtain a derivation of A in **NIS4**.

Finally, the case in which the assumption $\neg A$ is vacuous in the derivation of the right premiss of $\neg E$, in the first proof figure above: Here B must have the form $\Box D$. Its derivation in the left premiss of $\neg E$ must be by rule $\Box I$ and that can be, just as before, only if the assumption $\neg A$ was absent in the derivation of the left premiss. A then follows by *ex falso* in **NIS4**. QED.

Here is a normal derivation of our example in **NIS4** and its translation back to **NI**, by just deleting the $\Box$'s and their rules:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\overset{5}{\Box B} \qquad \dfrac{\overset{6}{\Box(\Box A \supset \Box\neg\Box B)} \qquad \dfrac{\overset{3}{\Box A \supset \Box\neg\Box B} \quad \overset{4}{\Box A} \quad \dfrac{\overset{2}{\Box\neg\Box B} \quad \overset{1}{\neg\Box B}}{\neg\Box B}{\scriptstyle \Box E,1}}{\neg\Box B}{\scriptstyle \supset E,2}}{\neg\Box B}{\scriptstyle \Box E,3}}{\neg\Box A}{\scriptstyle \neg I,4}}{\Box\neg\Box A}{\scriptstyle \Box I}}{\Box B \supset \Box\neg\Box A}{\scriptstyle \supset I,5}}{\Box(\Box B \supset \Box\neg\Box A)}{\scriptstyle \Box I}}{\Box(\Box A \supset \Box\neg\Box B) \supset \Box(\Box B \supset \Box\neg\Box A)}{\scriptstyle \supset I,6} \qquad \dfrac{\dfrac{\dfrac{\overset{5}{B} \qquad \dfrac{\overset{6}{A \supset \neg B} \quad \overset{4}{A} \quad \overset{2}{\neg B}}{\neg B}{\scriptstyle \supset E,2}}{\neg A}{\scriptstyle \neg I,4}}{B \supset \neg A}{\scriptstyle \supset I,5}}{(A \supset \neg B) \supset (B \supset \neg A)}{\scriptstyle \supset I,6}$$

References

[1] Gödel, K., Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums*, vol. 4 (1933), pp. 39–40. An English translation is found in Gödel's *Collected Works,* vol. 1.

[2] Gödel, K. *Collected Works*, vol. 1. Oxford U. P. 1986.

[3] Gödel, K., *Results on Foundations.* Edited by M. Hämeen-Anttila and J. von Plato. Springer 2023.

[4] Gödel, K. and H. Hahn, *Mathematical Logic in Vienna.* Edited by J. von Plato. Vienna Circle Institute Library, Springer 2025, in press.

[5] McKinsey, J, and A. Tarski, Some theorems about the sentential calculi of Lewis and Heyting. *The Journal of Symbolic Logic*, vol. 13 (1948), pp. 1–15.

[6] Negri, S. and J. von Plato (2025) Intuitionistic and modal logic in Gödel's *Resultate Grundlagen.* Submitted to a special issue of *The Bulletin of Symbolic Logic.*

[7] von Plato, J., Natural deduction with general elimination rules. *Archive for Mathematical Logic*, vol. 40 (2001), pp. 541–567.

[8] von Plato, J., Normal derivability in modal logic. *Mathematical Logic Quarterly*, vol. 51 (2005), pp. 632–638.

[9] von Plato, J. and A. Siders, Normal derivability in classical logic. *The Review of Symbolic Logic*, vol. 5 (2012), pp. 205–211.

Fusions of Canonical Predicate Modal Logics Are Canonical

Valentin Shehtman

Higher School of Modern Mathematics MIPT
National Research University Higher School of Economics
Moscow, Russia

Dmitry Shkatov

School of Computer Science and Applied Mathematics,
University of the Witwatersrand, Johannesburg,
South Africa

Abstract

We prove that canonicity of predicate modal logics transfers to their fusions.

Keywords: Predicate modal logic, fusion of logics, canonicity, strong Kripke completeness.

1 Introduction

In propositional modal logic, completeness through canonicity is a powerful technique for establishing Kripke completeness. A number of general results—among them, Sahlqvist canonicity theorem—show that modal propositional logics axiomatized by formulas of particular form are canonical and, therefore, Kripke complete. Moreover, it is known that canonicity and Kripke completeness of propositional logics transfer to their fusions [1,3]. By contrast, not much is known about canonicity of predicate modal (even monomodal) logics. The authors are only aware of the following general canonicity results for predicate modal logics: the Tanaka-Ono theorem for constant domains [5], canonicity of the minimal extensions of propositional one-way PTC logics with expanding domains [2, Theorem 6.1.29], and transfer of canonicity under boxing [4, Theorem 4.1]. Neither canonicity nor completeness transfers from logics to their fusions have been studied in the predicate setting. In this brief note, we show that, in predicate logic, just as in propositional logic, canonicity transfers to fusions.

2 Preliminaries

We consider logics in two languages: the monomodal predicate language $\mathcal{L}_1$ contains countably many free variables (denoted by $a, a_1, b, \ldots$), countably many bound variables (denoted by $x, y, x_1, \ldots$),[1] countably many predicate letters of every arity, the Boolean connectives $\neg$ and $\wedge$, the quantifier symbol $\forall$, and the unary modality $\Box_1$; the bimodal language $\mathcal{L}_2$ extends $\mathcal{L}_1$ with the unary modality $\Box_2$. Formulas are defined by recursion: atomic formulas are expressions of the form $P(a_1, \ldots, a_n)$; if A and B are formulas, then so is $(A \wedge B)$; if A is a formula, then so are $\neg A$, $\Box_i A$, and $\forall x\,[x/a]A$, where $[x/a]$ is a substitution of a bound variable x not occurring in A for a free variable a. As usual, $\bot$ abbreviates $(B \wedge \neg B)$, for some fixed B. An *occurrence of a free variable* a in a formula A is a triple (A, i, a) such that a is the ith symbol of A. The universal closure of a formula A, which is unique up to the renaming of variables, is denoted by $\bar{\forall} A$.

We denote by $\Box_2 Fma$ the set of all $\mathcal{L}_2$-formulas of the form $\Box_2 A$.

An *N-modal predicate logic* (in this paper, $N \in \{1, 2\}$) is a set of $\mathcal{L}_N$-formulas containing the minimal N-modal propositional logic $\mathbf{K}_N$ and closed under Substitution (Sub), Modus Ponens (MP), Generalization (Gen), and Necessitation for $\Box_1, \ldots, \Box_N$. The *fusion* $L_1 * L_2$ of 1-modal predicate logics L_1 and L_2 is the logic $\mathbf{K}_2 + L_1 \cup L_2^{+1}$, where L_2^{+1} is obtained from L_2 by replacing every occurrence of $\Box_1$ with $\Box_2$.

We work with Kripke semantics with expanding domains (see, e.g., [2, Chapter 3]). A *predicate Kripke N-frame with expanding domains* is a tuple $\boldsymbol{F} = (F, D)$ where $F = (W, R_1, \ldots, R_N)$ is a (propositional) Kripke N-frame and $D := \{D_w \mid w \in W\}$ is a system of non-empty domains over F such that, if $i \leqslant N$ and wR_iw', then $D_w \subseteq D_{w'}$. The following fact is well known:

Fact 2.1 *Let $\boldsymbol{F}$ be a predicate Kripke 2-frame, and let L_1 and L_2 be predicate monomodal logics. If $\boldsymbol{F} \models L_1$ and $\boldsymbol{F} \models L_2$, then $\boldsymbol{F} \models L_1 * L_2$.*

We write $(W, R) \subseteq (W', R')$ if a Kripke 1-frame (W, R) is a subframe of a Kripke 1-frame (W', R'), and $(W, R) \sqsubseteq (W', R')$ if (W, R) is a generated subframe of (W', R').

For the construction of canonical models, we use languages enriched with a countable set of *constants* (denoted by $c, c_1, \ldots$). Constants behave just like free variables, except that quantified formulas cannot be obtained by replacing constants with bound variables and prefixing a quantifier. A set of sentences possibly containing constants is called a *theory.* If Γ is a theory, the set of all constants occurring in Γ is denoted by $\mathcal{C}_\Gamma$ and the set of all sentences possibly containing constants from $\mathcal{C}_\Gamma$ is denoted by $\mathcal{L}(\Gamma)$. A theory Γ is called *negation-saturated* if, for every $A \in \mathcal{L}(\Gamma)$, either $A \in \Gamma$ or $\neg A \in \Gamma$. A theory Γ is called *Henkin* if, whenever $\exists x\, A(x) \in \mathcal{L}(\Gamma)$, there exists $c \in \mathcal{C}_\Gamma$ such that $\exists x A(x) \to A(c) \in \Gamma$.

We say that a formula, possibly with constants, A is *L-provable*, and write $\vdash_L A$, if $A = [\boldsymbol{c}/\boldsymbol{a}]B$, for some $B \in L$ and some renaming $[\boldsymbol{c}/\boldsymbol{a}]$ of some free variables of B with constants.

We say that a sentence A is *L-derivable from a theory* Γ, and write $\Gamma \vdash_L A$,

[1] Note that our syntax differs from that used in [2], where there are no separate stocks of free and bound variables.

if there exists a sequence $A_1, \dots, A_n$ of formulas, called an *L-derivation of A from Γ*, such that, for every A_i, either $A_i \in \Gamma$ or $\vdash_L A_i$ or else A_i is obtained from A_j, with $j < i$, by either (MP) or (Gen). A theory Γ is *L-consistent* if $\Gamma \nvdash_L \bot$. The following fact is well known and proven as for the classical logic:

Fact 2.2 *If Γ is a theory, a is a free variable, and c is a constant such that $c \notin \mathcal{C}_\Gamma$, then $\Gamma \vdash_L A$ implies $\Gamma \vdash_L [c/a]A$.*

Let L be an N-modal predicate logic. An *L-place* is a negation-saturated L-consistent Henkin theory Γ such that the set of constants not in $\mathcal{C}_\Gamma$ is infinite. The *canonical predicate frame* for L is the tuple $\boldsymbol{F}_L = (W_L, R_1, \dots, R_N, D_L)$ where W_L is the set of all L-places, $\Gamma R_i \Delta$ holds iff $\Box_i A \in \Gamma$ implies $A \in \Delta$, and the domain function is defined by $D_L(\Gamma) := \mathcal{C}_\Gamma$. It is well-known that every non-theorem of L is refuted on $\boldsymbol{F}_L$.

By analogy with propositional logic, we call a predicate logic *canonical* if it is validated by its canonical predicate frame. [2] Every canonical predicate logic is Kripke complete: if a predicate logic L is canonical, then $L = \{A \mid \boldsymbol{F}_L \models A\}$.

3 Main result

In view of Fact 2.1, our aim is to show that if L_1 and L_2 are canonical monomodal predicate logics, then the canonical predicate frame for $L_1 * L_2$ validates both L_1 and L_2. The arguments for L_1 and L_2 are symmetric, so we give only one in full detail.

Define a binary relation $\sim$ on *Fma* so that $A \sim B$ if B can be obtained by replacing occurrences of free variables in A with some free variables; e.g.,

$$\Box_2 \exists x (P(x,a,b) \wedge Q(x,b,c)) \sim \Box_2 \exists x (P(x,a,d) \wedge Q(x,a,b)), \qquad (*)$$

i.e, the only occurrence of a is replaced with a, the first occurrence of b with d, the second occurrence of b with a, and the only occurrence of c with b. It should be clear that $\sim$ is an equivalence; we write $[A]$ for $\{B \mid B \sim A\}$.

Enrich $\mathcal{L}_1$ with a countable set of predicate letters of each arity; denote by $\mathcal{S}$ the set of all newly introduced predicate letters and by $(\mathcal{L}_1 + \mathcal{S})$ the set of formulas of the resultant language; denote by $AF_\mathcal{S}$ the set of atomic formulas with predicate letters from $\mathcal{S}$.

Let $s\colon \Box_2 Fma/{\sim} \to \mathcal{S}$ be a bijection such that the arity of the letter $s([\Box_2 A])$ equals the number of occurrences of free variables in $\Box_2 A$; the map s is well defined since all formulas from $[\Box_2 A]$ have the same number of occurrences of free variables (e.g., four in formulas from $(*)$). We write $\Box_2 A(\boldsymbol{a})$ to mean that $\boldsymbol{a}$ is the list, with repetitions, of free variables with occurrences in $\Box_2 A$ (e.g., in the first formula from $(*)$, $\boldsymbol{a} = (a,b,b,c)$) and define a (unique) bijection $\bar{s}\colon \Box_2 Fma \to AF_\mathcal{S}$ so that $\bar{s}(\Box_2 A(\boldsymbol{a})) := s([\Box_2 A(\boldsymbol{a})])(\boldsymbol{a})$; the atomic formula $\bar{s}(\Box_2 A)$ is called the *surrogate* of $\Box_2 A$. Next, define a map $e\colon \mathcal{L}_2 \to (\mathcal{L}_1 + \mathcal{S})$

[2] Note that, in general, canonicity may depend on the cardinality of the set of constants used in the construction of a canonical predicate frame; for the purposes of this paper, however, this issue is immaterial.

by

$$\begin{array}{llll} e(A) & := A \quad \text{if } A \text{ is atomic;} & e(\forall x\,[x/a]A) & := \forall x[x/a]\,e(A); \\ e(\neg A) & := \neg e(A); & e(\Box_1 A) & := \Box_1 e(A); \\ e(A \wedge B) & := e(A) \wedge e(B); & e(\Box_2 A) & := \bar{s}(\Box_2 A). \end{array}$$

The formula $e(A) \in (\mathcal{L}_1 + \mathcal{S})$ is called the *ersatz* of $A \in \mathcal{L}_2$. If $\Gamma \subseteq \mathcal{L}_2$, then $\Gamma^e := \{e(A) \mid A \in \Gamma\}$. Note that the map e is a bijection; hence, we can define the map $r := e^{-1}$. The formula $r(A) \in \mathcal{L}_2$ is called the *reconstruction* of $A \in (\mathcal{L}_1 + \mathcal{S})$. Note that $r(A) = A$ if A does not contains predicate letters from $\mathcal{S}$. If $\Gamma \subseteq (\mathcal{L}_1 + \mathcal{S})$, then $\Gamma^r := \{r(A) \mid A \in \Gamma\}$.

Lemma 3.1 *For every $A \in (\mathcal{L}_1 + \mathcal{S})$, the formula $r(A)$ is a substitution instance of A.*

Proof. Induction on A. □

Let $L := L_1 * L_2$, and let

$$F_L := (W_L, R_L, D_L) \quad \text{and} \quad F_{L_1} := (W_{L_1}, R_{L_1}, D_{L_1})$$

be the canonical predicate frames of, respectively, L and L_1; let, also,

$$W_L^e := \{\Gamma^e \mid \Gamma \in W_L\}.$$

Lemma 3.2 *If $\Gamma \subseteq \mathcal{L}_2$ is negation-saturated and Henkin, then so is Γ^e.*

Proof. To see that Γ^e is negation-saturated, assume that $A \in \mathcal{L}(\Gamma) - \Gamma^e$. Then, $r(A) \notin \Gamma$, and so, since Γ is negation-saturated, $\neg r(A) (= r(\neg A)) \in \Gamma$. Thus, $e(r(\neg A)) (= \neg A) \in \Gamma^e$. The argument for Henkinness is similar. □

Lemma 3.3 *If $\Gamma \subseteq (\mathcal{L}_1 + \mathcal{S})$ is negation-saturated and Henkin, then so is Γ^r.*

Proof. Similar to the proof of Lemma 3.2. □

Lemma 3.4 $W_L^e \subseteq W_{L_1}$.

Proof. Let Γ be an L-place. By Lemma 3.2, Γ^e is negation-saturated and Henkin. Clearly, $\mathcal{C}_{\Gamma^e} = \mathcal{C}_\Gamma$. It remains to show that Γ^e is L_1-consistent.

Suppose not, i.e., $\Gamma^e \vdash_{L_1} \bot$. Then, there exists an L_1-derivation $A_1, A_2, \ldots, \bot$ of $\bot$ from Γ^e. Then, as we next show,

$$r(A_1), r(A_2), \ldots, r(\bot) (= \bot)$$

is an L-derivation of $\bot$ from Γ. Indeed, if $A_i \in \Gamma^e$, then $r(A_i) \in (\Gamma^e)^r (= \Gamma)$. If $\vdash_{L_1} A_i$, then there exists a renaming $[\boldsymbol{a}/\boldsymbol{c}]$ of the constants occurring in A_i into free variables such that $[\boldsymbol{a}/\boldsymbol{c}]A_i \in L_1$. Since $L_1 \subseteq L_1 * L_2 (= L)$, surely $[\boldsymbol{a}/\boldsymbol{c}]A_i \in L$; hence, $\vdash_L A_i$. Since L is closed under substitution, it follows, by Lemma 3.1, that $\vdash_L r(A_i)$. Lastly, the map r clearly commutes with both (MP) and (Gen). Thus, $\Gamma \vdash_L \bot$, contrary to L-consistency of Γ. □

If Γ is a set of formulas, we denote by $\overline{\Gamma}$ the set of sentences from Γ.

Lemma 3.5 *If* $\Gamma \in W_L^e$*, then* $\overline{L}^e \subseteq \Gamma$.

Proof. Let $\Gamma \in W_L^e$. Then, $\Gamma = \Delta^e$, for some $\Delta \in W_L$. Since, by [2, Lemma 6.1.4(1)], $\overline{L} \subseteq \Delta$, it follows that $\overline{L}^e \subseteq \Delta^e (= \Gamma)$. □

Lemma 3.6 *If* $\Gamma \in W_{L_1}$ *and* $\overline{L}^e \subseteq \Gamma$*, then* $\Gamma \in W_L^e$.

Proof. Suppose that $\overline{L}^e \subseteq \Gamma \in W_{L_1}$. We prove that $\Gamma^r \in W_L$; since $\Gamma = (\Gamma^r)^e$, it then follows that $\Gamma \in W_L^e$. By Lemma 3.3, Γ^r is negation-saturated and Henkin. Clearly, $\mathcal{C}_{\Gamma^r} = \mathcal{C}_\Gamma$. It remains to show that Γ^r is L-consistent.

Suppose not, i.e., $\Gamma^r \vdash_L \bot (= r(\bot))$. We will prove that, then, $\Gamma \vdash_{L_1} \bot$. To that end, we show that $\Delta \vdash_L A$ and $\overline{L}^e \subseteq \Delta^e$ imply $\Delta^e \vdash_{L_1} e(A)$, for every $\Delta \in W_L$ and every $A \in \mathcal{L}_2$. (The required conclusion then follows from the fact that $(\Gamma^r)^e = \Gamma$.) We proceed by induction on the L-derivation of A from Δ. If $B \in \Delta$, then $e(B) \in \Delta^e$. Suppose, next, that $\vdash_L B$. We may assume that B does not contain any constants from Δ, and hence any constants from Δ^e. Since $\vdash_L B$, there exists a renaming $[\boldsymbol{a}/\boldsymbol{c}]$ of constants into free variables such that $[\boldsymbol{a}/\boldsymbol{c}]B \in L$. Let $B' := [\boldsymbol{a}/\boldsymbol{c}]B$. By (Gen), $\bar{\forall} B' \in \overline{L}$. Since, by assumption, $\overline{L}^e \subseteq \Delta^e$, it follows that $e(\bar{\forall} B')(= \bar{\forall}\, e(B')) \in \Delta^e$. Thus, we can add $\bar{\forall}\, e(B')$ and the L_1-theorem $\bar{\forall}\, e(B') \to e(B')$ to any L_1-derivation from Δ^e; hence, $\Delta^e \vdash_{L_1} e(B')$. Now, let $[\boldsymbol{c}/\boldsymbol{a}] := [\boldsymbol{a}/\boldsymbol{c}]^{-1}$. Then, by Fact 2.2 and by our assumption about constants, $\Delta^e \vdash_{L_1} [\boldsymbol{c}/\boldsymbol{a}]e(B')(= e([\boldsymbol{c}/\boldsymbol{a}]B') = e(B))$. Lastly, the map e clearly commutes with (MP) and (Gen). Hence, $\Gamma \vdash_{L_1} \bot$, contrary to L_1-consistency of Γ. □

Let $R_L^e := R_{L_1} \upharpoonright W_L^e$.

Lemma 3.7 $(W_L^e, R_L^e) \sqsubseteq (W_{L_1}, R_{L_1})$.

Proof. It follows from Lemma 3.4 and the definition of R_L^e that $(W_L^e, R_L^e) \subseteq (W_{L_1}, R_{L_1})$. To see that $R_{L_1}(W_L^e) \subseteq W_L^e$, suppose that $\Gamma \in W_L^e$ and $\Gamma R_{L_1} \Delta$. Since $\Gamma \in W_L^e$, it follows, by Lemma 3.5, that $\overline{L}^e \subseteq \Gamma$. Since $\Box_1 \overline{L} \subseteq \overline{L}$, surely $(\Box_1 \overline{L})^e (= \Box_1 \overline{L}^e) \subseteq \Gamma$. Since $\Gamma R_{L_1} \Delta$, the definition of R_{L_1} implies that $\overline{L}^e \subseteq \Delta$. Hence, by Lemma 3.6, $\Delta \in W_L^e$. □

Proposition 3.8 *If* $\boldsymbol{F}_{L_1} \models L_1$*, then* $\boldsymbol{F}_{L_1 * L_2} \models L_1$.

Proof. Suppose that $\boldsymbol{F}_{L_1} \models L_1$. Since e is an embedding, Lemma 3.7 means that the frame $(W_{L_1 * L_2}, R_1)$ is isomorphic to a generated subframe of $\boldsymbol{F}_{L_1}$. Since validity of predicate formulas is preserved under generated subframes [2, Lemma 3.3.18], it follows that $(W_{L_1 * L_2}, R_1) \models L_1$ and hence $\boldsymbol{F}_{L_1 * L_2} \models L_1$. □

Lemma 3.9 *If* $\boldsymbol{F}_{L_2} \models L_2$*, then* $\boldsymbol{F}_{L_1 * L_2} \models L_2$.

Proof. The argument here is analogous to that of Proposition 3.8. We define surrogates of formulas of the form $\Box_1 A$ in the language $(\mathcal{L}_2 + \mathcal{S})$, and proceed as before, but swapping the roles of L_1 and L_2. □

Theorem 3.10 *Let* L_1 *and* L_2 *be predicate modal logics. Then,*

$$\boldsymbol{F}_{L_1} \models L_1 \ \&\ \boldsymbol{F}_{L_2} \models L_2 \implies \boldsymbol{F}_L \models L_1 * L_2.$$

In other words, the fusion of two canonical predicate modal logics is a canonical predicate modal logic.

Proof. Immediate from Proposition 3.8, Lemma 3.9 and Fact 2.1. □

Since an analogue of Theorem 3.10 can be proven for polymodal logics, we obtain the following:

Corollary 3.11 *Let $L_1, \ldots, L_n$ be predicate modal logics. Then,*

$$\boldsymbol{F}_{L_1} \models L_1 \;\&\; \ldots \;\&\; \boldsymbol{F}_{L_n} \models L_n \Longrightarrow \boldsymbol{F}_L \models L_1 * \ldots * L_n.$$

In other words, the fusion of any number of canonical predicate modal logics is a canonical predicate modal logic.

The previous treatment is likely to extend to logics of constant domains, with canonicity replaced by C-canonicity [2, Chapter 7]:

Conjecture 3.12 *For predicate modal logics, C-canonicity transfers to fusions.*

Also, we believe that additional techniques should enable us to prove transfer of Kripke completeness rather than simply canonicity:

Conjecture 3.13 *For predicate modal logics, strong Kripke completeness and Kripke completeness transfer to fusions.*

References

[1] Fine, K. and G. Schurz, *Transfer theorems for multimodal logics*, in: J. Copeland, editor, *Logic and Reality: Essays on the Legacy of Arthur Prior*, Cambridge University Press, 1996 pp. 169–213.

[2] Gabbay, D., V. Shehtman and D. Skvortsov, "Quantification in Nonclassical Logic, Volume 1," Studies in Logic and the Foundations of Mathematics **153**, Elsevier, 2009.

[3] Kracht, M. and F. Wolter, *Properties of independently axiomatizable bimodal logics*, Journal of Symbolic Logic **56** (1991), pp. 1469–1485.

[4] Shehtman, V., *On Kripke completeness of modal predicate logics around quantified K5*, Annals of Pure and Applied Logic 174 (2023), article 103202.

[5] Tanaka, Y. and H. Ono, *Rasiowa-Sikorski lemma and Kripke completeness of predicate and infinitary modal logics*, in: *Advances in Modal Logic 2*, CSLI Publications, 2001 pp. 401–420.

Substitution as Modality

Wen Tang, Yanjing Wang

Department of Philosophy, Peking University

Abstract

In the expositions of axiomatizations of first-order logic, the (admissible) substitution played an important role. However, it is usually defined only in the meta-language. In this paper, inspired by dynamic epistemic logic and modal logic with the assignment operators, we take *syntactic* substitutions as (dynamic) modalities in the logical language. Adding this operator to the first-order language complicates the logic, but we can avoid the meta-language notion of substitution in the axiomatizations. As the main result, we axiomatize first-order logic with equality enriched with such substitution modalities.

Keywords: substitution, first-order logic, dynamic epistemic logic, completeness

1 Introduction

In most modern expositions of the Hilbert-style proof systems of first-order logic (FOL), there are always some axiom schemata and rules that use the meta-language notion of *substitution*. However, the definitions of substitution and substitutability are provided in the meta-language, such as in [3]. A natural question is whether we can include the substitution in the object-language of logic, thus making $\alpha(t/x)$ a genuine formula schema just like $\forall x\alpha$. This motivates this paper.

A similar idea has been well-explored in the field of λ-calculus under the name of *explicit substitution*. By having the explicit substitution, one can bring λ-calculus closer to implementation in practice, without outsourcing a subtle meta-language notion of substitution. Moreover, this also allows us to delay substitution operations in reduction, which may be more efficient in practice. On the other hand, adding explicit substitutions also brings technical complications, as the substitutions can be nested in the term language, which may destroy strong normalization when accompanied by the most intuitive β-rules handling substitutions [10]. This can be overcome in various ways (e.g., [16,8]), and it influenced the closely related field of linear logic [4]. [1]

Coming back to FOL, there is a body of work in algebraic logic treating substitutions as algebraic operations based on Cylindric Algebras [7,11], Polyadic

[1] There also exists work on coding first-order logic using λ-calculus with substitution atoms [2].

Algebras [5], and Quantifier Algebras [12]. The substitution operations are usually parametrized by two (indexes of) variables, and substitutions with terms can be handled with more complicated extensions (e.g., [13]). In such works, the substitution operations are defined by a set of axioms that characterize some basic properties of substitutions.

In this work, we would like to make substitutions as first-class citizens in the object language of FOL. In particular, as modal logicians, we have the most natural way to bring a concept in the meta-language into the object-language: making it a *modality*.

In fact, making substitution-like operations as modalities is *not* a new idea, as already suggested in [15]. In [18,1], both quantifiers and (simultaneous) substitutions are treated like modalities, and a complete multi-type display calculus for a largely extended language of FOL is obtained based on algebraic semantics. There is also a large body of existing works in logic that make the *semantic counterparts* of substitutions as modalities, such as the assignment operator $[x := t]$ in dynamic logic [6] and epistemic logic [22], and various factual changing update operators in Dynamic Epistemic Logic [20,17]. However, the fundamental difference between our approach and these existing ones is that we treat substitution modalities as *syntax-changing operators*, while the above-mentioned works take those assignment-like modalities as model-changing operators.

Nevertheless, we are inspired by the axiomatic treatments of update modalities. For example, in public announcement logic [14], reduction axioms such as the following are used to axiomatize the logic by eliminating $[\varphi]$:

$$[\varphi]p \leftrightarrow (\varphi \to p), \quad [\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi), \quad [\varphi]\Box\psi \leftrightarrow (\varphi \to \Box[\varphi]\psi)$$

They can also be viewed as recursive definitions of the syntactic relativization [19]: by following the left-to-right direction of these axioms, a formula $[p]\psi$ will be rewritten into a formula ψ^p such that ψ^p holds in the original model iff ψ holds in the submodel consisting of the p-worlds. Since the dynamic operators can be nested, there are two ways to do reductions of $[\varphi]$, which can be described as inside-out and outside-in, as discussed in depth in [21], where the first relies on the replacement of equals and the second relies on the composition axiom combining two operations into one such as $[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$.

In this work, we follow these ideas to define the syntactic substitutions recursively in the system where nested substitutions are allowed in the language. Adding the substitution operator to propositional logic is relatively straightforward, where the rule of uniform substitution (from $\vdash \varphi$ infer $\vdash \varphi(\psi/p)$) also involves the meta-notion of substitution. We can extend the language with a substitution operator $[\varphi/p]$ and define the semantics by the corresponding syntactic substitution. The rule of uniform substitution becomes the rule of necessitation for the substitution operator. A sound and complete axiomatization of propositional logic with explicit substitution can be obtained without meta-language notions anymore. The resulting logic is identical to the logic of

[17] where $[\varphi/p]$ is interpreted as updating the value of p by the value of φ in the Kripke model. This is no surprise, as the syntactic substitution can be captured semantically in the propositional setting.

In this paper, we focus on the setting of FOL and define the semantics of substitution operator $[t/x]$ over both terms and formulas, naturally extending the so-called *standard substitution* in [9]. We obtain a sound and complete system without the notion of substitution in the meta-language and condition on the substitutability. Compared to the existing work in the algebraic settings such as [12,1], a notable feature of our work is to keep the language as close as possible to the standard language of FOL but allow terms with substitutions to occur in the substitution operators such as $[[t/x]g/y]$.

2 First-Order Logic with Substitution Operators

In this section, we introduce first-order logic with the explicit substitution modality. First, let us recall the so-called *standard substitution* for FOL in the meta-language [9], restricted to a single variable without relettering.

Definition 2.1 (Standard Substitution) *The standard substitution* (t/x) *in first-order logic is a function that maps terms and formulas of FOL with equality to terms and formulas respectively, defined as:*

$$y(t/x) := \begin{cases} x & \text{if } y \neq x \\ t & \text{if } y = x \end{cases} \qquad c(t/x) := c \qquad (ft_1 \dots t_n)(t/x) := ft_1(t/x) \dots t_n(t/x)$$

$$(Pt_1 \dots t_n)(t/x)) := Pt_1(t/x) \dots t_n(t/x) \qquad (t_1 \approx t_2)(t/x) := t_1(t/x) \approx t_2(t/x)$$
$$(\neg\varphi)(t/x) := \neg\varphi(t/x) \qquad (\varphi \wedge \psi)(t/x) := \varphi(t/x) \wedge \psi(t/x)$$

$$(\forall y\varphi)(t/x) := \begin{cases} \forall y\varphi(t/x) & \text{if } y \text{ does not occur in } t \\ \forall y\varphi & \text{otherwise} \end{cases}$$

Now we define the FOL language with explicit substitution modality. Note that the substitution operators can combine with both terms and formulas.

Definition 2.2 (First-Order Language with Substitutions) *Given a set of variables* $\mathbf{X}$*, a set of constants* $\mathbf{C}$*, a set of predicates* $\mathbf{S}$*, and a set of function symbols* $\mathbf{F}$*,* $\mathcal{L}_{\mathbf{FOS}}$*-terms and formulas are defined as follows:*

$$t ::= x \mid c \mid f\boldsymbol{t} \mid [t/x]t$$
$$\varphi ::= (t \approx t) \mid P\boldsymbol{t} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \forall x\varphi \mid [t/x]\varphi$$

where $x \in \mathbf{X}, c \in \mathbf{C}, f \in \mathbf{F}, P \in \mathbf{S}$*, and* $\approx$ *is the equality symbol.* $\boldsymbol{t}$ *denotes a list of terms. Let* $\mathbf{T}$ *be the set of terms, and let* $\mathbf{T_{FO}}$ *be the set of substitution-free terms, i.e., first-order terms. Let* $\mathcal{L}_{\mathbf{FO}}$ *be the substitution-free fragment of* $\mathcal{L}_{\mathbf{FOS}}$*, i.e. the standard first-order language.*

In this language, we can write formulas involving nested substitutions such as $[[y/x]fx/x]\forall z(Px \wedge \forall yRxyz)$, where $[y/x]fx$ is again a term. The semantics of $\mathcal{L}_{\mathbf{FOS}}$ is based on standard FO-structures $\mathfrak{A}$ with a domain and an interpretation for non-logical symbols. Given $\mathfrak{A}$, we denote the interpretations of $c \in \mathbf{C}, f \in \mathbf{F}, P \in \mathbf{S}$ as $x^{\mathfrak{A}}, f^{\mathfrak{A}}, P^{\mathfrak{A}}$ respectively. To determine the value of terms with (nested) substitutions, some care is needed.

Definition 2.3 Given a model $(\mathfrak{A}, \sigma)$ where σ is an assignment of variables, the interpretation $\mathcal{I}^\sigma$ of any $\mathcal{L}_{\mathbf{FOS}}$-term t is defined as follows:

- $\mathcal{I}^\sigma(x) = \sigma(x)$; $\mathcal{I}^\sigma(c) = c^{\mathfrak{A}}$; $\mathcal{I}^\sigma(ft_1 \dots t_n) = f^{\mathfrak{A}}(\mathcal{I}^\sigma(F(t_1)), \dots, \mathcal{I}^\sigma(F(t_n)))$;
- $\mathcal{I}^\sigma([t/x]t') = \mathcal{I}^\sigma(F([t/x]t'))$.

where F is defined as:

$$\begin{array}{rcl} F(x) & = & x \qquad F(c) = c \qquad F(ft_1 \dots t_n) = fF(t_1) \dots F(t_n) \\ F([g/x]x) & = & g \qquad F([g/x]y) = y \qquad F([g/x]c) = c \\ F([g/x]fg_1 \dots g_n) & = & fF([g/x]g_1) \dots F([g/x]g_n) \\ F([t/x]t') & = & F([F(t)/x]F(t')) \text{ if either } t \text{ or } t' \text{ is not in } \mathbf{T_{FO}} \end{array}$$

where $g \in \mathbf{T_{FO}}$, i.e., substitution-free.

We can show that for any $t \in \mathbf{T}$, $F(t) \in \mathbf{T_{FO}}$.

Remark 2.4 Although $[t/x]$ looks like a function "symbol" at the term level, it behaves differently. For example, for any (unary) function f, $\mathcal{I}^\sigma(x) = \mathcal{I}^\sigma(y)$ implies $\mathcal{I}^\sigma(fx) \approx \mathcal{I}^\sigma(fy)$, but $\mathcal{I}^\sigma(x) = \mathcal{I}^\sigma(y)$ does not imply $\mathcal{I}^\sigma([z/x]x) = \mathcal{I}^\sigma([z/x]y)$ which is equivalent to $\mathcal{I}^\sigma(z) = \mathcal{I}^\sigma(y)$.

The purpose of F is to execute the syntactic substitutions inside-out and eventually eliminate the substitutions. It is easy to see that if $g \in \mathbf{T_{FO}}$ then $F(g) = g$. Moreover, given $g, g' \in \mathbf{T_{FO}}$, $F([g/x]g')$ is exactly the result of doing substitution $g'(g/x)$ in the meta-language as in Definition 2.1. The last clause of F is to handle the nested substitutions. For example:

$$\begin{aligned} F([([f_1x/x]f_2x)/x]f_3x) &= F([F([f_1x/x]f_2x)/x]F(f_3x)) \\ &= F([f_2F([f_1x/x]x)/x]f_3F(x)) \\ &= F([f_2f_1x/x]f_3x) \\ &= f_3f_2f_1x \end{aligned}$$

Definition 2.5 (Semantics) *The semantics is defined as in standard FOL with the following clause for the substitution formulas:*

$$\mathfrak{A}, \sigma \vDash [t/x]\varphi \iff \mathfrak{A}, \sigma \vDash T_{x,F(t)}(\varphi)$$

where $T_{x,g} : \mathcal{L}_{\mathbf{FOS}} \cup \mathbf{T_{FO}} \to \mathcal{L}_{\mathbf{FOS}} \cup \mathbf{T_{FO}}$ *is defined as follows* ($g \in \mathbf{T_{FO}}$)*:*

$$\begin{array}{rcl} T_{x,g}(t_1 \approx t_2) & = & F([g/x]t_1) \approx F([g/x]t_2) \\ T_{x,g}(Pt_1 \dots t_n) & = & PF([g/x]t_1)) \dots F([g/x]t_n) \\ T_{x,g}(\neg\varphi) & = & \neg T_{x,g}(\varphi) \\ T_{x,g}(\varphi_1 \wedge \varphi_2) & = & T_{x,g}(\varphi_1) \wedge T_{x,g}(\varphi_2) \\ T_{x,g}(\forall y\varphi) & = & \begin{cases} \forall y T_{x,g}(\varphi) & \text{if } y \text{ does not occur in } g \\ \forall y\varphi & \text{otherwise} \end{cases} \\ T_{x,g}([t/y]\varphi) & = & T_{x,g}(T_{y,F(t)}(\varphi)) \end{array}$$

Remark 2.6 As $F(t) \in \mathbf{T_{FO}}$ thus we only need to define $T_{x,g}$ for $g \in \mathbf{T_{FO}}$. Like F, $T_{x,g}$ closely resembles the substitution on FO formulas as in Definition 2.1. It is not hard to verify $T_{x,g}(\alpha) = \alpha(g/x)$ when $g \in \mathbf{T_{FO}}$ and $\alpha \in \mathcal{L}_{\mathbf{FO}}$. The last clause of $T_{x,g}$ shows that we do substitution inside-out.

The tricky part about the semantics is again about nested substitutions, here is an example (x, y, z are distinct):

$$\begin{array}{l} \mathfrak{A}, \sigma \vDash [[y/x]fx/x]\forall z(Px \wedge \forall yRxyz)) \\ \iff \mathfrak{A}, \sigma \vDash T_{x,F([y/x]fx)}\forall z(Px \wedge \forall yRxyz) \\ \iff \mathfrak{A}, \sigma \vDash T_{x,fy}(\forall z(Px \wedge \forall yRxyz)) \\ \iff \mathfrak{A}, \sigma \vDash \forall z T_{x,fy}(Px \wedge \forall yRxyz) \\ \iff \mathfrak{A}, \sigma \vDash \forall z(T_{x,fy}(Px) \wedge T_{x,fy}(\forall yRxyz)) \\ \iff \mathfrak{A}, \sigma \vDash \forall z(Pfy \wedge \forall yRxyz) \end{array}$$

Note that the x in $\forall yRxyz$ is not substituted as $T_{x,fy}(\forall yRxyz) = \forall yRxyz$

Remark 2.7 Note that unlike F, $T_{x,g}(\varphi)$ might not be substitution-free as $T_{x,g}(\forall y\varphi) = \forall y\varphi$ where φ may still contain further substitutions, i.e., if not substitutable then we just stop, which also echos the spirit behind the standard substitution in Def. 2.1.

Note that in contrast to the propositional case, if $[t/x]$ is given the same semantics as the assignment operator $[x := t]$ as in dynamic logic [6], then the logic is different as the assignment operator can always change the model, in contrast with the syntactic substitution. To see the difference between the assignment operator and the substitution operator, consider a simple validity $[y/x]\forall yPx \leftrightarrow \forall yPx$ in our setting (as the substitution does nothing), the corresponding formula $[x := y]\forall yPx \leftrightarrow \forall yPx$ is not valid.

3 Axiom System RFSUB and its Completeness

We define the axiom system RFSUB in Table 3. Note that we no longer have meta-language substitutions, and the conditions for the axioms can be verified trivially. There are basically three groups of axioms and rules: the standard ones for first-order logic, the definition of substitutions (restricted to first-order terms), and the rules NEC, SUBFORM, SUBTERM taking care of the substitutions w.r.t. arbitrary terms such that we can rewrite them inside-out. Note that in SUBFORM and SUBTERM, we require the equality between terms in the premises to be *provable*, i.e., $\vdash ([fx/y]y) \approx fx$.

Axiom Schema	$\alpha \in \mathcal{L}_{\mathbf{FO}}$, $g \in \mathbf{T}_{\mathbf{FO}}$.
TAUT	All propositional logic axiom schemas
ID	$t \approx t$
SUBF	$t_0 \approx t'_0 \wedge \cdots \wedge t_n \approx t'_n \to ft_1 \dots t_n \approx ft'_1 \dots t'_n$
SUBI	$t_0 \approx t'_0 \wedge \cdots \wedge t_n \approx t'_n \to (Pt_0 \dots t_n \leftrightarrow Pt'_0 \dots t'_n)$
AK	$\forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$
AG	$\alpha \to \forall x\alpha$, if x is not free in α
INS	$\forall x\alpha \to [t/x]\alpha$
SSVARE	$([g/x]x) \approx g$
SSVARU	$([g/x]y) \approx y, y \neq x$
SSCONST	$([g/x]c) \approx c$
SSFUNC	$([g/x]ft_0 \dots t_k) \approx f[g/x]t_0 \dots [g/x]t_k$
SSATOM	$[g/x]Pt_0 \dots t_k \leftrightarrow P[g/x]t_0 \dots [g/x]t_k$
SSNEG	$[g/x]\neg\varphi \leftrightarrow \neg[g/x]\varphi$
SSCON	$[g/x](\psi \wedge \chi) \leftrightarrow ([g/x]\psi \wedge [g/x]\chi)$
SSQUANU	$[g/x]\forall y\varphi \leftrightarrow \forall y[g/x]\varphi$, if y does not occur in g
SSQUANE	$[g/x]\forall y\varphi \leftrightarrow \forall y\varphi$, if y occurs in g

Rule			
MP	$\dfrac{\varphi, \varphi \to \psi}{\psi}$	UG	$\dfrac{\vdash \varphi}{\vdash \forall x\varphi}$
NEC	$\dfrac{\vdash \varphi}{\vdash [t/x]\varphi}$		
SUBFORM	$\dfrac{\vdash t \approx t'}{\vdash [t/x]\varphi \leftrightarrow [t'/x]\varphi}$	SUBTERM	$\dfrac{\vdash t \approx t'}{\vdash [t/x]t'' \approx [t'/x]t''}$

Table 1
RFSUB system

We can show that the above axiom system is sound and complete via a reduction strategy similar to the one for DEL. We can syntactically reduce each term/formula by using axioms into a semantically equivalent substitution-free FO-term/formula. The details are left for the full version of this abstract.

As for future work, we can generalize the substitution operator to simultaneous substitutions for a list of variables. Then, it may be possible to compose sequential substitutions into one. We will also compare the logics based on different widely used notions of substitution for FOL. Note that the equality $\approx$ played an important role in our axiomatization in the reduction axioms for terms. What if we do not have the equality symbol in the language? Moreover, we can investigate first-order modal logic and propositional quantified modal logic extended with the corresponding substitution operators.

Acknowledgment We were inspired by a talk on substitutions of FOL given by Zhicheng Chen at Peking University. We thank the anonymous reviewers of AiML2024 for their very helpful comments and pointers, which improved the presentation of the abstract. We would also like to thank Zhiguang Zhao for explaining some technicalities of some related work in proof theory.

References

[1] Balco, S., G. Greco, A. Kurz, A. Moshier, A. Palmigiano and A. Tzimoulis, *First order logic properly displayed* (2021).
URL `http://arxiv.org/abs/2105.06877`

[2] Bednarczyk, M. A., *Logic of predicates with explicit substitutions*, in: *Proceedings of MFCS'96*, Springer, 1996, pp. 192–205.

[3] Enderton, H., "A Mathematical Introduction to Logic," Elsevier, 2001.

[4] Ghani, N., V. de Paiva and E. Ritter, *Linear explicit substitutions*, Log. J. IGPL **8** (2000), pp. 7–31.

[5] Halmos, P. R., "Algebraic logic," Chelsea Publishing Co., 1962.

[6] Harel, D., D. Kozen and J. Tiuryn, "Dynamic Logic," MIT Press, 2000.

[7] Henkin, L., J. D. Monk and A. Tarski, *Cylindric algebras* (1971).

[8] Kesner, D., *A theory of explicit substitutions with safe and full composition*, Log. Methods Comput. Sci. **5** (2009).

[9] Kracht, M., *Notes on substitution in first–order logic*, in: *First-order Logic Revisted*, Logos Verlag Berlin, 2004 .

[10] Mellies, P.-A., *Typed λ-calculi with explicit substitutions may not terminate*, in: *International Conference on Typed Lambda Calculi and Applications*, Springer, 1995, pp. 328–334.

[11] Pinter, C., *Cylindric algebras and algebras of substitutions*, Transactions of the American Mathematical Society **175** (1973), pp. 167–179.

[12] Pinter, C., *A simple algebra of first order logic*, Notre Dame Journal of Formal Logic **14** (1973), pp. 361–366.

[13] Pinter, C., *Terms in cylindric algebras*, Proceedings of the American Mathematical Society **40** (1973), pp. 568–572.

[14] Plaza, J. A., *Logics of public communications*, in: *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, 1989, pp. 201–216.

[15] Poigné, A., "Basic Category Theory," a **1**, Clarendon Press, Oxford, 1992 .

[16] Rose, K. H., "Explicit substitution: tutorial & survey," Computer Science Department, 1996.

[17] Tu, Y., S. Ghosh, F. Liu and D. Li, *A modal approach towards substitutions* (2023).

[18] Tzimoulis, A., "Algebraic and Proof-Theoretic Foundations of the Logics for Social Behaviour," Ph.D. thesis, TU Delft (2018).

[19] van Benthem, J., "Logical Dynamics of Information and Interaction," Cambridge University Press, 2011.

[20] van Benthem, J., J. van Eijck and B. P. Kooi, *Logics of communication and change*, Inf. Comput. **204** (2006), pp. 1620–1662.

[21] Wang, Y. and Q. Cao, *On axiomatizations of public announcement logic*, Synthese **190** (2013).

[22] Wang, Y., Y. Wei and J. Seligman, *Quantifier-free epistemic term-modal logic with assignment operator*, Ann. Pure Appl. Log. **173** (2022), p. 103071.

An Aretaic Approach to Deontic Logic

Joshua Yarrow [1]

Pembroke College, University of Cambridge

Sara L. Uckelman [2]

Dept. of Philosophy, Durham University

1 Virtue Ethics as a Foundation for Deontic Logic

This paper considers the following question: What would deontic logic look like if the underlying ethical theory was neither consequentialist nor deontological, but aretaic? [3] That is, our starting point is the Aristotelian idea that virtue consists in hitting the 'golden mean' between a vice of deficiency and a vice of excess of some characteristic. Upon this foundation, Rosalind Hursthouse made the first step towards a connection between the aretaic and the deontic:

Definition 1.1 (Criterion of Right Action (CRA)) *An action is* right *iff it is what a virtuous agent would characteristically (i.e., acting in character) do in the circumstances [2, p. 28].*

There are two advantages to this approach: Basing deontic logic on such a semantics answers a popular criticism of virtue ethics that it cannot provide moral guidance [2, p. 30] or that, if it can, the guidance that it does provide is inherently vague. At the same time, it would provide semantics for deontic logic with an easily understood interpretation, which is something that other deontic logics often struggle with.

2 Basic Language and Models

Taking inspiration from Meyer [4] we take a set of atomic propositions P and a set of atomic actions A and define an action language $\mathcal{A}$ which supports atomic actions (a), parallel execution ($\alpha \,\&\, \beta$), and free choice ($\alpha + \beta$). Our models are labeled transition systems $\mathfrak{M} = \langle W, R, V \rangle$, with the usual definitions. We define transition relations for action expressions:

$$\begin{aligned} R_a &:= \{\langle w, v\rangle \in W \times W \mid \langle w, a, v\rangle \in R\} \\ R_{\alpha\&\beta} &:= R_\alpha \cap R_\beta \\ R_{\alpha+\beta} &:= R_\alpha \cup R_\beta \end{aligned}$$

---

[1] jmky2@cam.ac.uk

[2] s.l.uckelman@durham.ac.uk

[3] This short paper derives from the undergraduate thesis of the first author, supervised by the second author.

This yields a multi-modal language $\mathcal{L}$ defined by the following grammar:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid [\alpha]\varphi$$

for which truth conditions are standard.

When adding an aretaic component to these semantics, the central conceptual move is to focus on *characteristics* of agents, which are defined as N-tuples (where $N \in \mathbb{N}$) of characteristics, C, which are (1) *ordered*, (2) *bounded*, and (3) *dense*. As such, the extent of possession of each characteristic in C is modelled as a value over the real interval $[-1, 1]$ where the lower bound of -1 represents limiting deficiency, the upper bound of 1 represents limiting excess, and the midpoint, 0, represents the 'golden mean'. `Char`, a characteristic assignment function, connects actions to the virtues and vices which they instantiate:

$$\texttt{Char} : W \times A \to [-1, 1]^N$$

`Char` takes a world and an atomic action and, for each characteristic in C gives the extent to which it is expected to have been instantiated in order to manifest the action in those circumstances. Thus we have a model $\mathfrak{M} = \langle W, R, V, \texttt{Char}\rangle$.

3 Strong, Weak, and Composite Obligation

Given the CRA, a virtuous agent would, by nature, do whichever action is the most balanced between the vices of excess and deficiency. This is equivalent to the output of the following function:

Definition 3.1 (`MinVice`) *The function* $\texttt{MinVice} : W \to \mathcal{P}(A)$*, which takes a world* w *and returns the set of executable atomic actions that have the minimum vice according to the norm of their characteristic assignment (the sum of the absolute values of their entries), is defined as follows:*

$$\begin{aligned} \texttt{MinVice}(w) = \{a \in A \mid & \text{ there is a } v \in W \text{ such that } wR_a v \\ & \text{and for all } b \in A \text{ where there is a } v' \in W \text{ such that } wR_b v', \\ & ||\texttt{Char}(w, a)|| \leq ||\texttt{Char}(w, b)||\} \end{aligned}$$

Due to ambiguities in the interpretation what a virtuous agent 'would do', with maximally-virtuous actions being executed either together or as alternatives, this affords both strong and weak definitions of obligation respectively:

Definition 3.2 (Strong Obligation)

$$\mathfrak{M}, w \vDash O^S(a_1 \,\&\, \cdots \,\&\, a_n) \quad \textit{iff} \quad \{a_1, \cdots, a_n\} \subseteq \texttt{MinVice}(w) \text{ and there is a } v \in W \text{ such that } wR_{(a_1 \& \cdots \& a_n)} v$$

Definition 3.3 (Weak Obligation)

$$\mathfrak{M}, w \vDash O^W(a_1 + \cdots + a_n) \quad \textit{iff} \quad \{a_1, \cdots, a_n\} = \texttt{MinVice}(w)$$

Both of these definitions track an important part of what obligation means, but are by themselves insufficient, as each definition is only able to accommodate one of the connectives from the action language, whereas we want *both*. Reflecting on these definitions, it becomes clear that the virtuousness of individual atomic actions is underdetermined as there are different plausible assignments based on what the agent does otherwise; consequently, we expand the domain of `Char` so that it naturally talks about combinations of atomic actions performed together:

$$\texttt{Char} : W \times \mathcal{P}(A) \rightarrow [-1, 1]^N$$

This allows us to define a composite definition of obligation, incorporating aspects of both weak and strong obligation, for which we need to define two auxiliary functions. The first takes a world and returns all the sets of atomic actions that can be jointly executed at that world. The second performs the equivalent of the `MinVice` function on these sets — it takes a world and returns all the sets of jointly-executable actions which together are assigned the minimal characteristic norm at that world:

Definition 3.4 *The function* $\textit{ActComplex} : W \rightarrow \mathcal{P}(\mathcal{P}(A))$ *is defined as*

$$\begin{aligned} \textit{ActComplex}(w) = \{&\{a_1, \cdots, a_n\} \in \mathcal{P}(A) \mid \textit{there is a } v \in W \\ &\textit{such that } wR_{(a_1 \& \cdots \& a_n)} v\} \end{aligned}$$

Definition 3.5 *The function* $\textit{MinViceComplex} : W \rightarrow \mathcal{P}(\mathcal{P}(A))$ *(abbreviated MVC) is defined as:*

$$\begin{aligned} \textit{MVC}(w) = \{&\boldsymbol{a} \in \textit{ActComplex}(w) \mid \textit{for all } \boldsymbol{b} \in \textit{ActComplex}(w), \\ &||\textit{Char}(w, \boldsymbol{a})|| \leq ||\textit{Char}(w, \boldsymbol{b})||\} \end{aligned}$$

The bold type used above indicates sets of atomic actions. We also introduce a normal form representation for action expressions, which we'll use to define truth conditions for the composite obligation of arbitrary action expressions:

Definition 3.6 (Choice Normal Form) *An action expression in* $\mathcal{A}$ *is in* Choice Normal Form *if it is of the form*

$$((a_1^1 \& \cdots \& a_{n_1}^1) + \cdots + (a_1^m \& \cdots \& a_{n_m}^m)).$$

The following lemmas are used to prove that all action expressions can be represented in Choice Normal Form (proofs of these are straightforward and omitted due to reasons of space):

Lemma 3.7 *Parallel execution is distributive over choice.*

Lemma 3.8 *Parallel execution is associative.*

Lemma 3.9 *Choice is associative.*

Lemma 3.10 *Any action expression in* $\mathcal{A}$ *is equivalent to one in Choice Normal Form.*

We then define a composite obligation by lifting weak obligation to the level of parallel-executed atomic actions rather than mere atomic actions:

Definition 3.11 (Composite Obligation) *Let* $\alpha := ((a_1^1 \,\&\, \cdots \,\&\, a_{n_1}^1) + \cdots + (a_1^m \,\&\, \cdots \,\&\, a_{n_m}^m))$ *be in Choice Normal Form. Then:*

$$\mathfrak{M}, w \vDash O(\alpha) \quad \textit{iff} \quad \{\{a_1^1, \cdots, a_{n_1}^1\}, \cdots, \{a_1^m, \cdots, a_{n_m}^m\}\} = \mathit{MVC}(w)$$

This definition brings together the benefits of both strong and weak obligation. It also forces us to take a more restricted reading of the obligation operator, which may at first seem unnatural — every mention of an action complex inside an obligatory choice must be prefixed with the words '*just*' or '*only*'. That is to say, free choice is to be interpreted as strictly exclusive such that $O(\alpha + \beta)$ means 'You ought to do either *just* α or *just* β'. This does not diminish the expressivity of the language, however, because the joint action can always be included as a third choice such as in $O(\alpha + \beta + (\alpha \,\&\, \beta))$.

4 Permission and Prohibition

In traditional deontic logics, obligation often is taken as primitive whilst permission is defined as its dual. Since we do not have action negation, it is not possible to define permission in this way. Instead, we treat permission as an independent, primitive operator — one which, nevertheless, still validates $P(\alpha+\beta) \to P(\alpha) \wedge P(\beta)$ (items in a permissible choice must all be individually permissible) and, at least at first, $O(\alpha) \to P(\alpha)$ (what is obligatory is permissible). The eventual rejection of this latter premise will become a central philosophical point. We consider several natural ways to define permission:

Permission as thresholding.

$$\mathfrak{M}, w \vDash P^1(a) \quad \text{iff} \quad ||\texttt{Char}(a)|| \leq k \times N$$

where $k \in [0, 1]$ is a user-defined threshold constant which determines the relative region of permissibility; k is then scaled by the maximum possible norm N (i.e. the size of the characteristic vector). This approach is motivated by the suggestion that virtuous agents are still fallible [6, p. 146]. However, this is arguably at odds with virtue ethics in terms of psychology.

Permission as virtue in development. Rather than having a region of actions which are simply 'good enough', we can consider the permissible as forming the realm of *virtue in development* — the idea being that on a virtue ethical account, the end goal is to become a perfectly virtuous agent, so an act is permitted if it contributes to the development of an agent's virtue towards that end of being perfectly virtuous. In other words, one often must pass through stages of imperfection to arrive at virtue:

$$\mathfrak{M}, w \vDash P^2(a_1 \,\&\, \cdots \,\&\, a_n) \quad \text{iff} \quad \text{there is some } \{b_1, \cdots, b_m\} \in \texttt{MVC}(w) \text{ such that } \{a_1, \cdots, a_n\} \subseteq \{b_1, \cdots, b_m\}$$

$$\mathfrak{M}, w \vDash P^3(a_1 \,\&\, \cdots \,\&\, a_n) \quad \text{iff} \quad \text{there is some } \{b_1, \cdots, b_m\} \in \texttt{MVC}(w) \text{ such that } \{a_1, \cdots, a_n\} \cap \{b_1, \cdots, b_m\} \neq \emptyset$$

P^2 begins to capture the idea of virtue in development, but is biased in favour of agents who are apathetic. P^3 fixes this asymmetry, but at the cost of being far too permissive. This leads us to our final possibility.

Improvement-based Permission. In order to give a good account of permission as *virtue in development*, we need representations of different agents' individual characteristics so that we can determine which actions for them count as a 'development' and which ones lead them further away from virtue than they already are. Let $\texttt{AgentChar} \in [-1,1]^N$ be a vector which represents the characteristics that the agent actually possesses. If all elements of `AgentChar` are zero then they are a perfectly virtuous agent. Our final definition claims that all actions on the agent's path towards virtue are permitted:

Definition 4.1 (Improvement-based Permission)

$$\begin{array}{lll} \mathfrak{M}, w \vDash P(\alpha) & \textit{iff} & ((a_1^1 \,\&\, \cdots \,\&\, a_{n_1}^1) + \cdots + (a_1^m \,\&\, \cdots \,\&\, a_{n_m}^m)) \\ & & \textit{is the Choice Normal Form of } \alpha \\ & & \textit{and for all } \mathbf{a} \in \{\{a_1^1, \cdots, a_{n_1}^1\}, \cdots, \{a_1^m, \cdots, a_{n_m}^m\}\} \\ & & \textit{and all } i \in [\![1, N]\!], |\mathtt{Char}(w, \mathbf{a})_i| \leq |\mathtt{AgentChar}_i| \end{array}$$

This makes permission a relative concept whilst keeping obligation fixed for all agents, and divorces permission from obligation to the extent that $O(\alpha) \to P(\alpha)$ is no longer a theorem. We claim that obligation and permission demonstrate two distinct connections between virtue and the justification of actions — obligatory actions are justified as immediate manifestations of virtue whilst permissible actions are justified as contributions to the cultivation of virtue over a longer time scale. As such, our logic exposes the reality of cases when an agent's impermissible obligations will inevitably harm their moral character.

To complete the set of deontic operators, we can now use propositional negation to define prohibition (denoted by F for 'forbid') similar to the way it is done traditionally but now excluding both obligation and permission due to their independence, as $F(\alpha) := \neg P(\alpha) \wedge \neg O(\alpha)$.

5 Conditional Obligation

To illustrate the potential our system has for expressive extensions, we consider obligations that are conditional on either some aspect of a state of affairs or partially specified actions. Dyadic deontic logics such as Hansson's DSDL3 [5] accommodate the former kind of conditional obligations by giving the deontic operators a second propositional parameter that serves as the conditions. The notation is inspired by conditional probability in the sense that 'It is obligatory that φ given ψ' is written as $O(\varphi|\psi)$ [3] and the system works as follows:

> "Hansson presents a possible worlds semantics in which all worlds are ordered by a preference (betterness) relation. $O(\varphi|\psi)$ is then defined true if φ is true in the best ψ-worlds." [1]

In our system, however, it is actions which are subject to a preference ordering rather than worlds. It is quite clear that a ψ-world is a world where ψ is true, but what is a ψ-action? It seems reasonable to suggest that we resolve this ambiguity by turning to alethic modalities. This gives us four types of conditions: necessary consequence; possible consequence; atomic action inclusion; and atomic action exclusion. In Definition 5.3 these are denoted using box, diamond, plus, and minus symbols respectively, with the former two applied to a formulae of $\mathcal{L}$ and the latter two applied to sets of atomic actions.

Definition 5.1 *The function* $\texttt{CondActComplex}: W \times \mathcal{L} \times \mathcal{L} \times \mathcal{P}(A) \times \mathcal{P}(A) \to \mathcal{P}(\mathcal{P}(A))$ *is defined as*

$$\begin{aligned}\texttt{CondActComplex}\ (w, \varphi, \psi, \mathbf{c}, \mathbf{d}) = \{ & \{a_1, \cdots, a_n\} \in \mathcal{P}(A) \mid \\ & \mathbf{c} \subseteq \{a_1, \cdots, a_n\} \\ & \textit{and } \mathbf{d} \cap \{a_1, \cdots, a_n\} = \emptyset \\ & \textit{and there is a } v \in W \textit{ such that } wR_{(a_1 \& \cdots \& a_n)}v \textit{ and } \mathfrak{M}, v \vDash \psi \\ & \textit{and for all } v' \in W \textit{ such that } wR_{(a_1 \& \cdots \& a_n)}v',\ \mathfrak{M}, v' \vDash \varphi \quad \}\end{aligned}$$

Definition 5.2 *The function* $\texttt{CondMVC}: W \times \mathcal{L} \times \mathcal{L} \times \mathcal{P}(A) \times \mathcal{P}(A) \to \mathcal{P}(\mathcal{P}(A))$ *is defined as*

$$\begin{aligned}\texttt{CondMVC}(w, \varphi, \psi, \mathbf{c}, \mathbf{d}) = \{\ & \mathbf{a} \in \texttt{CondActComplex}(w, \varphi, \psi, \mathbf{c}, \mathbf{d}) \mid \\ \textit{for all}\ & \mathbf{b} \in \texttt{CondActComplex}(w, \varphi, \psi, \mathbf{c}, \mathbf{d}), \\ & ||\texttt{Char}(w, \mathbf{a})|| \leq ||\texttt{Char}(w, \mathbf{b})|| \quad \}\end{aligned}$$

Definition 5.3 (Conditional Obligation)

$$\begin{aligned}\mathfrak{M}, w \vDash\ & O(\alpha \mid \Box\varphi, \Diamond\psi, +\{c_1, \cdots, c_{l^+}\}, -\{d_1, \cdots, d_{l^-}\}) \\ \textit{iff}\ & ((a_1^1 \,\&\, \cdots \,\&\, a_{n_1}^1) + \cdots + (a_1^m \,\&\, \cdots \,\&\, a_{n_m}^m)) \\ & \textit{is the Choice Normal Form of } \alpha \\ \textit{and}\ & \{\{a_1^1, \cdots, a_{n_1}^1\}, \cdots, \{a_1^m, \cdots, a_{n_m}^m\}\} \\ =\ & \texttt{CondMVC}(w, \varphi, \psi, \{c_1, \cdots, c_{l^+}\}, \{d_1, \cdots, d_{l^-}\})\end{aligned}$$

We have found in our wider exploration that this conditional notation supports the expression of sophisticated moral affairs, such as contrary-to-duty obligations. Empty or trivial conditions may be omitted to streamline the notation; in the most extreme case we have that $O(\alpha) \triangleq O(\alpha \mid \Box\top, \Diamond\top, +\emptyset, -\emptyset)$, which behaves exactly the same as the composite obligation from Definition 3.11.

6 Conclusion and Future work

The primary contributions of this paper are (1) providing a semantics for deontic logic rooted in Aristotelian virtue ethics and (2) three definitions of deontic operators using these semantics. The semantics presented here generate distinctively few theorems to connect propositional and deontic operators; this is a strength in disguise as such theorems have often been the downfall of other deontic logics where, for example, trivial disjunctions permit diabolical choices. In the wider work that this short paper is based on, we also explored dynamically updating an agent's virtue profile to model virtue acquisition and its effect on deontic statements. Other areas we considered for future work include (1) expanding our semantics to virtue epistemology [7]; (2) integration with a logic of norms [1]; (3) multi-agent models; (4) modelling virtue with neural networks.

References

[1] Hansen, J., G. Pigozzi and L. van der Torre, *Ten Philosophical Problems in Deontic Logic*, in: G. Boella, L. van der Torre and H. Verhagen, editors, *Normative Multi-agent Systems*, Dagstuhl Seminar Proceedings (DagSemProc) **7122** (2007), pp. 1–26.

[2] Hursthouse, R., "On Virtues Ethics," Oxford University Press, 2000.

[3] McNamara, P. and F. Van De Putte, *Deontic Logic*, in: E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2022, Fall 2022 edition .

[4] Meyer, J.-J., *Deontic dynamic logic: a retrospective*, Filosofiska Notiser **6** (2019), pp. 63–76.

[5] Spohn, W., *An analysis of hansson's dyadic deontic logic*, Journal of Philosophical Logic **4** (1975), pp. 237–252.

[6] Swanton, C., "Target Centred Virtue Ethics," Oxford University Press, 2021.

[7] Zagzebski, L. T., "Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge," Cambridge: Cambridge University Press, 1996.