Journal of applied logics (Print) ISSN 2631-9810
Journal of applied logics (Online) ISSN 2631-9829

Journal of Applied Logics

The IfCoLog Journal of Logics and their Applications

Volume 10 • Issue 4 • July 2023

Available online at
www.collegepublications.co.uk/journals/ifcolog/
Free open access

JOURNAL OF APPLIED LOGICS - IFCOLOG
JOURNAL OF LOGICS AND THEIR APPLICATIONS

Volume 10, Number 4

July 2023

Disclaimer
Statements of fact and opinion in the articles in Journal of Applied Logics - IfCoLog Journal of Logics and their Applications (JALs-FLAP) are those of the respective authors and contributors and not of the JALs-FLAP. Neither College Publications nor the JALs-FLAP make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability for any errors or omissions that may be made. The reader should make his/her own evaluation as to the appropriateness or otherwise of any experimental technique described.

ISBN 978-1-84890-434-7
ISSN (E) 2631-9829
ISSN (P) 2631-9810

College Publications
Scientific Director: Dov Gabbay
Managing Director: Jane Spurr

http://www.collegepublications.co.uk

Editorial Board

Scope and Submissions

This journal considers submission in all areas of pure and applied logic, including:

pure logical systems
proof theory
constructive logic
categorical logic
modal and temporal logic
model theory
recursion theory
type theory
nominal theory
nonclassical logics
nonmonotonic logic
numerical and uncertainty reasoning
logic and AI
foundations of logic programming
belief change/revision
systems of knowledge and belief
logics and semantics of programming
specification and verification
agent theory
databases
dynamic logic
quantum logic
algebraic logic
logic and cognition
probabilistic logic
logic and networks
neuro-logical systems
complexity
argumentation theory
logic and computation
logic and language
logic engineering
knowledge-based systems
automated reasoning
knowledge representation
logic in hardware and VLSI
natural language
concurrent computation
planning

This journal will also consider papers on the application of logic in other subject areas: philosophy, cognitive science, physics etc. provided they have some formal content.

Submissions should be sent to Jane Spurr (jane@janespurr.net) as a pdf file, preferably compiled in LaTeX using the IFCoLog class file.

Contents

ARTICLES

Extending a Call-by-Value Calculus Based on Bilateralism with Dynamic Binding

Tatsuya Abe
STAIR Lab, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba, 275-0016, Japan.
abet@stair.center

Daisuke Kimura
Department of Information Science, Toho University, 2-2-1 Miyama, Funabashi, Chiba, 274-8510, Japan.
kmr@is.sci.toho-u.ac.jp

Abstract

Bilateral natural deduction has judgments about not only acceptance but also rejection. A typed lambda-calculus corresponding to bilateral natural deduction on the formulae-as-types notion has not only first-class expressions but also first-class continuations. Functions on continuations in the calculus have co-implication types. In this paper, we extend the call-by-value variant of the calculus with dynamic binding. In the extended calculus, we can more delicately operate and observe continuations in the computation. We demonstrate that a language with control operators for delimited continuations can be compiled into the extended calculus.

Keywords: formulae-as-types notion, bilateralism, dynamic binding, delimited continuation, control operator

1 Introduction

Natural deduction for minimal logic consists of the harmonious pairs of the introduction and elimination rules about the connectives in the sense of [17, 4]. The normalization procedure also makes sense. However, natural deduction for classical logic has often been criticized for having the inharmonious pairs in proof-theoretic semantics [8]. Nevertheless, Prawitz modified classical natural deduction formulated by Gentzen and gave a normalization procedure [16], which has been studied and improved [2, 23].

Bilateral natural deduction is obtained by repairing the inharmony [18]. It has judgments about not only (ordinary) acceptance but also rejection. The normalization procedure in bilateral natural deduction can be more naturally provided. We constructed a typed λ-calculus λ_{conf} corresponding to bilateral natural deduction on the formulae-as-types notion [1].

The λ_{conf}-calculus has not only (ordinary) first-class expressions but also first-class continuations. It has also (ordinary) expression variables and continuation variables. It has also abstractions of expressions and continuations by variables symmetrically. Surely, abstractions of expressions by expression variables are functions on expressions of implication types. Similarly, abstractions of continuations by continuation variables are functions on continuations of co-implication types. Each reduction corresponds to a step of the normalization procedure in bilateral natural deduction.

A frequently asked question is what abstractions of continuations by continuation variables of co-implication types are used for. Sakaue and Asai, and Ueda and Asai adopted the negation type $\neg T$ to denote the type of a continuation for an expression of the type T [19, 22]. Functions on continuations have not co-implication types but implication types consisting of negation types. Also, a well-known calculus with first-class continuations $\bar{\lambda}\mu\tilde{\mu}$ corresponding to sequent calculus for classical logic on the formulae-as-types notion has no abstractions of continuations by continuation variables of co-implication types, although they mentioned such an extension [5].

Moreover, although the control operator `call/cc` for unlimited continuations [21] characterized by the reduction $\langle \mathtt{call/cc}\, V \mid C\rangle \rightsquigarrow_v{}^+ \langle V \ulcorner \lambda\_.C \urcorner \mid C\rangle$ is definable in the call-by-value variant of λ_{conf} where the notation will be explained in Section 2, abstractions of continuations by continuation variables of the co-implication types are not essentially used because the current continuation is just discarded, as seen in $\lambda\_.C$ when the captured continuation C is used. With these facts, we have even received a comment that λ_{conf} based on bilateralism (moreover, bilateral natural deduction itself) is just a well-known system based on unilateralism in disguise.

However, the λ_{conf}-calculus corresponding to bilateral natural deduction has an advantage that first-class continuations are reasoned in logic from the beginning. In conventional work based on unilateralism to extend the simply typed λ-calculus corresponding to minimal logic, coterms and reductions are minimally added to the minimal calculus [15]. Similarly, coterms and reductions are added in the existing extensions with delimited continuations [3, 7]. The extended reduction relations including the added coterms are logically reasoned through multiple continuation passing style (CPS) transformations.

We do not care about the minimality. From the beginning, we adopt bilateral natural deduction for classical logic as a base to be extended. In extending λ_{conf} with delimited continuations, we maintain the reduction rules, which have already been reasoned in logic. Our previous work has also shown that the control operator for unlimited continuations is definable in bilateral natural deduction [1]. This paper aims to identify what functionality extends unlimited continuations to delimited continuations.

We first investigated why λ_{conf} cannot support delimited continuations. Characteristics of delimited continuations, distinguished from unlimited continuations, are replicability and composability. Because continuations are first-class objects in λ_{conf} corresponding to a non-linear logic, replicability has been obtained. Because functions on continuations are also first-class objects in λ_{conf} and functions are surely composable, composability has also been obtained. Finally, we realized that λ_{conf} cannot support delimited continuations because it cannot identify continuations from control operators until delimiters in the computation. In other words, it is not possible to dynamically define functions on continuations in the computation in λ_{conf}.

In order to enable continuation function definitions dynamically in the computation, let us remember (ordinary) expression function definitions dynamically in the computation. Abstractions of expressions by expression variables are functions on expressions of implication types. An expression $\lambda x.E$ is a function throughout the computation because xs that occur in E are statically bound by λx. For example, $\lambda x.y$ is a constant function that returns y and $(\lambda y.\lambda x.y)(x+1)$ is reduced to a constant function $\lambda z.(x+1)$ that returns $x+1$ by renaming x to z. Dynamic binding provides expression function definitions dynamically in the computation of an expression. The expression $(\lambda y.\lambda x.y)(x+1)$ is reduced to the successor function $\lambda x.(x+1)$ because the x in $x+1$ is dynamically bound.

We apply it to functions on continuations. Dynamic binding by continuation variables provides continuation function definitions dynamically in the computation of a continuation. Specifically, we extend the call-by-value variant of λ_{conf} with dynamic binding of a continuation variable.

We use the variable level notion which provides dynamic binding by expression variables in Sato et al.'s calculus [20]. Each variable has a level, and each level has infinitely many variables. They adopt a higher-level reduction strategy different from the call-by-value strategy. Therefore, the variable level notion cannot be immediately applied to λ_{conf}. We elaborate on the technical details, in particular α-equivalence and substitution, to extend the call-by-value calculus with dynamic binding using the variable level notion.

The remainder of this paper is organized as follows: Section 2 introduces the

$$
\begin{array}{ll}
\text{(types)} & T ::= o \mid T \to T \mid T \leftarrow T \\
\text{(expressions)} & E ::= x \mid \lambda x.E \mid EE \mid \mu a.D \\
\text{(continuations)} & C ::= a \mid \lambda a.C \mid CC \mid \mu x.D \\
\text{(configurations)} & D ::= \langle E \mid C \rangle \\
\text{(syntactical objects)} & B ::= E \mid C \mid D
\end{array}
$$

Figure 1: The λ-calculus with first-class configurations λ_{conf}.

$$
\frac{\Xi \vdash_{+} E\colon T \qquad \Xi \vdash_{-} C\colon T}{\Xi \vdash_{\mathrm{o}} \langle E \mid C \rangle}\ \text{(Non-contradiction)}
$$

$$
\frac{\Xi, a\colon T \vdash_{\mathrm{o}} D}{\Xi \vdash_{+} \mu a.D\colon T}\ (\text{Reductio}_{+}) \qquad \frac{\Xi, x\colon T \vdash_{\mathrm{o}} D}{\Xi \vdash_{-} \mu x.D\colon T}\ (\text{Reductio}_{-})
$$

$$
\frac{}{\Xi, x\colon T \vdash_{+} x\colon T} \quad \frac{\Xi, x\colon T_0 \vdash_{+} E\colon T_1}{\Xi \vdash_{+} \lambda x.E\colon T_0 \to T_1} \quad \frac{}{\Xi, a\colon T \vdash_{-} a\colon T} \quad \frac{\Xi, a\colon A_1 \vdash_{-} C\colon T_0}{\Xi \vdash_{-} \lambda a.C\colon T_0 \leftarrow T_1}
$$

$$
\frac{\Xi \vdash_{+} E_0\colon T_0 \to T_1 \qquad \Xi \vdash_{+} E_1\colon T_0}{\Xi \vdash_{+} E_0 E_1\colon T_1} \qquad \frac{\Xi \vdash_{-} C_0\colon T_0 \leftarrow T_1 \qquad \Xi \vdash_{-} C_1\colon T_1}{\Xi \vdash_{-} C_0 C_1\colon T_0}
$$

Figure 2: The type system.

call-by-value variant of λ_{conf} corresponding to bilateral natural deduction. Section 3 introduces the variable level notion, which can be used to manage a dynamically bound variable. Section 4 extends the call-by-value calculus with dynamic binding using the variable level notion. Section 5 shows that a typed language with control operators for delimited continuations can be compiled into the extended calculus. Section 6 compares related work to clarify the contributions of this paper. Section 7 concludes the paper by identifying future research directions.

2 The typed lambda-calculus based on bilateralism

Figures 1 and 2 show types, expressions, continuations, configurations, and the type system of the typed λ-calculus with first-class configurations λ_{conf} [1]. Two kinds of variables, x and a, exist for expressions and continuations.

First, we ignore syntactical objects in the type system in Figure 2 to regard it as a proof system. A judgment is of the form $\Xi \vdash_+ \_\!: T$, $\Xi \vdash_- \_\!: T$, or $\Xi \vdash_\mathrm{o} \_$ where Ξ is a function from variables to types. The first and second judgments mean that T is accepted and rejected, respectively. The third judgment is neutral. Reductio ad absurdum is split into (Reductio$_+$) and (Reductio$_-$) in Figure 2. We can derive the converse of an assumption from the third judgment using the inference rules.

The introduction and elimination rules of the implication connective $\rightarrow$ are standard. We do not consider other standard logical connectives such as $\wedge$, $\vee$, and $\neg$ because they are beyond the scope of this paper. It is noteworthy that it has the introduction and elimination rules of the co-implication connective $\leftarrow$. Because $T \leftarrow T'$ has the same truth value as $T \wedge \neg\, T'$ in classical logic, the co-implication connective is also called but-not. Rumfitt did not consider the co-implication connective when he proposed bilateral natural deduction [18]. However, we adopt it as the dual of the implication connective $\rightarrow$ and use it for giving function types on continuations.

Next, we see syntactical objects in the type system. Expression variables, λ-abstractions of expressions, and applications of expressions are standard. Those of continuations are symmetrically defined. A configuration consists of an expression and a continuation of the same type. A μ-abstraction of a configuration by a continuation variable is an expression. A μ-abstraction of a configuration by an expression variable is a continuation.

Finally, we reason about a computational aspect of λ_{conf}. In the following, we give up maintaining the symmetry between expressions and continuations. We add expression constants c and the continuation constant $\bullet$ of the type o, for example, corresponding to the integers and the bottom of the stack consisting of continuations of the integer type, respectively, to λ_{conf} in Figure 1. We also define the set of values as shown in Figure 3.

The call-by-value and call-by-name equalities are provided in our previous work. In addition, a call-by-value reduction relation of an extended λ-calculus is also defined. The reduction relation $\rightsquigarrow_v$ in Figure 3 is its restriction to the base calculus. $D \not\rightsquigarrow_v$ means no D' such that $D \rightsquigarrow_v D'$. The reduction relation $\rightsquigarrow_v$ is defined based on the following strategy:

1. If the argument of the function is a value, then the function is applied to the value.

2. If the argument of the function is not a value, then the continuation including the function is pushed onto the stack consisting of the continuations for evaluating the argument.

3. If the expression on the function position has not yet been a function, then the

$$
\begin{array}{lll}
\text{(expressions)} & E ::= \cdots \mid c & \text{(typing rules)} \\
\text{(continuations)} & C ::= \cdots \mid \bullet & \\
\text{(values)} & V ::= c \mid x \mid \lambda x.E & \dfrac{}{\varXi \vdash_{+} c \colon o} \qquad \dfrac{}{\varXi \vdash_{-} \bullet \colon o}
\end{array}
$$

$$
\begin{array}{rll}
\langle (\lambda x.E)V \mid C\rangle & \rightsquigarrow_v \langle [V/x]E \mid C\rangle & \\
\langle V E \mid C\rangle & \rightsquigarrow_v \langle E \mid \mu x.\langle V x \mid C\rangle\rangle & \\
& & \text{if } E \text{ is not a value and } x \text{ is fresh} \\
\langle E_0 E_1 \mid C\rangle & \rightsquigarrow_v \langle E_0 \mid \mu x.\langle x E_1 \mid C\rangle\rangle & \\
& & \text{if } E_0 \text{ is not a value and } x \text{ is fresh} \\
\langle \mu a.D_0 \mid C\rangle & \rightsquigarrow_v \langle \mu a.D_1 \mid C\rangle & \text{if } D_0 \rightsquigarrow_v D_1 \\
\langle \mu a.D \mid C\rangle & \rightsquigarrow_v [C/a]D & \text{if } D \not\rightsquigarrow_v \\
\langle V \mid (\lambda a.C_0) C_1 C_{2,0} \cdots C_{2,n-1}\rangle & \rightsquigarrow_v \langle V \mid ([C_1/a]C_0) C_{2,0} \cdots C_{2,n-1}\rangle & \\
\langle V \mid (\mu x.D) C_1 C_{2,0} \cdots C_{2,n-1}\rangle & \rightsquigarrow_v [\mu a.\langle V \mid a C_1 C_{2,0} \cdots C_{2,n-1}\rangle/x]D & \text{if } a \text{ is fresh} \\
\langle V \mid \mu x.D\rangle & \rightsquigarrow_v [V/x]D &
\end{array}
$$

Figure 3: The call-by-value λ-calculus with first-class configurations.

continuation including the argument is pushed onto the stack for evaluating the expression.

4. The μ-abstraction by the continuation variable is not a value, and the configuration abstracted by the continuation variable is reduced under the condition that the continuation variable is free.

5. If the expression is a value, then the continuation is reduced, and the top continuation of the stack consisting of continuations is popped.

The reduction relation $\rightsquigarrow_v$ has the following basic properties:

Proposition 2.1. *The reduction relation $\rightsquigarrow_v$ is deterministic.*

Proof. This is because each configuration can be applied to the unique reduction rule. □

Lemma 2.2. *The substitution lemma holds, that is,*

1. *$\varXi \vdash_{+} E \colon T'$ and $\varXi, x \colon T' \vdash B \colon T$ imply $\varXi \vdash [E/x]B \colon T$ and*

2. $\varXi \vdash_{\_} C\colon T'$ *and* $\varXi, a\colon T' \vdash B\colon T$ *imply* $\varXi \vdash [C/a]B\colon T$*.*

Proof. We prove the lemma by induction on derivations of $\varXi, x\colon T' \vdash B\colon T$ and $\varXi, a\colon T' \vdash B\colon T$.
1) Let B be a variable. By the definition of substitution, it is obvious.
2) Let B and T be $\lambda x_0.E_0$ and $T_0 \to T_1$, respectively, where we take x_0 so as not to crash x. By induction hypothesis, $\varXi, x_0\colon T_0 \vdash [E/x]B\colon T_1$ holds. By the introduction rule of $\to$, $\varXi \vdash \lambda x_0.[E/x]B\colon T_0 \to T_1$ holds. By the definition of substitution, $\varXi \vdash [E/x]\lambda x_0.B\colon T_0 \to T_1$ holds. Cases of a λ-abstraction by a continuation variable, a μ-abstraction by a continuation variable, and a μ-abstraction by an expression variable are similar. Cases of substituting a continuation variable with a continuation are also similar.
3) Let B be E_0E_1. There exists T_0 such that $\varXi, x\colon T' \vdash E_0\colon T_0 \to T$ such that $\varXi, x\colon T' \vdash E_1\colon T_0$. By induction hypothesis, $\varXi \vdash [E/x]E_0\colon T_0 \to T$ and $\varXi \vdash [E/x]E_0\colon T_0$ hold. By the elimination rule of $\to$, $\varXi \vdash [E/x]E_0[E/x]E_1\colon T$ holds. By the definition of substitution, $\varXi \vdash [E/x](E_0E_1)\colon T$ holds. Cases of applying a continuation function to a continuation and a configuration consisting of an expression and a continuation are similar. Cases of substituting a continuation variable with a continuation are also similar. ☐

Theorem 2.3. $\varXi \vdash_{\mathrm{o}} D$ *and* $D \rightsquigarrow_v D'$ *imply* $\varXi \vdash_{\mathrm{o}} D'$*.*

Proof. We prove it by cases of the reduction relation. We use Lemma 2.2 in cases where reduction rules contain substitutions. The other cases are routines that just see types regarding the definition of typing rules. ☐

Theorem 2.4. *Assume* $\varnothing \vdash_{\mathrm{o}} D_0$*. Then, for any* D_1 *that occurs in its reduction sequence and is not a configuration consisting of a value and* $\bullet$*, there exists* D_2 *such that* $D_1 \rightsquigarrow_v D_2$*.*

Proof. We prove it by cases of configurations. First, we consider $\langle E \mid C\rangle$ where E is not a value. We can apply a reduction rule to it. Next, we consider $\langle V \mid C\rangle$. If it cannot be reduced, C is $\bullet$ or $\lambda a.C'$ for some a and C'. The assumption of the proposition satisfies neither case because $\langle V \mid \lambda a.C'\rangle$ has no type. ☐

We define encodings from expressions and continuations to continuations and expressions, respectively, as follows:

$$\llcorner E \lrcorner \equiv \lambda a.\mu x.\langle Ex \mid a\rangle \qquad\qquad \ulcorner C \urcorner \equiv \lambda x.\mu a.\langle x \mid Ca\rangle$$

where x and a are fresh.

Proposition 2.5. *The following inferences are derivable:*

$$\frac{\varXi \vdash_+ E\colon T_0 \to T_1}{\varXi \vdash_- \llcorner E \lrcorner\colon T_0 \leftarrow T_1} \qquad \frac{\varXi \vdash_- C\colon T_0 \leftarrow T_1}{\varXi \vdash_+ \ulcorner C \urcorner\colon T_0 \to T_1} \ .$$

Proof. They are derivable as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\varXi' \vdash_+ E\colon T_0 \to T_1 \qquad \varXi' \vdash_+ x\colon T_0}{\varXi' \vdash_+ Ex\colon T_1} \qquad \varXi' \vdash_- a\colon T_1}{\varXi, x\colon T_0, a\colon T_1 \vdash_{\mathrm{o}} \langle Ex \mid a\rangle}}{\varXi, a\colon T_1 \vdash_- \mu x.\langle Ex \mid a\rangle\colon T_0}}{\varXi \vdash_- \lambda a.\mu x.\langle Ex \mid a\rangle\colon T_0 \leftarrow T_1}$$

$$\dfrac{\dfrac{\dfrac{\varXi' \vdash_+ x\colon T_0 \qquad \dfrac{\varXi' \vdash_- C\colon T_0 \leftarrow T_1 \qquad \varXi' \vdash_- a\colon T_1}{\varXi' \vdash_- Ca\colon T_0}}{\varXi, x\colon T_0, a\colon T_1 \vdash_{\mathrm{o}} \langle x \mid Ca\rangle}}{\varXi, x\colon T_0 \vdash_+ \mu a.\langle x \mid Ca\rangle\colon T_1}}{\varXi \vdash_+ \lambda x.\mu a.\langle x \mid Ca\rangle\colon T_0 \to T_1}$$

where $\varXi'$ is $\varXi, x\colon T_0, a\colon T_1$. □

Proposition 2.6. *The encodings between expressions and continuations work under the reduction relation as follows:*

1. $\langle \ulcorner C_0 \urcorner V \mid C_1\rangle \rightsquigarrow_v \langle V \mid C_0\, C_1\rangle$,
2. $\langle V \mid \llcorner E \lrcorner C\rangle \rightsquigarrow_v \langle EV \mid C\rangle$,
3. $\langle \ulcorner \llcorner E \lrcorner \urcorner V \mid C\rangle \rightsquigarrow_v \langle EV \mid C\rangle$, *and*
4. $\langle V \mid \llcorner \ulcorner C_0 \urcorner \lrcorner C_1\rangle \rightsquigarrow_v \langle V \mid C_0\, C_1\rangle$.

Proof. The first and second statements hold immediately from the definition of $\rightsquigarrow_v$, $\ulcorner C \urcorner$, and $\llcorner E \lrcorner$ as follows:

$$\begin{array}{ll}
\langle \ulcorner C_0 \urcorner V \mid C_1\rangle & \langle V \mid \llcorner E \lrcorner C\rangle \\
\equiv \langle (\lambda x.\mu a.\langle x \mid C_0\, a\rangle) V \mid C_1\rangle & \equiv \langle V \mid (\lambda a.\mu x.\langle Ex \mid a\rangle) C\rangle \\
\rightsquigarrow_v \langle \mu a.\langle V \mid C_0\, a\rangle \mid C_1\rangle & \rightsquigarrow_v \langle V \mid \mu x.\langle Ex \mid C\rangle\rangle \\
\rightsquigarrow_v \langle V \mid C_0\, C_1\rangle & \rightsquigarrow_v \langle EV \mid C\rangle \ .
\end{array}$$

The third and fourth statements are derived from the first and second statements. □

The control operator `call/cc` can be defined as $\lambda x.\mu a.\langle x\ulcorner\lambda\_.a\urcorner \mid a\rangle$ in the call-by-value calculus, that is, $\langle \mathtt{call/cc}\, V \mid C\rangle \rightsquigarrow_v{}^+ \langle V\ulcorner\lambda\_.C\urcorner \mid C\rangle$ holds where $\_$ is a placeholder denoting a fresh continuation variable and $\rightsquigarrow_v{}^+$ is the transitive closure of $\rightsquigarrow_v$.

3 The variable level

For meta-variable studies, Sato et al. constructed a typed λ-calculus $\lambda\mathcal{M}$, which included dynamic binding of (expression) variables [20]. They used the level notion of variables and provided the definitions of α-equivalence, substitution, and reductions.

Levels are non-positive integers. Each level has infinitely many variables. Each variable x has a level i. We write $x^{\langle i\rangle}$ for it. We write level(E) for the maximum number of levels of variables that occur in E. The reduction strategy of $\lambda\mathcal{M}$ is that a λ-abstraction by a variable of the higher level is reduced. In this paper, we call it the higher-level reduction strategy. The reduction relation of $\lambda\mathcal{M}$ requires a function that the level of its body is not more than the level of its argument as follows:

$$(\lambda x^{\langle i\rangle}.E_0)E_1 \rightsquigarrow [E_1/x^{\langle i\rangle}]E_0 \qquad\qquad \text{if } \max\{\mathrm{level}(E_0), \mathrm{level}(E_1)\} \leq i.$$

It is noteworthy that the reduction strategy is non-deterministic.

The α-equivalence $E_0 \equiv_\alpha E_1$ in $\lambda\mathcal{M}$ is defined as $id \vdash E_0 \simeq E_1$ where $\simeq$ is defined by an inductive definition that includes

$$\frac{f \vdash E_0 \simeq E_1 \qquad f(x_0) = x_1 \qquad {x_1}^{\langle i\rangle} \text{ is fresh to } E_0 \qquad \max\{\mathrm{level}(E_0), \mathrm{level}(E_1)\} \leq i}{f \restriction (\mathrm{dom}\, f \setminus \{x_0\}) \vdash \lambda {x_0}^{\langle i\rangle}.E_0 \simeq \lambda {x_1}^{\langle i\rangle}.E_1}$$

id is the identity function, and f is an injective partial function. The α-equivalence relation distinguishes $\lambda {x_0}^{\langle 1\rangle}.({x_1}^{\langle 2\rangle}\lambda {x_2}^{\langle 1\rangle}.{x_0}^{\langle 1\rangle}{x_2}^{\langle 1\rangle})$ from $\lambda {x_3}^{\langle 1\rangle}.({x_1}^{\langle 2\rangle}\lambda {x_2}^{\langle 1\rangle}.{x_3}^{\langle 1\rangle}{x_2}^{\langle 1\rangle})$ because an expression containing ${x_0}^{\langle 1\rangle}$ may be substituted to ${x_1}^{\langle 2\rangle}$ on its reduction relation and the replacement of ${x_0}^{\langle 1\rangle}$ by ${x_3}^{\langle 1\rangle}$ may change its semantics. Intuitively, the α-equivalence relation 1) ignores variable names of the highest level, that is, the α-equivalence relation behaves as the ordinary α-equivalence relation for the highest-level variables, and 2) distinguishes variable names of the lower levels.

The substitution $[E_1/x^{\langle i\rangle}]E_0$ in $\lambda\mathcal{M}$ is defined only if $\max\{\mathrm{level}(E_0), \mathrm{level}(E_1)\} \leq i$ holds. The substitution is more restricted than usual because the α-equivalence relation is defined as described above. Because a variable crash in a capture-avoiding substitution needs another α-equivalent expression, the condition $\max\{\mathrm{level}(E_0), \mathrm{level}(E_1)\} \leq i$ is necessary. The condition is also sufficient because the $\lambda\mathcal{M}$ adopts the higher-level reduction strategy.

Thus, the α-equivalence relation of $\lambda\mathcal{M}$ works well for the higher-level reduction strategy. However, the α-equivalence relation cannot be adopted for the call-by-value calculus because redexes consisting of variables of lower levels can be reduced even though the higher-level variables remain. It also seems difficult to assign appropriate levels to variables so that the higher-level reduction strategy coincides with the call-by-value strategy. Therefore, we restrict the number of levels to 2. We also restrict the number of continuation variables of the low level to 1. The restrictions enable (1) and (2) in the call-by-value calculus straightforwardly.

4 Dynamic binding extension

We extend the call-by-value variant with dynamic binding of a continuation. We add a special continuation variable d of the type o distinguished from the other continuation variables. We think that d has the low level and the other variables have the high level, respectively.

Similarly to the type system in Figure 2, we define $\Xi \vdash^d_+ E\colon T$, $\Xi \vdash^d_- C\colon T$, and $\Xi \vdash^d_{\mathrm{o}} D$. It is noteworthy that we do not necessarily have to remove the variable x from $\operatorname{dom}\Xi$ in the $\to$-introduction rule for λ-abstractions for expressions. The other rules about λ-abstractions and μ-abstractions are similar; that ensures the derivability of α-equivalent syntactical objects such as $\lambda x_0.\lambda x_0.x_0$ and $\lambda x_0.\lambda x_1.x_1$. Furthermore, that also allows abstractions by d at multiple times.

We define the α-equivalence in a standard manner. To be precise, the α-equivalence considering the variable level is the standard α-equivalence under the assumption that the number of levels is 2, and the number of variables of the low level is 1.

We define substitutions $[E/x]^d B$, $[C/a]^d B$, and $[C/d]^d B$ as shown in Figures 4 and 5. We do not rename $[E/x]^d \mu d.D$, $[E/x]^d \lambda d.C$, $[C/a]^d \mu d.D$, and $[C/a]^d \lambda d.C$ because the ds in their substituted expressions and continuations are dynamically bound.

Proposition 4.1. *The definition of substitutions is well-defined; that is, $B_0 \equiv_\alpha B_1$ implies $[E/x]^d B_0 \equiv_\alpha [E/x]^d B_1$, $[C/a]^d B_0 \equiv_\alpha [C/a]^d B_1$, and $[C/d]^d B_0 \equiv_\alpha [C/d]^d B_1$.*

Proof. It is noteworthy that substitutions do not affect the scopes of μd and λd as seen in $[C/d]^d \mu d.D$ and $[C/d]^d \lambda d.C$. The other cases are trivial because the α-equivalence $\equiv_\alpha$ is defined in a standard manner, and we take fresh variables in cases of λ and μ-abstractions of the definition of substitution. □

We provide reduction rules as shown in Figure 6, which are unchanged from Figure 3, and substitution is slightly modified, as seen in Figures 4 and 5. It is noteworthy that d is also a continuation variable.

$$
\begin{array}{ll}
[E/x]^d x \equiv E & [E/x]^d a \equiv a \\
[E/x]^d x_0 \equiv x_0 \qquad \text{if } x_0 \not\equiv x & [E/x]^d d \equiv d \\
[E/x]^d \lambda x_0.E_0 \equiv \lambda x_1.[E/x]^d [x_1/x_0]^d E_0 & [E/x]^d \lambda a_0.C \equiv \lambda a_1.[E/x]^d [a_1/a_0]^d C \\
\qquad \text{if } x_1 \text{ is fresh} & \qquad \text{if } a_1 \text{ is fresh} \\
 & [E/x]^d \lambda d.C \equiv \lambda d.[E/x]^d C \\
[E/x]^d (E_0 E_1) \equiv [E/x]^d E_0 [E/x]^d E_1 & [E/x]^d (C_0 C_1) \equiv [E/x]^d C_0 [E/x]^d C_1 \\
[E/x]^d \mu a_0.D \equiv \mu a_1.[E/x]^d [a_1/a_0]^d D & [E/x]^d \mu x_0.D \equiv \mu x_1.[E/x]^d [x_1/x_0]^d D \\
\qquad \text{if } a_1 \text{ is fresh} & \qquad \text{if } x_1 \text{ is fresh} \\
[E/x]^d \mu d.D \equiv \mu d.[E/x]^d D &
\end{array}
$$

$$
[E/x]^d (\langle E_0 \mid C \rangle) \equiv \langle [E/x]^d E_0 \mid [E/x]^d C \rangle
$$

Figure 4: Substitution of expressions.

Proposition 4.2. *The reduction relation $\rightsquigarrow_v^d$ is deterministic.*

Proof. This is because each configuration can be applied to the unique reduction rule. □

We adopt the following uniform notation:

$$
\varXi \vdash^d B \colon T = \begin{cases} \varXi \vdash^d_+ E \colon T & \text{if } B \text{ is } E \\ \varXi \vdash^d_- C \colon T & \text{if } B \text{ is } C \\ \varXi \vdash^d_{\mathrm{o}} D & \text{if } B \text{ is } D \end{cases}
$$

for convenience of presentation. If B is a configuration, we ignore its type T.

Lemma 4.3. *The substitution lemma holds, that is,*

1. $\varXi \vdash^d_+ E \colon T'$ *and* $\varXi, x \colon T' \vdash^d B \colon T$ *imply* $\varXi \vdash^d [E/x]^d B \colon T$,
2. $\varXi \vdash^d_- C \colon T'$ *and* $\varXi, a \colon T' \vdash^d B \colon T$ *imply* $\varXi \vdash^d [C/a]^d B \colon T$, *and*
3. $\varXi \vdash^d_- C \colon o$ *and* $\varXi, d \colon o \vdash^d B \colon T$ *imply* $\varXi \vdash^d [C/d]^d B \colon T$.

Proof. We prove the lemma by induction on derivations of $\varXi, x \colon T' \vdash^d B \colon T$, $\varXi, a \colon T' \vdash^d B \colon T$, and $\varXi, d \colon T' \vdash^d B \colon T$. In the case that B is an abstraction by d, the d in the assumption was not removed when the typing rule was applied. The other cases are similar to those in the proof of Lemma 2.2. □

$$
\begin{array}{rlrl}
[C/a]^d x & \equiv x & [C/d]^d x & \equiv x \\
[C/a]^d \lambda x_0.E & \equiv \lambda x_1.[C/a]^d[x_1/x_0]^d E & [C/d]^d \lambda x_0.E & \equiv \lambda x_1.[C/d]^d[x_1/x_0]^d E \\
& \qquad \text{if } x \text{ is fresh} & & \qquad \text{if } x_1 \text{ is fresh} \\
[C/a]^d (E_0 E_1) & \equiv [C/a]^d E_0 [C/a]^d E_1 & [C/d]^d (E_0 E_1) & \equiv [C/d]^d E_0 [C/d]^d E_1 \\
[C/a]^d \mu a_0.D & \equiv \mu a_1.[C/a]^d[a_1/a_0]^d D & [C/d]^d \mu a_0.D & \equiv \mu a_1.[C/d]^d[a_1/a_0]^d D \\
& \qquad \text{if } a_1 \text{ is fresh} & & \qquad \text{if } a_1 \text{ is fresh} \\
[C/a]^d \mu d.D & \equiv \mu d.[C/a]^d D & [C/d]^d \mu d.D & \equiv \mu d.D \\
[C/a]^d a & \equiv C & [C/d]^d a & \equiv a \\
[C/a]^d a_0 & \equiv a_0 \qquad \text{if } a_0 \not\equiv a & & \\
[C/a]^d d & \equiv d & [C/d]^d d & \equiv C \\
[C/a]^d \lambda a_0.C_0 & \equiv \lambda a_1.[C/a]^d[a_1/a_0]^d C_0 & [C/d]^d \lambda a_0.C_0 & \equiv \lambda a_1.[C/d]^d[a_1/a_0]^d C_0 \\
& \qquad \text{if } a_1 \text{ is fresh} & & \qquad \text{if } a_1 \text{ is fresh} \\
[C/a]^d \lambda d.C_0 & \equiv \lambda d.[C/a]^d C_0 & [C/d]^d \lambda d.C_0 & \equiv \lambda d.C_0 \\
[C/a]^d (C_0 C_1) & \equiv [C/a]^d C_0 [C/a]^d C_1 & [C/d]^d (C_0 C_1) & \equiv [C/d]^d C_0 [C/d]^d C_1 \\
[C/a]^d \mu x_0.D & \equiv \mu x_1.[C/a]^d[x_1/x_0]^d D & [C/d]^d \mu x_0.D & \equiv \mu x_1.[C/d]^d[x_1/x_0]^d D \\
& \qquad \text{if } x_1 \text{ is fresh} & & \qquad \text{if } x_1 \text{ is fresh} \\
[C/a]^d (\langle E \mid C_0 \rangle) & \equiv \langle [C/a]^d E \mid [C/a]^d C_0 \rangle & [C/d]^d (\langle E \mid C_0 \rangle) & \equiv \langle [C/d]^d E \mid [C/d]^d C_0 \rangle
\end{array}
$$

Figure 5: Substitution of continuations.

Theorem 4.4. *The reduction relation enjoys the subject reduction property, that is,* $\Xi \vdash^d_{\mathrm{o}} D$ *and* $D \leadsto^d_v D'$ *imply* $\Xi \vdash^d_{\mathrm{o}} D'$.

Proof. We prove it by cases of the reduction relation. We use Lemma 4.3 in cases where reduction rules contain substitutions. The other cases are routines that just see types regarding the definition of typing rules. □

Theorem 4.5. *Assume* $\varnothing \vdash^d_{\mathrm{o}} D_0$. *Then, for any* D_1 *that occurs in its reduction sequence and is not a configuration consisting of a value and* $\bullet$, *there exists* D_2 *such that* $D_1 \leadsto^d_v D_2$.

Proof. It is by the same argument as for Theorem 2.4. □

We can easily confirm that the encodings described in Section 2 also work under $\leadsto^d_v$.

$$
\begin{array}{rll}
\langle(\lambda x.E)V \mid C\rangle \rightsquigarrow_v^d \langle[V/x]^d E \mid C\rangle & & \\
\langle V E \mid C\rangle \rightsquigarrow_v^d \langle E \mid \mu x.\langle V x \mid C\rangle\rangle & & \\
& & \text{if } E \text{ is not a value and } x \text{ is fresh} \\
\langle E_0 E_1 \mid C\rangle \rightsquigarrow_v^d \langle E_0 \mid \mu x.\langle x E_1 \mid C\rangle\rangle & & \\
& & \text{if } E_0 \text{ is not a value and } x \text{ is fresh} \\
\langle \mu a.D_0 \mid C\rangle \rightsquigarrow_v^d \langle \mu a.D_1 \mid C\rangle & & \text{if } D_0 \rightsquigarrow_v^d D_1 \\
\langle \mu a.D \mid C\rangle \rightsquigarrow_v^d [C/a]^d D & & \text{if } D \not\rightsquigarrow_v^d \\
\langle V \mid (\lambda a.C_0)\, C_1\, C_{2,0} \cdots C_{2,n-1}\rangle \rightsquigarrow_v^d \langle V \mid ([C_1/a]^d C_0)\, C_{2,0} \cdots C_{2,n-1}\rangle & & \\
\langle V \mid (\mu x.D)\, C_1\, C_{2,0} \cdots C_{2,n-1}\rangle \rightsquigarrow_v^d [\mu a.\langle V \mid a C_1\, C_{2,0} \cdots C_{2,n-1}\rangle/x]^d D & & \text{if } a \text{ is fresh} \\
\langle V \mid \mu x.D\rangle \rightsquigarrow_v^d [V/x]^d D & &
\end{array}
$$

Figure 6: A reduction relation $\rightsquigarrow_v^d$.

5 Definability of control operators

We call a control operator definable if there exists a syntactical object that causes the reduction in the context (soundness) and causes no other reduction in any context (completeness).

In order to show the definability of `prompt`/`control` [9], and `reset`/`shift` [6], we define a typed language with them, as shown in Figure 7. The delimiting operator `prompt` is also used for `reset`. The language is based on unilateralism.

The control operator `prompt` installs a delimiter distinguishing an inner continuation from the global continuation. If we reach a value inside the delimiter, then the delimiter disappears. The control operator `control` captures the delimited continuation by the innermost delimiter. The control operator `shift` also captures the delimited continuation by the innermost delimiter. The only difference is that a new delimiter is added when the delimited continuation is used.

A typical value that `control` receives is $\lambda x_0.\mathscr{E}_0[x_0 v]$. At this time, we expect that v is computed in the context $\mathscr{E}_0[\mathscr{E}]$, that is,

$$\mathtt{prompt}(\mathscr{E}[\mathtt{control}(\lambda x.\mathscr{E}_0[xv])]) \rightsquigarrow_v{}^+ \mathtt{prompt}(\mathscr{E}_0[\mathscr{E}[v]]) \ .$$

It is derived from the first, third, and fourth reduction rules in Figure 7. Similarly,

$$\mathtt{prompt}(\mathscr{E}[\mathtt{shift}(\lambda x.\mathscr{E}_0[xv])]) \rightsquigarrow_v{}^+ \mathtt{prompt}(\mathscr{E}_0[\mathtt{prompt}(\mathscr{E}[v])])$$

$$
\begin{array}{ll}
\text{(types)} & t ::= o \mid t \to t \\
\text{(expressions)} & e ::= c \mid x \mid \lambda x.e \mid ee \mid \texttt{prompt}(e) \mid \texttt{control}(v) \mid \texttt{shift}(v) \\
\text{(values)} & v ::= c \mid x \mid \lambda x.e \\
\text{(type environments)} & \Gamma ::= \varnothing \mid \Gamma, x\colon t \\
\text{(contexts)} & \mathscr{E} ::= [\cdot] \mid v\mathscr{E} \mid \mathscr{E}e
\end{array}
$$

$$
\frac{}{\Gamma \vdash c\colon o} \qquad \frac{}{\Gamma, x\colon t \vdash x\colon t} \qquad \frac{\Gamma, x\colon t_0 \vdash e\colon t_1}{\Gamma \vdash \lambda x.e\colon t_0 \to t_1} \qquad \frac{\Gamma \vdash e_0\colon t_0 \to t_1 \qquad \Gamma \vdash e_1\colon t_0}{\Gamma \vdash e_0 e_1\colon t_1}
$$

$$
\frac{\Gamma \vdash e\colon o}{\Gamma \vdash \texttt{prompt}(e)\colon o} \qquad \frac{\Gamma \vdash v\colon (o \to o) \to o}{\Gamma \vdash \texttt{control}(v)\colon o} \qquad \frac{\Gamma \vdash v\colon (o \to o) \to o}{\Gamma \vdash \texttt{shift}(v)\colon o}
$$

$$
\begin{array}{rcll}
\mathscr{E}[(\lambda x.e)v] & \rightsquigarrow_v & \mathscr{E}[[v/x]e] & \\
\texttt{prompt}(v) & \rightsquigarrow_v & v & \\
\texttt{prompt}(e_0) & \rightsquigarrow_v & \texttt{prompt}(e_1) & \text{if } e_0 \rightsquigarrow_v e_1 \\
\texttt{prompt}(\mathscr{E}[\texttt{control}(v)]) & \rightsquigarrow_v & \texttt{prompt}(v\lambda x.\mathscr{E}[x]) & \text{if } x \text{ is fresh} \\
\texttt{prompt}(\mathscr{E}[\texttt{shift}(v)]) & \rightsquigarrow_v & \texttt{prompt}(v\lambda x.\texttt{prompt}(\mathscr{E}[x])) & \text{if } x \text{ is fresh}
\end{array}
$$

Figure 7: A typed language with control operators for delimited continuations.

$$
\begin{array}{rcll}
\Phi(c) & \equiv & c & \qquad \Phi(x) \equiv x \\
\Phi(\lambda x.e) & \equiv & \lambda x.\Phi(e) & \qquad \Phi(e_0 e_1) \equiv \Phi(e_0)\Phi(e_1) \\
\Phi(\texttt{prompt}(e)) & \equiv & \mu d.\langle \Phi(e) \mid d \rangle & \\
\Phi(\texttt{control}(v)) & \equiv & \mu a.\langle \Phi(v) \ulcorner \lambda d.a \urcorner \mid d \rangle & \text{where } a \text{ is fresh} \\
\Phi(\texttt{shift}(v)) & \equiv & \mu a.\langle \Phi(v)(\lambda x.\mu d.\langle x \mid a \rangle) \mid d \rangle & \text{where } a \text{ and } x \text{ are fresh}
\end{array}
$$

Figure 8: Compilation of the language with control operators.

is derived from the first, third, and fifth reduction rules in Figure 7.

We provide a compilation into the extended calculus as shown in Figure 8. We also define $\Phi(\mathscr{E})$ consistently. We define contexts for continuations and a transformation from contexts for expressions to contexts for continuations as shown in

$$
\begin{aligned}
\mathscr{C} &::= [\cdot] \mid \mu x.\langle V x \mid \mathscr{C}\rangle \mid \mu x.\langle x E \mid \mathscr{C}\rangle \\
([\cdot])^{\Rightarrow} &\equiv [\cdot] \\
(V\Phi(\mathscr{E}))^{\Rightarrow} &\equiv (\Phi(\mathscr{E}))^{\Rightarrow}[\mu x.\langle V x \mid [\cdot]\rangle] && \text{where } x \text{ is fresh} \\
(\Phi(\mathscr{E})E)^{\Rightarrow} &\equiv (\Phi(\mathscr{E}))^{\Rightarrow}[\mu x.\langle x E \mid [\cdot]\rangle] && \text{where } x \text{ is fresh}
\end{aligned}
$$

Figure 9: A transformation from contexts for expressions to contexts for continuations

Figure 9. Intuitively, the transformation maps an expression until a control operator to a continuation pushed onto the stack. The compilation of $\texttt{prompt}(\mathscr{E}[e])$ by Φ deterministically reaches $\mu d.\langle e \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle$.

Proposition 5.1. $\langle \Phi(\texttt{prompt}(\mathscr{E}[e])) \mid C\rangle \rightsquigarrow_v^{d+} \langle \mu d.\langle e \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid C\rangle$.

Therefore, we can also regard Φ as including the optimization in its compilation.

The completeness of definability is derived from Proposition 4.2 and the definition of the compilation. In the following, we focus on soundness.

Theorem 5.2. *The following reductions hold:*

1. $\langle \Phi(\texttt{prompt}(v)) \mid C\rangle \rightsquigarrow_v^{d} \langle \Phi(v) \mid C\rangle$ *and*
2. $\Phi(e_0) \rightsquigarrow_v^{d} \Phi(e_1)$ *implies* $\langle \Phi(\texttt{prompt}(e_0)) \mid C\rangle \rightsquigarrow_v^{d} \langle \Phi(\texttt{prompt}(e_1)) \mid C\rangle$.

Proof. 1) It is derived from the following:

$$\langle \Phi(\texttt{prompt}(v)) \mid C\rangle \equiv \langle \mu d.\langle \Phi(v) \mid d\rangle \mid C\rangle \rightsquigarrow_v^{d} \langle \Phi(v) \mid C\rangle \ . \qquad (\because \Phi(v) \not\rightsquigarrow_v^{d})$$

2) By the definitions of Φ and $\rightsquigarrow_v^{d}$. □

Theorem 5.3. *The following reductions hold:*

1. $\langle \Phi(\texttt{prompt}(\mathscr{E}[\texttt{control}(v)])) \mid C\rangle \rightsquigarrow_v^{d+} \langle \mu d.\langle \Phi(v)\ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \mid d\rangle \mid C\rangle$ *and*
2. $\langle V \mid (\lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d])(\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle \rightsquigarrow_v^{d} \langle V \mid (\Phi(\mathscr{E}))^{\Rightarrow}[(\Phi(\mathscr{E}_0))^{\Rightarrow}[d]]\rangle$.

Proof. 1) The reductions are derived from the following:

$$
\begin{aligned}
\langle \Phi(\texttt{prompt}(\mathscr{E}[\texttt{control}(v)])) \mid C\rangle &\equiv \langle \mu d.\langle \Phi(\mathscr{E})[\Phi(\texttt{control}(v))] \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d+} \langle \mu d.\langle \mu a.\langle \Phi(v)\ulcorner \lambda d.a\urcorner \mid d\rangle \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(v)\ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \mid d\rangle \mid C\rangle \ .
\end{aligned}
$$

2) $(\Phi(\mathscr{E}))^{\Rightarrow}$ and $(\Phi(\mathscr{E}_0))^{\Rightarrow}$ have no free d. □

Corollary 5.4. $\texttt{prompt}(\mathscr{E}[\texttt{control}(\lambda x.\mathscr{E}_0[xv])]) \rightsquigarrow_v{}^+ \texttt{prompt}(\mathscr{E}_0[\mathscr{E}[v]])$ *is preserved by Φ in the sense of Proposition 5.1.*

Proof. The reductions are derived from the following:

$$
\begin{aligned}
&\langle \Phi(\texttt{prompt}(\mathscr{E}[\texttt{control}(\lambda x.\mathscr{E}_0[xv])])) \mid C\rangle \\
&\rightsquigarrow_v^{d+} \langle \mu d.\langle \Phi(\lambda x.\mathscr{E}_0[xv])\ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \mid d\rangle \mid C\rangle \\
&\equiv \langle \mu d.\langle (\lambda x.\Phi(\mathscr{E}_0)[x\Phi(v)])\ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle (\Phi(\mathscr{E}_0)[\ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \Phi(v)]) \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d+} \langle \mu d.\langle \ulcorner \lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d]\urcorner \Phi(v) \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(v) \mid (\lambda d.(\Phi(\mathscr{E}))^{\Rightarrow}[d])(\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(v) \mid (\Phi(\mathscr{E}))^{\Rightarrow}[(\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle \mid C\rangle \ .
\end{aligned}
$$

□

Theorem 5.5. *The following reductions hold:*

1. $\langle \Phi(\texttt{prompt}(\mathscr{E}_0[\texttt{shift}(v)])) \mid C\rangle \ \rightsquigarrow_v^{d+} \ \langle \mu d.\langle \Phi(v)\lambda x.\mu d.\langle x \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid d\rangle \mid C\rangle$ *and*
2. $\langle \mu d.\langle V \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle \rightsquigarrow_v^{d} (\Phi(\mathscr{E}))^{\Rightarrow}[(\Phi(\mathscr{E}_0))^{\Rightarrow}[d]]$.

Proof. 1) The reductions are derived from the following:

$$
\begin{aligned}
&\langle \Phi(\texttt{prompt}(\mathscr{E}_0[\texttt{shift}(v)])) \mid C\rangle \\
&\equiv \langle \mu d.\langle \Phi(\mathscr{E})[\Phi(\texttt{shift}(v))] \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d+} \langle \mu d.\langle \mu a.\langle \Phi(v)\lambda x.\mu d.\langle x \mid a\rangle \mid d\rangle \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(v)\lambda x.\mu d.\langle x \mid (\Phi(\mathscr{E}))^{\Rightarrow}[d]\rangle \mid d\rangle \mid C\rangle \ .
\end{aligned}
$$

2) $(\Phi(\mathscr{E}))^{\Rightarrow}$ and $(\Phi(\mathscr{E}_0))^{\Rightarrow}$ have no free d. □

Corollary 5.6. $\texttt{prompt}(\mathscr{E}[\texttt{shift}(\lambda x.\mathscr{E}_0[xv])]) \rightsquigarrow_v{}^+ \texttt{prompt}(\mathscr{E}_0[\texttt{prompt}(\mathscr{E}[v])])$ *is preserved by Φ in the sense of Proposition 5.1.*

Proof. The reductions are derived from the following:

$$
\begin{aligned}
&\langle \Phi(\texttt{prompt}(\mathscr{E}_0[\texttt{shift}(\lambda x.\mathscr{E}_1[xv])])) \mid C\rangle \\
&\rightsquigarrow_v^{d+} \langle \mu d.\langle \Phi(\lambda x.\mathscr{E}_1[xv])(\lambda x.\mu d.\langle x \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle) \mid d\rangle \mid C\rangle \\
&\equiv \langle \mu d.\langle (\lambda x.\Phi(\mathscr{E}_1)[x\Phi(v)])(\lambda x.\mu d.\langle x \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle) \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(\mathscr{E}_1)[(\lambda x.\mu d.\langle x \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle)\Phi(v)] \mid d\rangle \mid C\rangle \\
&\rightsquigarrow_v^{d} \langle \mu d.\langle \Phi(\mathscr{E}_1)[\mu d.\langle \Phi(v) \mid (\Phi(\mathscr{E}_0))^{\Rightarrow}[d]\rangle] \mid d\rangle \mid C\rangle \ .
\end{aligned}
$$

□

It is tedious to define a compilation that syntactically and exactly preserves the reduction relation because the language with control operators for delimited continuations is based on unilateralism, and continuations are not first-class objects in the language. We show that the fourth and fifth reduction rules in Figure 7 can be simulated by $\rightsquigarrow_v^d$ as seen in Theorems 5.3 and 5.5.

6 Related work

Our calculus is based on bilateralism, and we use the co-implication connective to denote types of functions on continuations. There is no directly related work. Sakaue and Asai, and Ueda and Asai used the negation connective for types of continuations [19, 22]. However, its occurrences are strongly restricted. Types of functions on continuations are of the form $\neg\, T \rightarrow \neg\, T'$. It means rejection that we adopted rather than implication between negations. We think they adopt the negation connective because they thought that the implication connective should conventionally define not only expression function types but also continuation function types.

The top-level variable has been used to extend calculi with delimited continuations. Ariola et al. investigated calculi corresponding to classical logic and added a single dynamically-scoped variable denoting the top-level continuation to construct $\lambda_{\mathcal{C}\widehat{\mathtt{tp}}}$ [3]. They followed Griffin's approach that adds control operators to the simply typed λ-calculus [12], whereas we followed Filinski's approach that constructs a symmetric λ-calculus [11]. Herbelin and Ghilezan redefined $\lambda_{\mathcal{C}\widehat{\mathtt{tp}}}$ to $\lambda\mu\widehat{\mathtt{tp}}$, which is based on Parigot's $\lambda\mu$ [15], in their work to study call-by-name delimited continuations [13]. Downen and Ariola constructed $\lambda\widehat{\mu}$ and $\lambda\widehat{\mu_0}$, which support multiple prompts [7].

In their approaches based on unilateralism, continuations are formally represented as contexts consisting of expressions. Each time a calculus is extended, its reduction relation is extended. For example, the $\lambda\mu\widehat{\mathtt{tp}}$-calculus, which is obtained by replacing the top-level constant of the $\lambda\mu$-calculus with the top-level variable, has additional reductions $\mu\widehat{\mathtt{tp}}.[\widehat{\mathtt{tp}}]V \rightsquigarrow V$. The extension is logically reasoned through a CPS transformation. In our approach based on bilateralism, the reduction rules are logically reasoned from the beginning. In the extended calculus, the reduction rules are unchanged and included by the normalization procedure in bilateral natural deduction, as described in Section 1.

It also seems unsatisfactory that they do not refer to the α-equivalence in their calculi even though studies on adding dynamically bound variables to calculi are also studies on how to define α-equivalence and substitution in calculi. We elaborated on the technical details to extend the call-by-value calculus with dynamic binding to support delimited continuations on bilateralism on which first-class continuations can directly express stacks. We also clarified that there is a technical difficulty in α-equivalence to support multiple prompts using the variable level. We have not found how to define appropriate α-equivalence and substitution in calculi like $\lambda\widehat{\mu}$, which supports multiple prompts.

In Curien and Herbelin's $\bar{\lambda}\mu\tilde{\mu}$ corresponding to sequent calculus, stacks can be expressed by first-class continuations more directly than calculi corresponding to unilateral natural deduction. Munch-Maccagnoni extended the $\bar{\lambda}\mu\tilde{\mu}$-calculus with delimited continuations using the top-level variable [14]. His calculus has commands c and a special variable $\widehat{\mathtt{tp}}$. He introduced pairs $c[\sigma]$ of a command c and a list of negative terms $\sigma = (t^0_\ominus, \ldots, t^{n-1}_\ominus)$ as follows:

$$\begin{aligned} c[\,] &\equiv c \\ c[t^0_\ominus, \ldots, t^{n-1}_\ominus] &\equiv \langle \mu\widehat{\mathtt{tp}}.c \parallel t^0_\ominus\rangle[t^1_\ominus, \ldots, t^{n-1}_\ominus] \ , \end{aligned}$$

by which the stack can be intuitively grasped. The reduction relation is also extended to include

$$\begin{aligned} \langle V_+ \parallel \widehat{\mathtt{tp}}\rangle[t_\ominus, \sigma]\rangle &\rightsquigarrow \langle V_+ \parallel t_\ominus\rangle[\sigma] \\ c[\sigma, \sigma'] &\rightsquigarrow c[\sigma, \widehat{\mathtt{tp}}, \sigma'] \\ c[\sigma, \mu q.(c'[t_\ominus]), \sigma'] &\rightsquigarrow c[\sigma, \mu q.c', t_\ominus, \sigma'] \end{aligned}$$

and the control operators `shift` and `reset` are definable in the calculus.

However, he just noted the variable's specialty: "*It can be bound in a special way by the binder $\mu\widehat{\mathtt{tp}}.c$. ... The operator $\mu\widehat{\mathtt{tp}}$ differs from a binder $\mu\alpha$ because $\widehat{\mathtt{tp}}$ is not a standard variable. Therefore it is not subject to standard renaming conventions nor subject to a capture-avoiding substitution* (p.164 in [14])" and did not describe the α-equivalence and substitution including the special variable clearly. We clarify a technical aspect of the construction of a calculus with such a variable by extending the call-by-value calculus to support delimited continuations on bilateralism on which first-class continuations can directly express stacks.

7 Conclusion and future work

We extended the call-by-value variant of λ_{conf} corresponding bilateral natural deduction with dynamic binding of a continuation variable in which the typical control operators for delimited continuations are definable. Stacks can be expressed by abstractions by continuation variables of the co-implication types, and the inspiration that dynamic binding defines delimited continuations naturally arises. We think that typed λ-calculi corresponding to bilateral natural deduction have the potentiality to be further studied and extended.

We used the extended calculus just as a target of the language with the control operators. Providing an abstract machine for the extended calculus based on bilateralism is significant as Herbelin and Ghilezan, and Downen and Ariola provided abstract machines for their calculi.

In order to define the substitutions for arbitrary syntactical objects on the call-by-value strategy, the number of continuation variables that are dynamically bound is limited to 1. This limitation prevents us from defining prompts with tags implemented in practical use (e.g., [10]). Relaxing the limitation is a theoretical challenge.

Acknowledgments. We thank Masatomo Hashimoto and Masahiko Sato, who answered several questions respectfully. We also thank the reviewer, who read the submitted manuscript carefully, suggested logical gaps, and provided many comments for improvement.

References

[1] Tatsuya Abe and Daisuke Kimura. A typed lambda-calculus with first-class configurations. *Journal of Logic and Computation*. `https://doi.org/10.1093/logcom/exac062`.

[2] Yuuki Andou. A normalization-procedure for the first order classical natural deduction with full logical symbols. *Tsukuba Journal of Mathematics*, 19(1):153–162, 1995.

[3] Zena M. Ariola, Hugo Herbelin, and Amr Sabry. A type-theoretic foundation of delimited continuation. *Higher-Order and Symbolic Computation*, 22(3):233–273, 2009.

[4] Nuel D. Belnap. Tonk, plonk and plink. *Analysis*, 22(6):130–134, 1962.

[5] Pierre-Louis Curien and Hugo Herbelin. The duality of computation. In *Proc. ICFP*, pages 233–243, 2000.

[6] Olivier Danvy and Andrzej Filinski. Abstracting control. In *Proc. LISP and Functional Programming*, pages 151–160, 1990.

[7] Paul Downen and Zena M. Ariola. Delimited control and computational effects. *Journal of Functional Programming*, 24(1):1–55, 2014.

[8] Michael Dummett. *The Logical Basis of Metaphysics*. Duckworth, 1991.

[9] Matthias Felleisen. The theory and practice of first-class prompts. In *Proc. POPL*, pages 180–190, 1988.

[10] Matthias Felleisen et al. Racket. `https://racket-lang.org/`, 1995.

[11] Andrzej Filinski. Declarative continuations: An investigation of duality in programming language semantics. In *Proc. CTCS*, volume 389 of *LNCS*, pages 224–249, 1989.

[12] Timothy G. Griffin. A formulae-as-types notion of control. In *Proc. POPL*, pages 47–58, 1990.

[13] Hugo Herbelin and Silvia Ghilezan. An approach to call-by-name delimited continuations. *ACM SIGPLAN Notices*, 43(1):383–394, 2008.

[14] Guillaume Munch-Maccagnoni. *Syntax and Models of a Non-Associative Composition of Programs and Proofs*. PhD thesis, Paris Diderot University, 2013.

[15] Michel Parigot. $\lambda\mu$-calculus: An algorithmic interpretation of classical natural deduction. In *Proc. LPAR*, volume 624 of *LNAI*, pages 190–201, 1992.

[16] Dag Prawitz. *Natural Deduction*. Almqvist & Wiksell, 1965.

[17] Arthur N. Prior. The runabout inference-ticket. *Analysis*, 21(2):38–39, 1960.

[18] Ian Rumfitt. “Yes” and “no”. *Mind*, 109(477):781–823, 2000.

[19] Sari Sakaue and Kenichi Asai. The foundations of symmetric lambda calculus (in Japanese). *JSSST Computer Software*, 26(2):3–17, 2009.

[20] Masahiko Sato, Takafumi Sakurai, Yukiyoshi Kameyama, and Atsushi Igarashi. Calculi of meta-variables. *Frontiers of Computer Science in China*, 2(1):12–21, 2008.

[21] Alex Shinn et al. *Revised^7 Report on the Algorithmic Language Scheme*, 2013.

[22] Yayoi Ueda and Kenichi Asai. Reinvestigation of symmetric lambda calculus. In *Proc. the 4th DIKU-IST Joint Workshop on Foundations of Software*, pages 10–26, 2011.

[23] Jan von Plato and Annika Siders. Normal derivability in classical natural deduction. *The Review of Symbolic Logic*, 5(2):205–211, 2012.

 Received 17 November 2022

A Special Type of Ideals in MV-algebras of Continuous Functions

Mahta Bedrood, Giuseppina Gerarda Barbieri, Giacomo Lenzi, Farhad Sajadian, Arsham Borumand Saeid

Bedrood.m@gmail.com, gibarbieri@unisa.it, gilenzi@unisa.it, fsajadian@bam.ac.ir, arsham@uk.ac.ir

Abstract

In this paper, Z-ideals and Z°-ideals in MV-algebras of continuous functions $C(X)$ are investigated. We introduce a special type of Z-ideals and Z°-ideals in MV-algebras of continuous functions $C(X)$, and call them Z_J-ideals and Z_J°-ideals. In particular, equivalent definitions have been provided for them and the relationship between them has been examined; their relationship with the minimal prime ideal has been investigated, too. The meet and the join of two Z_J-ideals and Z_J°-ideals are studied and we prove that the join of two Z_J-ideals is a Z_J-ideal whenever the continuous functions are defined on the real numbers $\mathbb{R}$ or the functions are defined on a compact metric space X. We also study when the annihilator of a subset of an MV-algebras of continuous functions is a Z_J-ideal or a Z_J°-ideal.

Keywords: Z-ideals/Z°-ideal, Z_J-ideals/Z_J°-ideals, MV-algebras, continuous functions
AMS Classification 2020: 06D35, 06B10, 54C05.

1 Introduction

MV-algebras were introduced by Chang in [8] and then intensively investigated and applied in many-valued logic. A trivial example of MV-algebra is a Boolean algebra; given an MV-algebra A and a set X, the set A^X of all functions $f : X \to A$ becomes an MV-algebra; in particular the set $C(X)$ of all continuous functions $f : X \to [0, 1]$ becomes an MV-algebra, that is studied in this paper. Y. Imai and K. Iseki introduced BCK-algebras as a common abstraction of the algebras corresponding to the implicative fragments of several logics existing in the literature as classical and intuitionistic logic. Then it was proved by Font, Rodriguez and Torrens, and independently, by Daniele Mundici, that bounded commutative BCK-algebras coincide with MV-algebras. Moreover, it is well-known that the classes of Abelian ℓ-groups and MV-algebras are deeply connected to each other. Mundici (see [10]) proved that there is a natural equivalence between the category of MV-algebras and the category of Abelian ℓ-groups with a strong unit. Now our setting is inspired by Di Nola-Sessa's paper (see [12]) where the authors begin to study MV-algebras of continuous functions. Indeed, we recall that a semisimple MV-algebra is (up to isomorphisms) a subalgebra of the MV-algebra of all $[0,1]$-valued continuous functions defined on a compact Hausdorff topological space. So our background is the MV-algebra of all continuous functions defined on a topological space X with values in the standard MV-algebra $[0,1]$.

Note that unlike Boolean algebras, MV-algebras cannot be recovered from their prime spectrum. However, the MV-algebras with a given prime spectrum are always a set, whereas MV-algebras with a given maximal spectrum are always a proper class, see [1].

Whereas spaces of prime ideals of Boolean algebras are known as Boolean spaces by Stone duality, spaces of prime ideals of MV-algebras are more difficult to understand. A characterization of these spaces is given in [11] in terms of the lattice of compact open sets of the space.

In this paper, we deal with Z-ideals/Z°-ideals, and Z_J-ideals/Z_J°-ideals in MV-algebras of continuous functions. Our aim is to find a characterization of ideals by means of Z_J-ideals or Z_J°-ideals. It is well-known the importance of ideals in algebra, and how the study of ideals in algebras could give information concerning the whole structure. Just as an example in Boolean algebras, since the seminal work by Stone [23], by force of his duality it is clear that: Open sets correspond to arbitrary ideals, clopen sets correspond to principal ideals, and in this sense the original Boolean algebra can be recovered from its space of prime ideals.

After giving notations and preliminaries, we offer some results about Z-ideals and Z°-ideals introduced in [6, 5]. Then we introduce Z_J-ideals and Z_J°-ideals and we give some properties; the zero sets, their relationship with Z-ideals and Z°-ideals, maximal ideals and minimal prime ideals are investigated. We emphasize that Z-ideals (Z°-ideal, respectively) are special kind of Z_J-ideals (Z_J°-ideals, respectively), exactly they are $Z_{C(X)}$-ideals ($Z_{C(X)}^{\circ}$-ideals, respectively). We prove that if I is an ideal of $C(X)$, then $M_f \cap I$ ($P_f \cap I$, respectively), for every $f \in I$, is a Z-ideal (Z°-ideal, respectively); if J is not a Z-ideal (Z°-ideal, respectively), then it contains a Z_J-ideal (Z_J°-ideal, respectively); if I is a subset of $C(X)$ and J is an ideal of $C(X)$ which contains $Ann(I)$, then $Ann(I)$ is a Z_J-ideal; and if each element of I is a boolean element, then $Ann(I)$ is a Z_J°-ideal. Moreover, we are able to prove that in MV-algebras of continuous functions defined on $\mathbb{R}$, the join of two Z_J-ideals is always a Z_J-ideal. For topological spaces different from $\mathbb{R}$, the problem is still open but showed not necessary join of two Z_J°-ideal is not a Z_J°-ideal . Nevertheless we give some partial results.

2 Preliminaries

We recollect some definitions and results which will be used in the sequel:

Definition 2.1. *[8] An MV-algebra is a structure (A, ⊕, *, 0) where ⊕ is a binary operation, * is a unary operation, and* 0 *is a constant such that the following axioms are satisfied for any* $x, y \in A$
(MV1) (A, ⊕, 0*) is an abelian monoid;*
(MV2) $(x^*)^* = x$*;*
(MV3) $0^* \oplus x = 0^*$*;*
(MV4) $(x^* \oplus y)^* \oplus y = (y^* \oplus x)^* \oplus x$.

From now on, A is an MV-algebra.
We put $1 := 0^*$, and the operation $\odot$ is defined as follows:

$$x \odot y = (x^* \oplus y^*)^*.$$

Moreover, we can define an order $\leq$ in such a way, for any two elements $x, y \in A$, $x \leq y$ if and only if $x^* \oplus y = 1$ if and only if $x \odot y^* = 0$. This order is called the natural order, and A becomes a bounded distributive lattice such that

$$x \vee y = x \oplus (x^* \odot y) = y \oplus (x \odot y^*) \quad \text{and} \quad x \wedge y = x \odot (x^* \oplus y) = y \odot (y^* \oplus x).$$

We will say that A is an MV-chain if it is linearly ordered.

Lemma 2.1. *[10] The following items hold for all* $x, y, z \in A$*:*
(1) If $x \leq y$*, then* $x \oplus z \leq y \oplus z$ *and* $x \odot z \leq y \odot z, x \wedge z \leqslant y \wedge z$,
(2) $x, y \leq x \oplus y$ *and* $x \odot y \leq x, y$, $x \leq nx = x \oplus x \oplus \cdots \oplus x$ *and* $x^n = x \odot x \odot \cdots \odot x \leq x$,
(3) If $x \leq y$ *and* $z \leq t$*, then* $x \oplus z \leq y \oplus t$,
(4) $x \wedge (y \oplus z) \leqslant (x \wedge y) \oplus (x \wedge z), x \wedge (x_1 \oplus ... \oplus x_n) \leqslant (x \wedge x_1) \oplus ... \oplus (x \wedge x_n)$*, for all* $x_1, ..., x_n \in A$*; in particular* $(mx) \wedge (ny) \leqslant mn(x \wedge y)$*, for every* $m, n \geq 0$.

Definition 2.2. *[22] Let* $e \in A$*. If there exists* $x \in A$ *such that* $e \vee x = 1$ *and* $e \wedge x = 0$*, then* e *is called a complemented element of* x.

We denote by $B(A)$ the set of all complemented elements of A.

Theorem 2.1. *[22] For every element e in A, the following conditions are equivalent:*
(1) $e \in B(A)$,
(2) $e \vee e^* = 1$,
(3) $e \wedge e^* = 0$,
(4) $e \oplus e = e$,
(5) $e \odot e = e$.

Definition 2.3. *[22] An ideal of A is a nonempty subset I of A satisfying the following conditions:*
(I1) If $x \in I,\ y \in A$ and $y \leq x$, then $y \in I$,
(I2) If $x, y \in I$, then $x \oplus y \in I$.

We denote by $Id(A)$ the set of all ideals of A: $Id(A)$ is a lattice with respect to the inclusion.

Remark 2.1. *[22]* (1) *Let $X \subseteq A$. Denote by $(X]$ the ideal generated by X. Then we have*
$(X] = \{a \in A \mid a \leqslant x_1 \oplus x_2 \oplus ... \oplus x_n,\ for\ some\ n \in \mathbb{N}\ and\ x_1, ..., x_n \in X\}$. *In particular,* $(a] = \{x \in A \mid x \leqslant na,\ for\ some\ n \in \mathbb{N}\}$.
(2) *For $I_1, I_2 \in Id(A), I_1 \wedge I_2 = I_1 \cap I_2$, $I_1 \vee I_2 = (I_1 \cup I_2] = \{a \in A : a \leq x \oplus y\ ;\ x \in I_1, y \in I_2\}$.*

Definition 2.4. *[22] Let I be an ideal of A. If $I \neq A$, then I is a proper ideal of A. A proper ideal I of A is called prime if for all $x, y \in A, x \wedge y \in I$, then $x \in I$ or $y \in I$.*

We denote by $Spec(A)$ the set of all prime ideals of an MV-algebra A.

Definition 2.5. *[22] An ideal I of A is called a minimal prime ideal of A:*
1) $I \in Spec(A)$;
2) If there exists $Q \in Spec(A)$ such that $Q \subseteq I$, then $I = Q$.

We denote by $Min(A)$ the set of all minimal prime ideals of A.

Remark 2.2. *[22] Minimal prime ideal P of A is called minimal prime ideal over ideal I, if*
1) $I \subseteq P$;
2) $If\ there\ exists\ Q \in Spec(A)\ such\ that\ I \subseteq Q \subseteq P,\ then\ P = Q.$

We denote by $Min(I)$ the set of all minimal prime ideals over ideal I.

Definition 2.6. *[22] A proper ideal I of A is called maximal if there exists no other proper ideal J of A so that $I \subseteq J$.*

We denote by $Max(A)$ the set of all maximal ideals of A.

Theorem 2.2. *[22] For a proper ideal $P \in Id(A)$ the following items are equivalent:*
(1) If $I \cap J \subseteq P$, then either $I \subseteq P$ or $J \subseteq P$, for all $I, J \in Id(A)$.
(2) $P \in Spec(A)$.
(3) If $x \wedge y \in P$, then either $x \in P$ or $y \in P$.
(4) If $x \wedge y = 0$, then either $x \in P$ or $y \in P$.

Definition 2.7. *[4] Let X be a nonempty subset of A. Then $Ann(X)$ is the annihilator of X defined by:*

$$Ann(X) = \{a \in A : a \wedge x = 0, \forall x \in X\}.$$

Theorem 2.3. *[3] Let $P \in Min(A)$ and I be finitely generated ideal. Then $I \subseteq P$ if and only if $Ann_A(I) \nsubseteq P$.*

Theorem 2.4. *[6] Let I be a proper ideal of A and $P \in Spec(A)$ such that $I \subseteq P$. Then there exists $P^* \in Min(I)$ such that $P^* \subseteq P$.*

Corollary 2.1. *[22] Every prime ideal of A is contained in a unique maximal ideal of A.*

3 Z-ideals and Z°-ideals in MV-algebras

In [6] and [5] the authors introduced and extensively studied Z-ideals and Z°-ideals in MV-algebras and in MV-algebras of continuous functions. Here we recall their definitions and give some other properties.
Let X be a completely regular topological space. We denote by $C(X)$ the set of all continuous functions on X to the standard MV-algebra $([0,1], \oplus, *)$. For every $f, g \in C(X)$ we define $(f \oplus g)(x) = f(x) \oplus g(x)$, $f^*(x) = (f(x))^*$ and $0(x) = 0$, for all $x \in X$. Then $(C(X), \oplus, *, 0)$ is an MV-algebra.

Notation 3.1. *Let $f \in C(X)$ and I be an ideal of $C(X)$, and $Y \subseteq P(X)$, put*

$$Z(f) = \{x \in X : f(x) = 0\}$$
$$Z(X) = \{Z(f) : f \in C(X)\}$$
$$Z(I) = \{Z(f) : f \in I\}$$
$$Z^{-1}(Y) = \{f \in C(X) : Z(f) \in Y\}.$$

Notation 3.2. *Let $a \in A$ and $I \in Id(A)$. Put*

$$M(a) = \{M : M \in Max(A), a \in M\} \quad M_a = \{b \in A | \forall M \in M(a).b \in M\}$$

$$P(a) = \{P : P \in Min(A), a \in P\} \quad P_a = \{b \in A | \forall P \in P(a).b \in P\}$$

$$M_I = \{b \in A | \forall M \in Max(A).I \subseteq M \to b \in M\}$$

$$P_I = \{b \in A | \forall P \in Min(A).I \subseteq P \to b \in P\}$$

Lemma 3.1. *Let $f \leq h$, then $M(h) \subseteq M(f)$ and $P(h) \subseteq P(f)$.*

Proof. Let $M \in M(h)$, then $h \in M$, since M is an ideal we get $f \in M$, therefore $M \in M(f)$. Likewise $P(h) \subseteq P(f)$.

□

Lemma 3.2. *Let f, $g \in C(X)$. Then $M_f \cap M_g = M_{f\wedge g}$ and $P_f \cap P_g = P_{f\wedge g}$.*

Proof. We prove the lemma for $P \in Min(C(X))$, the proof goes analogously for $M \in Max(C(X))$. (i) First observe that $P_f \cap P_g \supseteq P_{f\wedge g}$, since P_f, $P_g \supseteq P_{f\wedge g}$.

(ii) Let $h \in P_f \cap P_g$, we have to prove that $h \in P_{f\wedge g}$.

By way of contradiction, suppose that $h \notin P_{f\wedge g}$, then there exists an ideal $P \in Min(C(X))$ such that $f \wedge g \in P$ and $h \notin P$. By the definition of a prime ideal since $f \wedge g \in P$ we get either $f \in P$ or $g \in P$. In the first case we have $f \in P$ and $h \notin P_f$, in the second case we have $g \in P$ and $h \notin P_g$. In both cases we get a contradiction, so the proof is complete. □

Corollary 3.1. *Let I be a proper ideal of A. Then $P_I \subseteq M_I$.*

Proof. It is clear by Corollary 2.1 and Theorem 2.4. □

Lemma 3.3. *If I and J are ideals of $C(X)$ such that $I \subseteq J$ and $M_f \cap J \subseteq I$, for every $f \in I$, then $I = P_I \cap J$.*

Proof. It is clear that $I \subseteq P_I$. Then $I \subseteq P_I \cap J$.
Let $f \in I$. Then $M_I \subseteq M_f$, we get $M_I \cap J \subseteq M_f \cap J$, so $M_I \cap J \subseteq I$. By Corollary 3.1, we deduce $P_I \cap J \subseteq I$. Therefore $I = P_I \cap J$. □

Lemma 3.4. *[5] Let I be an ideal of $C(X)$. Then $Z(I)$ is closed under finite intersections, and $Z(I)$ is closed under extension, i.e., if $Z \in Z(I)$, $Z \subseteq Z'$, $Z' \in Z(X)$, then $Z' \in Z(I)$.*

Lemma 3.5. *[5] Let $f, g \in C(X)$. Then the following statements are equivalent:*
(1) $M_g \subseteq M_f$,
(2) $M(f) \subseteq M(g)$.

Lemma 3.6. *[5] If $f, g \in C(X)$ and $M(f) \subseteq M(g)$, then $Z(f) \subseteq Z(g)$.*

Definition 3.1. *[6]*

- *A proper ideal I of A is called a Z-ideal if $M_a \subseteq I$, for every $a \in I$.*
- *A proper ideal I of A is called a Z°-ideal if $P_a \subseteq I$, for each $a \in I$.*

Theorem 3.1. *[6] Let I be a Z-ideal (Z°-ideal, respectively) of A. Then every minimal prime ideal over I is a Z-ideal (Z°-ideal, respectively).*

Proposition 3.1. *[5] Let I be an ideal of $C(X)$. Then the following statements are equivalent:*
(1) *I is a Z-ideal,*
(2) *$Z(f) \subseteq Z(g)$ and $f \in I$ imply $g \in I$.*

Notation 3.3. *In the sequel we denote by $Z^\circ(f)$ the interior of $Z(f)$ where $f \in C(X)$.*

Proposition 3.2. *[5] Let I be an ideal of $C(X)$. Then the following statements are equivalent:*
(1) *I is a Z°-ideal,*
(2) *$Z^\circ(f) \subseteq Z^\circ(g)$ and $f \in I$ imply $g \in I$.*

Trivially, if X is endowed with the discrete topology, then every Z-ideal is a Z°-ideal and vice versa.

Theorem 3.2. *[5] If I and J are Z-ideals of $C(X)$, then $I \vee J$ is a Z-ideal of $C(X)$.*

Lemma 3.7. *[5] Let $f, g \in C(X)$. Then the following statements are equivalent:*
(1) *$P_g \subseteq P_f$,*
(2) *$P(f) \subseteq P(g)$,*
(3)*$Z^\circ(f) \subseteq Z^\circ(g)$,*
(4) *$Ann(f) \subseteq Ann(g)$.*

Lemma 3.8. *[5] Let $f_1, f_2 \in C(X)$. Then*
(1) $Z(f_1 \oplus f_2) = Z(f_1) \cap Z(f_2)$,
(2) $Z^\circ(f_1 \oplus f_2) = Z^\circ(f_1) \cap Z^\circ(f_2)$.

We recall that a space X is basically disconnected if every cozero-set has an open closure.

Theorem 3.3. *[5] If I and J are Z°-ideals of $C(X)$ and X is a basically disconnected space, then $I \vee J$ is either a Z°-ideal or $C(X)$.*

Theorem 3.4. *Suppose that I is an ideal of $C(X)$ and $I_z = \{g \in C(X) : g \in M_f$ for some $f \in I\}$. Then*
(1) $I_z = \bigvee_{f \in I} M_f$.
(2) I_z is a Z-ideal.
(3) $I \subseteq I_z$.
(4) Let $\{I_\alpha\}_{\alpha\in\Gamma}$ be a finite family of ideals. Then $\bigcap_{\alpha\in\Gamma}(I_\alpha)_z = (\bigcap_{\alpha\in\Gamma} I_\alpha)_z$
(5) If I is a Z-ideal then $I = I_z$.

Proof. (1) Let $g \in I_z$. Then there exists $f \in I$ such that $g \in M_f$. Hence $g \in \bigvee_{f\in I} M_f$. We deduce that $I_z \subseteq \bigvee_{f\in I} M_f$. Let $g \in \bigvee_{f\in I} M_f$. Then there exist $f_1, ..., f_n \in I$ such that $g \leq a_1 \oplus ... \oplus a_n$ and $a_i \in M_{f_i}$, for all $1 \leq i \leq n$. Put $f := f_1 \oplus ... \oplus f_n$. Thus $M_{f_i} \subseteq M_f$, for all $1 \leq i \leq n$, imply $a_i \in M_f$, for all $1 \leq i \leq n$. So $a_1 \oplus ... \oplus a_n \in M_f$ hence $g \in M_f$. Therefore $g \in I_z$ and the thesis comes.

(2) We have to prove that $M_h \subseteq I_z$ for every $h \in I_z$.

Let $h \in I_z$ and $g \in M_h$. So $h \in M_f$ for some $f \in I$, hence $M_h \subseteq M_f$ and, since $M_f \subseteq I_z$, we have $M_h \subseteq I_z$. Summing up, $g \in I_z$ and $M_h \subseteq I_z$. So the thesis comes.

(3) The proof is straightforward.

(4) We have to prove that
$\bigcap_{\alpha\in\Gamma} \{g \in C(X) : g \in M_f \ for\ some\ f \in I_\alpha\} = \bigcap_{\alpha\in\Gamma}(I_\alpha)_z = (\bigcap_{\alpha\in\Gamma} I_\alpha)_z = \{ g \in C(X) : g \in M_f \ for\ some\ f \in \bigcap I_\alpha\}$.

(i) First we prove $\bigcap_{\alpha\in\Gamma}(I_\alpha)_z \subseteq (\bigcap_{\alpha\in\Gamma} I_\alpha)_z$.

Let $h \in \bigcap_{\alpha\in\Gamma}(I_\alpha)_z$. Then $h \in (I_\alpha)_z$, for every $\alpha \in \Gamma$. Therefore, for every $\alpha \in \Gamma$, there exists $f_\alpha \in I_\alpha$ such that $h \in M_{f_\alpha}$. Therefore, for every $\alpha \in \Gamma$, there exists $f_\alpha \in I_\alpha$ such that $h \in M_{f_\alpha}$. Put $f := \bigwedge_{\alpha\in\Gamma} f_\alpha$. Then $\bigwedge_{\alpha\in\Gamma} f_\alpha \in \bigcap_{\alpha\in\Gamma}(I_\alpha)$. We claim that $h \in M_{\bigwedge_{\alpha\in\Gamma} f_\alpha}$. By way of contradiction suppose that $h \notin M_{\bigwedge_{\alpha\in\Gamma} f_\alpha}$. Then there exists $M \in Max(C(X))$ such that $h \notin P$ and $\bigwedge_{\alpha\in\Gamma} f_\alpha \in M$. Hence there exists $\alpha \in \Gamma$ such that $h \notin P$ and $f_\alpha \in P$, so $h \notin P_{f_\alpha}$, a contradiction. Therefore item (i) is proved.
(ii) Now we prove that $\bigcap_{\alpha\in\Gamma}(I_\alpha)_z \supseteq (\bigcap_{\alpha\in\Gamma} I_\alpha)_z$.

Let $h \in (\bigcap_{\alpha\in\Gamma} I_\alpha)_z$, then $h \in M_{\bar{f}}$ for some $\bar{f} \in \cap I_\alpha$.

Put $f_\alpha := \bar{f}$, then $h \in M_{f_\alpha}$ for $f_\alpha \in I_\alpha$, for every $\alpha \in \Gamma$.

Hence $h \in \bigcap_{\alpha\in\Gamma} \{g \in M_{f_\alpha}$ for some $f_\alpha \in I_\alpha\} = \bigcap_{\alpha\in\Gamma} (I_\alpha)_z$.

From (i) and (ii) we derive the proof.

(5) Suppose I is a Z-ideal. It is enough to show that $I_z \subseteq I$. Let $g \in I_z$. Then $g \in M_f$ for some $f \in I$. Since I is a Z-ideal, $M_f \subseteq I$. So $g \in I$. □

Remark 3.1. *The point (4) above is false in general if Γ is infinite.*

In fact, take $X = [0,1]$ and Γ be the set of all continuous functions $f : [0,1] \to [0,1]$ such that $Z(f) = \{0\}$. Since $[0,1]$ is compact, in $C([0,1])$ maximal ideals coincide with sets of functions which are zero in some fixed point $a \in [0,1]$. So, $g \in M_f$ is equivalent to say $Z(f) \subseteq Z(g)$.

Let $I_f = (f]$ and $g(x) = x$. For every $f \in \Gamma$, since f and g have the same zeros, we have $g \in M_f$, so $g \in (I_f)_z$.

Let now $h \in \bigcap_{f\in\Gamma} I_f$. Then $h(0) = 0$ since $h(x) \le nx$ for some n. Moreover $h = 0$ in a neighborhood of 0. In fact, suppose this is false. Then we would have a decreasing sequence $x_m \to 0$ with $h(x_m) > 0$ and $h(x_m) \to 0$. Take any function $k \in \Gamma$ such that $k(x_m) = h(x_m)$ for every m. We would have $k^2 \in \Gamma$ and $h \le nk^2$ for some n, so $h(x_m) \le nk(x_m)^2 = nh(x_m)^2$ for all m, which is impossible since $h(x_m)$ tends to 0 so $h(x_m) < 1/n$ for some m.

So $Z(h)$ is not contained in $Z(g)$ and $g \notin M_h$. Summing up, $g \notin (\bigcap_f I_f)_z$.

Remark 3.2. *By the theorem above, the map sending I to I_z is a closure operator in the sense of [15] (this remark is due to a referee).*

Remark 3.3. *Let I be an ideal of $C(X)$. By Lemma 3.5 we have*

$$\begin{aligned} I_z &= \{g \in C(X) : g \in M_f \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : M_g \subseteq M_f \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : M(f) \subseteq M(g) \ \textit{for some}\ f \in I\} \end{aligned}$$

Remark 3.4. *Suppose that I is an ideal of $C(X)$. Put $I_{z^\circ} = \{g \in C(X) : g \in P_f$ for some $f \in I\}$. By Lemma 3.7 we have*

$$\begin{aligned} I_{z^\circ} &= \{g \in C(X) : g \in P_f \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : P_g \subseteq P_f \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : P(f) \subseteq P(g) \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : Z^\circ(f) \subseteq Z^\circ(g) \ \textit{for some}\ f \in I\} \\ &= \{g \in C(X) : Ann(f) \subseteq Ann(g) \ \textit{for some}\ f \in I\}\,. \end{aligned}$$

Theorem 3.5. *Suppose that I is an ideal of $C(X)$, let $I_{z^\circ} := \{g \in C(X) : g \in P_f$ for some $f \in I\}$. Then*
(1) $I_{z^\circ} = \bigvee_{f\in I} P_f$.
(2) I_{z° is a Z°-ideal.
(3) $I \subseteq I_{z^\circ}$.
(4) $I_{z^\circ} = \cap\{Q : I \subseteq Q$ and Q is a Z°-ideal $\}$.
(5) If I is a Z°-ideal, then $I = I_{z^\circ}$
(6) If X is a basically disconnected, I and K are Z°-ideals such that $I \vee K \neq C(X)$, then $I_{z^\circ} \vee K_{z^\circ} = (I \vee K)_{z^\circ}$.
(7) Let $\{I_\alpha\}_{\alpha\in\Gamma}$ be a finite family of ideals. Then $\bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ} = (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ}$

Proof. (1) Let $g \in I_{z^\circ}$. Then there exists $f \in I$ such that $g \in P_f$. Hence $g \in \bigvee_{f\in I} P_f$. We deduce that $I_{z^\circ} \subseteq \bigvee_{f\in I} P_f$. Let $g \in \bigvee_{f\in I} P_f$. Then there exist $f_1, ..., f_n \in I$ such that $g \leq a_1 \oplus ... \oplus a_n$ and $a_i \in P_{f_i}$, for all $1 \leq i \leq n$. Put $f := f_1 \oplus ... \oplus f_n$. Thus $P_{f_i} \subseteq P_f$, for all $1 \leq i \leq n$, imply $a_i \in P_f$, for all $1 \leq i \leq n$. So $a_1 \oplus ... \oplus a_n \in P_f$ hence $g \in P_f$. Therefore $g \in I_{z^\circ}$ and the thesis comes.
(2) Let $Z^\circ(f_1) \subseteq Z^\circ(f_2)$ and $f_1 \in I_{z^\circ}$. Then there exists $f \in C(X)$ such that $f_1 \in P_f$. Therefore $P_{f_1} \subseteq P_f$. By Lemma 3.7 we get $P(f_1) \subseteq P(f_2)$ and $P_{f_2} \subseteq P_{f_1}$. Thus $P_{f_2} \subseteq P_f$, so $f_2 \in P_f$. Then $f_2 \in I_{z^\circ}$.
(3) The proof is straightforward.
(4) Put
$$H = \cap\{Q : I \subseteq Q \textit{ and } Q \textit{ is a } Z^\circ - \textit{ideal}\}.$$
By (2) and (3), we get $H \subseteq I_{z^\circ}$. We claim that $I_{z^\circ} \subseteq H$. By way of contradiction, let $h \in I_{z^\circ}$ but $h \notin H$. Then there exists K such that K is a Z°-ideal, $I \subseteq K$ and $h \notin K$. By Proposition 3.2, there does not exist any $g \in K$ such that $Z^\circ(g) \subseteq Z^\circ(h)$. It follows from Lemma 3.7 that there does not exist any $g \in K$ such that $P_h \subseteq P_g$ and there does not exist any $g \in K$ such that $h \in P_g$ so $h \notin I_{z^\circ}$ which is impossible.
(5) It follows from (3) and (4).
(6) By (5) and Theorem 3.3, we deduce that $I_{z^\circ} \vee K_{z^\circ} = I \vee K = (I \vee K)_{z^\circ}$.
(7) We have to prove that $\bigcap_{\alpha\in\Gamma} \{g \in C(X) : g \in P_f$ *for some* $f \in I_\alpha\} = \bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ} = (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ} = \{g \in C(X) : g \in P_f$ *for some* $f \in \cap I_\alpha\}$.

(i) First we prove that $\bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ} \subseteq (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ}$.

Let $h \in \bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ}$, we have to prove that $h \in (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ}$.

Since $h \in \bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ}$, then, for every $\alpha \in \Gamma$, we get $h \in (I_\alpha)_{z^\circ}$. Therefore, for every $\alpha \in \Gamma$, there exists $f_\alpha \in I_\alpha$ such that $h \in P_{f_\alpha}$. Put $f := \bigwedge_{\alpha\in\Gamma} f_\alpha$. Then $\bigwedge_{\alpha\in\Gamma} f_\alpha \in \bigcap_{\alpha\in\Gamma}(I_\alpha)$. We claim that $h \in P_{\bigwedge_{\alpha\in\Gamma} f_\alpha}$. By way of contradiction, suppose that $h \notin P_{\bigwedge_{\alpha\in\Gamma} f_\alpha}$. Then there exists $P \in Min(C(X))$ such that $h \notin P$ and $\bigwedge_{\alpha\in\Gamma} f_\alpha \in P$. Hence there exists $\alpha \in \Gamma$ such that $h \notin P$ and $f_\alpha \in P$, so $h \notin P_{f_\alpha}$, a contradiction. Therefore the proof of item (i) is complete.
(ii) Vice versa, we have to prove that $\bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ} \supseteq (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ}$.
Let $h \in (\bigcap_{\alpha\in\Gamma} I_\alpha)_{z^\circ}$ Then $h \in P_{\bar{f}}$ for some $\bar{f} \in \cap I_\alpha$, then $\bar{f} \in I_\alpha$, for every $\alpha \in \Gamma$. Put $f_\alpha := \bar{f}$. Then we get $h \in P_{f_\alpha}$, with $f_\alpha \in I_\alpha$, for every $\alpha \in \Gamma$.

Hence $h \in \bigcap_{\alpha\in\Gamma}\{g \in C(X) : g \in P_{f_\alpha} \text{ with } f_\alpha \in I_\alpha\} = \bigcap_{\alpha\in\Gamma}(I_\alpha)_{z^\circ}$. □

Remark 3.5. *The point (7) is false in general if Γ is infinite. Actually the counterexample is very similar to the one in Remark 3.1. In fact, again we use $X = [0,1]$ and Γ is the set of all continuous functions $f : [0,1] \to [0,1]$ such that $Z(f) = \{0\}$. Now, $g \in P_f$ is equivalent to say $Z^\circ(f) \subseteq Z^\circ(g)$.*

Let $I_f = (f]$ and $g(x) = x$. For every $f \in \Gamma$, since f and g have the same zeros, we have $g \in P_f$ so $g \in (I_f)_{z^\circ}$.

Let now $h \in \bigcap_{f\in\Gamma} I_f$. Then, as we know, $h = 0$ in a neighborhood of 0. So $Z^\circ(h)$ is not contained in $Z^\circ(g)$ and $g \notin P_h$. Summing up, $g \notin (\bigcap_f I_f)_{z^\circ}$.

Remark 3.6. *By the theorem above, the map sending I to I_{z° is a closure operator in the sense of [15] (this remark is due to a referee).*

4 Z_J-ideals of $C(X)$

In this section we introduce and study Z_J-ideals of $C(X)$.

Definition 4.1. *Let I, J be two ideals of $C(X)$. Then I is called a Z_J-ideal if*
(1) $I \subseteq J$
(2) $Z(f) \subseteq Z(g), f \in I$ *and* $g \in J$ *imply* $g \in I$.

Theorem 4.1. *Let I and J be two ideals of $C(X)$. Then the following items are equivalent:*

(1) I is a Z_J-ideal.
(2) (i) $I \subseteq J$
(ii)If $Z(f) = Z(g), f \in I$ and $g \in J$, then $g \in I$.
(3) (i) $I \subseteq J$
(ii)$M_f \cap J \subseteq I$, for every $f \in I$

Proof. (1) $\Rightarrow$ (2) The proof is straightforward.
(2) $\Rightarrow$ (1) Let $f \in I$, $k \in J$ and $Z(f) \subseteq Z(k)$, we have to prove that $k \in I$. Then $Z(f \oplus k) = Z(f)$ and $f \oplus k \in J$. Since $f \in I$, by (2) we get $(f \oplus k) \in I$. Since I is an ideal, also $k \in I$.
(1) $\Rightarrow$ (3) Let $t \in M_f \cap J$, for every $f \in I$. Then $t \in M_f$ and $t \in J$. So $M_t \subseteq M_f$, by Lemma 3.5 and Lemma 3.6, we have $Z(f) \subseteq Z(t)$. Therefore $t \in I$.
(3) $\Rightarrow$ (1) Let $Z(f) \subseteq Z(g), g \in J$ and $f \in I$. Then $Z^{\circ}(f) \subseteq Z^{\circ}(g)$. It follows from Lemma 3.7, that $Ann(f) \subseteq Ann(g)$. Since $f \in I$ by Lemma 3.3, we have $f \in P_I \cap J$, i.e, $f \in J$ and $f \in P$, for each $P \in Min(C(X))$ such that $I \subseteq P$. It follows from Theorem 2.3, that $Ann(f) \nsubseteq P$, thus $Ann(g) \nsubseteq P$. From Theorem 2.3 we derive $g \in P$. Hence $g \in P_I$, so $g \in P_I \cap J$. Therefore $g \in I$. □

Corollary 4.1. *Let I be an ideal of $C(X)$. Then for every $f \in C(X)$, $M_f \cap I$ is a Z_I-ideal.*

Theorem 4.2. *If I and J are two ideals of $C(X)$ and J is not a Z-ideal, then there exists an ideal H which is a Z_J-ideal and $H \subsetneq J$*

Proof. By hypothesis there exist $f, g \in C(X)$ such that $Z(f) \subseteq Z(g), f \in J$ but $g \notin J$. Put $H = M_g \cap J$. From Corollary 4.1 we derive that H is a Z_J-ideal. We will prove that H is a proper subset of J. By way of contradiction suppose that $H = J$. Then $M_g \cap J = J$, so $M_g \subseteq J$, we get $g \in J$, which is a contradiction. Therefore $H \subsetneq J$. □

From now on, if not otherwise specified, I, J are ideals of $C(X)$.

Remark 4.1. • *Every ideal I of $C(X)$ is a Z_I-ideal.*
• *The zero ideal is the only $Z_{\{0\}}$-ideal.*
• *Every Z-ideal is a $Z_{C(X)}$-ideal.*
• *An arbitrary intersection of Z_J-ideals is a Z_J-ideal.*

Example 4.1. (1) *Let X be any topological space and A, B be two subsets of X such that $B \subseteq A$. Put*

$$I = \{h \in C(X) : A \subseteq Z(h)\} \text{ and } J = \{h \in C(X) : B \subseteq Z(h)\}.$$

Obviously, I and J are ideals of $C(X)$ and I is a subset of J. We claim that I is a Z_J-ideal. By way of contradiction, suppose that there are $f, g \in C(X)$ such that $Z(f) \subseteq Z(g), f \in I, g \in J$ but $g \notin I$. Then $A \nsubseteq Z(g)$. On the other hand $f \in I$ so $A \subseteq Z(f)$ imply that $A \subseteq Z(g)$, a contradiction. Hence I is a Z_J-ideal.
(2) Let $X = [0,1], I = (f]$ with $f(x) = x$ and $J = (g]$ with $g(x) = \sqrt{x}$, for all $x \in X$. Obviously, $I \subseteq J$ and $Z(f) = Z(g) = \{0\}$. We claim that $g \notin I$. If $g \in I$, then there would exist $n \in \mathbb{N}$ such that $\sqrt{x} \leq nx$, for each $x \in [0,1]$. Hence $1 \leq n\sqrt{x}$, for each $x \in [0,1]$ which is impossible. Therefore I is not a Z_J-ideal.

Theorem 4.3. *Let I be a Z_J-ideal and $I^{'}$ be a $Z_{J^{'}}$-ideal of $C(X)$. Then $I \cap I^{'}$ is a $Z_{J \cap J^{'}}$-ideal of $C(X)$.*

Proof. By assumption I is Z_J-ideal, we get $I \subseteq J$ and, analogously, since $I^{'}$ is a $Z_{J^{'}}$-ideal, we get $I^{'} \subseteq J^{'}$, so $I \cap I^{'} \subseteq J \cap J^{'}$. Let $f, g \in C(X)$ such that $Z(f) \subseteq Z(g), f \in I \cap I^{'}$ and $g \in J \cap J^{'}$. We have to prove $g \in I \cap I^{'}$ Since I is Z_J-ideal, $f \in I$ and $g \in J$, we get $g \in I$. Since $I^{'}$ is a $Z_{J^{'}}$-ideal, $f \in I'$ and $g \in J'$, we get $g \in I'$. Therefore $g \in I \cap I^{'}$. □

Theorem 4.4. *Let I_z be as in Theorem 3.4. The following statements are equivalent:*
(1) I is a Z_J-ideal of $C(X)$,
(2) $I_z \cap J = I$,
(3) There exists a Z-ideal H in $C(X)$ such that $H \cap J = I$.

Proof. (1) $\Rightarrow$ (2) Obviously, $I \subseteq I_z \cap J$. Now let $g \in I_z \cap J$. Since $g \in I_z$, there exists $f \in I$ such that $g \in M_f$. So $M_g \subseteq M_f$. It follows, from Lemma 3.5, that $M(f) \subseteq M(g)$ and by Lemma 3.6, $Z(f) \subseteq Z(g)$. Since I is a Z_J-ideal we have $g \in I$.
(2) $\Rightarrow$ (3) It follows from Theorem 3.4(2).
(3) $\Rightarrow$ (1) By assumption $I \subseteq J$. Let $Z(f) \subseteq Z(g), f \in I$ and $g \in J$. We have to prove $g \in I$. Since $f \in H$ and H is a Z-ideal, we get $g \in H$, hence $g \in I$. □

Theorem 4.5. *Let I be a Z_J-ideal and I_z be as in Theorem 3.4.*
(1) If I is a maximal ideal, then either $I = J$ or $J = C(X)$.
(2) If I is a prime ideal, then either $I = J$ or $I = I_z$.
(3) If J is a Z-ideal, then I is a Z-ideal.
(4) If $J = C(X)$, then either $I = C(X)$ or $Z(f) \neq \emptyset$, for all $f \in I$.

Proof. (1) The proof is straightforward.

(2) By Theorem 4.4, we have $I_z \cap J = I$. It follows from Theorem 2.2 that either $I_z \subseteq I$ or $J \subseteq I$. By Theorem 3.4(3) we get either $I = I_z$ or $J = I$.

(3) Let $Z(f) \subseteq Z(g)$ such that $f \in I$ and $g \in C(X)$, we have to prove that $g \in I$.

Since I is Z_J-ideal we have $I \subseteq J$, so $f \in J$ and that implies $g \in J$, as J is a Z-ideal. As I is a Z_J-ideal we conclude that $g \in I$.

(4) Let $f \in I$ such that $Z(f) = \emptyset$. Then $Z(f) = Z(i)$ with $i(x) = 1$, for all $x \in X$. So $i \in I$ and that implies $I = C(X)$. $\square$

Proposition 4.1. *Let I be a Z_J-ideal, $P \in Min(I)$. Then P is either a Z-ideal or $J \subseteq P$.*

Proof. By Theorem 4.4, we get $I_z \cap J = I$. On the other hand $I \subseteq P$, so either $J \subseteq P$ or $I_z \subseteq P$. Suppose $J \nsubseteq P$. By Remark 2.2 we have $P \in Min(I_z)$. It follows from Theorem 3.4(2) and Theorem 3.1 that P is a Z-ideal. $\square$

Theorem 4.6. *Suppose that I is a subset of $C(X)$, J is an ideal of $C(X)$ such that $Ann_A(I) \subseteq J$. Then $Ann_A(I)$ is a Z_J-ideal.*

Proof. Let $Z(f) \subseteq Z(g), f \in Ann_A(I)$ and $g \in J$ but $g \notin Ann_A(I)$. Then there exists $k \in I$ such that $g \wedge k \neq 0$. Thus there exists $x \in X$, such that $(g \wedge k)(x) \neq 0$, so $g(x) \neq 0$ and $k(x) \neq 0$. We deduce $x \notin Z(g)$ and $x \notin Z(k)$, then $x \notin Z(f)$. So $(f \wedge k)(x) \neq 0$ implies that $f \notin Ann_A(I)$, which is a contradiction. Therefore $Ann_A(I)$ is a Z_J-ideal. $\square$

We conclude the section with a question. Thanks to the next sections we can answer YES whenever $X = \mathbb{R}$ or X is a compact metric space.

Question 4.1. *Let I and K be Z_J-ideals. Is $I \vee K$ a Z_J-ideal?*

4.1 Z_J-ideals of $C(X)$ where X is compact

If we suppose X compact, we have more additional results, we confine them in this section. A fundamental fact, in order to obtain these results, is that the inclusion of zero-sets of two functions f, g implies the inclusion $M(f) \subseteq M(g)$. In the compact case that holds true since we have the following representation of maximal ideals.

Theorem 4.7. *[10] If X is a compact space, then every maximal ideal M of $C(X)$ has the form M^x for a unique $x \in X$ where*

$$M^x = \{f \in C(X) : f(x) = 0\}.$$

Proposition 4.2. *If $f, g \in C(X)$ and $Z(f) \subseteq Z(g)$, then $M(f) \subseteq M(g)$.*

Proof. Let $M \in M(f)$. Then there exists $x \in X$ such that $M = M^x$ and $f \in M^x$. So $f(x) = 0$. By assumptions $x \in Z(f)$ implies $x \in Z(g)$. Thus $g(x) = 0$, we get $g \in M^x$. Therefore $M \in M(g)$. $\square$

Corollary 4.2. *Let $f, g \in C(X)$. Then the following statements are equivalent:*
(1) $M_g \subseteq M_f$,
(2) $M(f) \subseteq M(g)$,
(3)$Z(f) \subseteq Z(g)$.

Theorem 4.8. *Let I be an ideal of $C(X)$ and I_z be as in Theorem 3.4. Then*
(1) $I_z = \cap\{Q : I \subseteq Q$ and Q is a Z-ideal $\}$.
(2) If I is a Z-ideal, then $I = I_z$.
(3) If I and K are Z-ideals, then $I_z \vee K_z = (I \vee K)_z$.

Proof. (1) Put

$$H = \cap\{Q : I \subseteq Q \text{ and } Q \text{ is a } Z-\text{ideal}\}.$$

By Theorem 3.4(2) and (3), we get $H \subseteq I_z$. We claim that $I_z \subseteq H$. By way of contradiction, let $h \in I_z$ but $h \notin H$. Then there exists a Z-ideal K such that $I \subseteq K$ and $h \notin K$. Thus there does not exist any $g \in K$ such that $Z(g) \subseteq Z(h)$. It follows from Corollary 4.2 that there does not exist any $g \in K$ with $M_h \subseteq M_g$. Therefore we obtain $h \notin M_g$, so $h \notin I_z$, a contradiction.
(2) It follows from Theorem 3.4 and (1).
(3) By (2) and Theorem 3.2, we deduce that $I_z \vee K_z = I \vee K = (I \vee K)_z$. □

We conclude this section giving a partial answer to Question 4.1.

Proposition 4.3. *Let I and K be Z-ideals and Z_J-ideals. Then $I \vee K$ is a Z_J-ideal.*

Proof. Since I and K are Z_J-ideals, we have $I \subseteq J$ and $K \subseteq J$, hence $I \vee K \subseteq J$.
Thanks to Theorems 4.8 and 4.4(2) we have

$$I \vee K = (I \vee K) \cap J = (I_z \cap J) \vee (K_z \cap J) = (I_z \vee K_z) \cap J = (I \vee K)_z \cap J.$$

By Theorem 4.4(2) we can conclude that $I \vee K$ is a Z_J-ideal. □

5 The join of Z_J-ideals of $C(\mathbb{R})$: the case of principal ideals

In this section we will prove that whenever J is a principal ideal, the join of two Z_J-ideals is a Z_J-ideal.

Theorem 5.1. *Let $X = \mathbb{R}, I = (f], J = (g]$ and $f \leq g$. Then I is a Z_J-ideal if and only if there is $n \in \mathbb{N}$ such that for every x, $f(x) > 0$ implies $g(x) \leq nf(x)$.*

Proof. By way of contradiction for every n there is a point sequence x_n such that $g(x_n) > nf(x_n) > 0$. Up to passing to a subsequence we can suppose x_n converging or going to infinity.

So $f(x_n) < (1/n)g(x_n)$. Let k be a continuous function such that:

- $k(x_n) = \sqrt{n}f(x_n)$ for every n
- $k \leq g$

Now the ratio k/f is not bounded, because $k(x_n) = \sqrt{n}f(x_n)$. So we have $k \notin (f]$, so I is not a Z_J-ideal. Conversely, suppose $f(x) > 0$ implies $g(x) \leq nf(x)$. Let $f_0 \leq f$, $Z(f_0) = Z(k)$, $k \leq g$. We want to show $k \leq nf$ for some n. Let $x \in \mathbb{R}$. Likewise if $k(x) = 0$ then clearly $k(x) \leq f(x)$. Suppose $k(x) > 0$. Then $f_0(x) > 0$ and $f(x) > 0$. By hypothesis, $g(x) \leq nf(x)$, where n is independent of x. So $k(x) \leq nf(x)$, where n is independent of x. So $k \in I$ and this concludes the proof. □

Corollary 5.1. *Let $X = \mathbb{R}$, let $J = (g]$ for some $g \in C(X)$, and let $I = (f]$, $K = (h]$ be Z_J-ideals, for some f, $h \in C(X)$. Then $I \vee K$ is a Z_J-ideal.*

Proof. Suppose $I = (f]$, $K = (h]$ are Z_J-ideals. Let $(f \vee h)(x) > 0$. Then $f(x) > 0$ or $h(x) > 0$. If $f(x) > 0$ then $g(x) \leq nf(x) \leq n(f \vee h)(x)$ for some n. Likewise If $h(x) > 0$ then $g(x) \leq nh(x) \leq n(f \vee h)(x)$ for some n. By the previous theorem, $(f] \vee (h]$ is a Z_J-ideal. □

Before generalizing the previous results, we find it useful to introduce the following technical lemma. The function $nf(x)$ in Theorem 5.1 suggests us the existence of a function H which will play a crucial role in the "domination condition " contained in Theorem 5.2, this condition will be useful for proving that the join of two Z_J-ideals is a Z_J-ideal in $C(\mathbb{R})$.

Lemma 5.1. *For every continuous function $f : X \to [0, 1]$ there is a function $p_f : X \to [0, 1]$ (not continuous in general) such that:*

- *$f \leq p_f \leq 2f$ (hence $Z(f) = Z(p_f)$);*
- *p_f takes values in $\{0\} \cup \{1/2^n | n = 0, 1, 2, \ldots\}$*

Proof. We take $p_f(x) = 0$ if $f(x) = 0$, otherwise, $p_f(x)$ is equal to any $1/2^n$ such that

$$f(x) \leq 1/2^n \leq 2f(x)$$

(there are at least one and at most two of them). □

The next lemma can be seen as a characterization of the relation $Z(f) \subseteq Z(k)$, at least when f, k are functions on a compact metric space:

Lemma 5.2. *Let K be a compact metric space. Let f, $k \in C(K)$ with $Z(f) \subseteq Z(k)$. Then there is an increasing continuous function $H : [0,1] \to [0,1]$, with $H(0) = 0$, and $k(x) \leq H(f(x))$ for every $x \in K$.*

Proof. Let K be a compact metric space. Let f, $k \in C(K)$ with $Z(f) \subseteq Z(k)$. Let $p = p_f, q = p_k$ be as in Lemma 5.1. Let h_n be the sequence of the values of q. Then the sequence h_n tends to zero. In fact, otherwise there is a sequence $x_n \in K$ with $p(x_n)$ tending to zero and $q(x_n) \geq \epsilon > 0$. Since K is a compact metric space, x_n has a limit point x. So $f(x) = 0$ and $k(x) > 0$, contrary to the hypothesis $Z(f) \subseteq Z(k)$.

We can construct an increasing continuous function $H_0(y)$ such that we have $H_0(1/2^n) \geq h_m$ for every $m \leq n$. So $q(x) \leq H_0(p(x))$, for every $x \in K$. Since $p \leq 2f$ and $k \leq q$ we conclude $k(x) \leq q(x) \leq H_0(p(x)) \leq H_0(2f(x))$ since H_0 is increasing. Now it is enough to take $H(y) = H_0(2y)$. □

Now we will pass from compact sets to $\mathbb{R}$.

Corollary 5.2. *Let $f, k \in C(\mathbb{R})$ with $Z(f) \subseteq Z(k)$. Then there is a continuous function $H(x, y)$ increasing in y such that $H(x, 0) = 0$ and $k(x) \leq H(x, f(x))$ for every $x \in \mathbb{R}$.*

Proof. Let $f \in I$ and let $H_z(y)$ the function of the previous lemma defined on the compact $[z, z+1]$, where $z \in \mathbb{Z}$. Up to taking sums, we can suppose $H_n(y) = H_{-n}(y)$ and $H_{n+1}(y) - H_n(y)$ are continuous and increasing. For $x \in [n, n+1]$, $n \in \mathbb{N}$, set

$$H^+(x, y) := H_n(x, y) + (x - n)(H_{n+1}(x, y) - H_n(x, y)).$$

This function extends to a continuous function $H^+(x, y)$ for $x \in \mathbb{R}^+$. Likewise we define
$H^-(x, y) := H^+(-x, y)$ for $x \in \mathbb{R}^-$. Note that $H^+(0, y) = H^-(0, y)$, so H^+ and H^- extend to a continuous function $H(x, y)$ for every $x \in \mathbb{R}$. □

Theorem 5.2. *Let $J = (g]$ be a principal ideal. Then an ideal I of $C(\mathbb{R})$ is a Z_J-ideal if and only if $I \subseteq (g]$ and the following "domination condition" holds:*

if $f \in I$, $k \leq g$, and there is a continuous function $H(x, y)$ increasing in y, $H(x, 0) = 0$, $k(x) \leq H(x, f(x))$, then $k \in I$.

Proof. This is because the domination condition is equivalent to $Z(f) \subseteq Z(k)$. □

Corollary 5.3. *Let $J = (g]$ be a principal ideal. In $C(\mathbb{R})$ the join of two Z_J-ideals is a Z_J-ideal.*

Proof. Let I, K be two Z_J-ideals. We show that $I \vee K$ is a Z_J-ideal. Let $f \in I \vee K$. Then $f = f_I \vee f_K$ with $f_I \in I$ and $f_K \in K$. Let $t \leq g$ and

$$t(x) \leq H(x, f(x)) = H(x, f_I \vee f_K(x)) = H(x, f_I(x)) \vee H(x, f_K(x)).$$

Then $t(x) \wedge H(x, f_I(x)) \leq H(x, f_I(x))$ so $t(x) \wedge H(x, f_I(x)) \in I \subseteq I \vee K$. Likewise $t(x) \wedge H(x, f_K(x)) \in K \subseteq I \vee K$. Summing up, $t = (t \wedge H(x, f_I(x)) \vee (t \wedge H(x, f_K(x))) \in I \vee K$. By the previous theorem, $I \vee K$ is a Z_J-ideal. □

6 The join of Z_J-ideals in $C(\mathbb{R})$: from principal ideals to the general case

In this section we will prove that the join of two Z_J-ideals in $C(\mathbb{R})$ is always a Z_J-ideal.

Lemma 6.1. *I is a Z_J-ideal if and only if*
(i) $I \subseteq J$,
(ii) I is a $Z_{I\vee(g]}$-ideal for every $g \in J$.

Proof. $\Rightarrow$ Obviously, if I is a Z_J, then $I \subseteq J$ and I is a $Z_{I\vee(g]}$-ideal for every $g \in J$.
$\Leftarrow$ Suppose that I is a $Z_{I\vee(g]}$-ideal for every $g \in J$. We have to prove that I is a Z_J-ideal.
By assumptions we have
(i) $I \subseteq J$
(ii) $Z(f) \subseteq Z(k), f \in I$, $k \in I \vee (g]$, then $k \in I$
Let $Z(f) \subseteq Z(k), f \in I, k \in J$. We have to prove that $k \in I$.
But this comes from (ii). □

Lemma 6.2. *I is a $Z_{I\vee(g]}$-ideal for every $g \in C(X)$ if and only if for every $f, f_1 \in I$, $Z(f_1 \vee g_1) = Z(f)$, $g_1 \leq g$ imply $g_1 \in I$.*

Proof. $\Rightarrow$: Suppose that I is a $Z_{I\vee(g]}$-ideal, i.e.
(i) $I \subseteq I \vee (g]$
(ii) $Z(f) \subseteq Z(k)$, $f \in I$, $k \in I \vee (g]$ imply $k \in I$.
Now suppose $Z(f_1 \vee g_1) = Z(f)$, with $f, f_1 \in I$, then $f_1 \vee g_1 \in I \vee (g]$, and $g_1 \leq g$. We have to prove $g_1 \in I$. Since I is a $Z_{I\vee(g]}$-ideal we get $f_1 \vee g_1 \in I$, hence $g_1 \in I$.
$\Leftarrow$: Suppose that for every $f, f_1 \in I$, $Z(f_1 \vee g_1) = Z(f)$, $g_1 \leq g$ imply $g_1 \in I$. We have to prove that I is a $Z_{I\vee(g]}$-ideal.
Obviously, (i) $I \subseteq I \vee (g]$
(ii) Suppose $Z(f) = Z(k)$, $f \in I, k \in I \vee (g]$, we have to prove that $k \in I$.

By (ii) we get $Z(f \vee k) = Z(f)$. Since $k \in I \vee (g]$, we have $k \leq i \oplus ng$, for some $n \in \mathbb{N}$. Set

$$\tilde{g} := i \oplus ng$$

Then $I \vee (g] = I \vee (\tilde{g}]$. Now apply assumptions with $f = f$, $f_1 = f$, $g_1 = k$ and $\tilde{g} = g$, and we can conclude that $g_1 = k \in I$. □

Lemma 6.3. *I is a $Z_{I\vee(g]}$-ideal if and only if for every $f \in I$, $Z(g_1) \supseteq Z(f)$, $g_1 \leq g$ imply $g_1 \in I$.*

Proof. $\Rightarrow$: Just apply the definition of $Z_{I\vee(g]}$-ideal.
$\Leftarrow$: Apply Lemma 6.2 with $f_1 = f$: Indeed $Z(g_1 \vee f) = Z(g_1) \cap Z(f) = Z(f)$. □

Lemma 6.4. *I is a $Z_{I\vee(g]}$-ideal if and only if for every $f \in I$, $Z(g_1) = Z(f)$, $g_1 \leq g$ imply $g_1 \in I$.*

Proof. $\Rightarrow$ The proof comes Theorem 4.1.
$\Leftarrow$: Apply Lemma 6.2 with $f_1 = f$: Indeed $Z(g_1 \vee f) = Z(g_1) \cap Z(f) = Z(f)$. □

Lemma 6.5. *I is a $Z_{I\vee(g]}$-ideal if and only if the following condition holds: If $f \in I$, $k \leq g$ there is a continuous function $H(x, y)$ increasing in y and with $H(x, 0) = 0$, $k(x) \leq H(x, f(x))$, then $k \in I$.*

Proof. This is because the condition above is equivalent to $Z(f) \subseteq Z(k)$. □

Theorem 6.1. *If I, K are Z_J-ideals then $I \vee K$ is a Z_J-ideal.*

Proof. Let I, K be Z_J-ideals and $g \in J$. By Lemma 6.1 we obtain that I is $Z_{I\vee(g]}$-ideal and K is $Z_{K\vee(g]}$-ideal. It is sufficient to show that $I \vee K$ is $Z_{I\vee K\vee(g]}$-ideal, then apply Lemma 6.1 to complete the proof. Let $f \in I \vee K$. Then $f = f_I \vee f_K$ with $f_I \in I$ and $f_K \in K$. Let $t \leq g$ and $t(x) \leq H(x, f(x)) = H(x, f_I \vee f_K(x))$. Then $t(x) \wedge H(x, f_I(x)) \leq H(x, f_I(x))$ so $t(x) \wedge H(x, f_I(x)) \in I \subseteq I \vee K$. Likewise $t(x) \wedge H(x, f_K(x)) \in K \subseteq I \vee K$. Summing up, $t = (t \wedge H(x, f_I(x))) \vee (t \wedge (H(x, f_K(x))) \in I \vee K$. By the previous lemma, $I \vee K$ is a Z_J-ideal. □

6.1 Z_J°-ideals of $C(X)$

In this section we introduce and study Z_J°-ideals of $C(X)$.

Definition 6.1. *Let I, J be two ideals of $C(X)$. Then I is called a Z_J°-ideal if*
(1) $I \subseteq J$
(2) $Z^\circ(f) \subseteq Z^\circ(g), f \in I$ *and* $g \in J$ *imply* $g \in I$.

Theorem 6.2. *Let I and J be two ideals of $C(X)$. Then the following items are equivalent:*
(1) I is a Z°_J-ideal.
(2) (i) $I \subseteq J$
(ii)If $Z^\circ(f) = Z^\circ(g), f \in I$ and $g \in J$, then $g \in I$.
(3) (i) $I \subseteq J$
(ii)$P_f \cap J \subseteq I$, for every $f \in I$

Proof. $1 \Rightarrow 2$) The proof is straightforward.
$2 \Rightarrow 1$) Let $f \in I$, $f_1 \in J$ and $Z^\circ(f) \subseteq Z^\circ(f_1)$, we have to prove that $f_1 \in I$.

By Lemma 3.8, $Z^\circ(f \oplus f_1) = Z^\circ(f) \cap Z^\circ(f_1)$, thus $Z^\circ(f \oplus f_1) = Z^\circ(f)$ and $f \oplus f_1 \in J$. Since $f \in I$, by (2) we get $f \oplus f_1 \in I$. Since I is an ideal, also $f_1 \in I$.
$1 \Rightarrow 3$) Let $f, g \in C(X)$ such that $Z^\circ(f) \subseteq Z^\circ(g), f \in I$ and $g \in J$. It follows from Lemma 3.7, $P_g \subseteq P_f$ hence $P_g \cap J \subseteq P_f \cap J$. So $P_g \cap J \subseteq I$, then $g \in I$.
$3 \Rightarrow 1$) We have to prove that $P_f \cap J \subseteq I$, for every $f \in I$. Let $t \in P_f \cap J$. Then $t \in P_f$ and $t \in J$. So $P_t \subseteq P_f$, by Lemma 3.7, we have $Z^\circ(f) \subseteq Z^\circ(t)$. Thus $t \in I$. ☐

Corollary 6.1. *If I is an ideal of $C(X)$, then $P_f \cap I$, for every $f \in C(X)$, is a Z°_I-ideal.*

First we recall that an extremally disconnected topological space is a topological space in which the closure of every open set is open.

Remark 6.1. • *Every ideal I of $C(X)$ is a Z°_I-ideal.*
• *The zero ideal is the only $Z^\circ_{\{0\}}$-ideal.*
• *Every Z°-ideal is a $Z^\circ_{C(X)}$-ideal.*
• *Every Z°_J-ideal is a Z_J-ideal.*
• *If X is endowed with discrete topology or it is a P-space (see [13, on p.340] for definition) or it is an extremally disconnected topological space, then every Z°_J-ideal is a Z_J-ideal.*
• *An arbitrary intersection of Z°_J-ideals is a Z°_J-ideal.*

Easily from Example 4.1 we derive the following examples.

Example 6.1. *(1) Let $X = \mathbb{R}$ and A, B be two subsets of $\mathbb{R}$, such that $B \subseteq A$. Put*

$$I = \{h \in C(X) : A \subseteq Z^\circ(h)\} \text{ and } J = \{h \in C(X) : B \subseteq Z^\circ(h)\}.$$

Obviously, I and J are ideals of $C(X)$ and I is a subset of J. We claim that I is a Z°_J-ideal.
By way of contradiction suppose $f, g \in C(X)$ such that $Z^\circ(f) \subseteq Z^\circ(g), f \in I, g \in J$ but $g \notin I$. Then $A \nsubseteq Z^\circ(g)$. On the other hand $f \in I$ so $A \subseteq Z^\circ(f)$ imply that $A \subseteq Z^\circ(g)$, which is a contradiction. Hence I is a Z°_J-ideal.

(2) Let $X = [0,1], I = (f]$ where $f(x) = x$ and $J = (g]$ where $g(x) = \sqrt{x}$, for all $x \in X$. We have $I \subseteq J$ and $Z^{\circ}(f) = Z^{\circ}(g) = \emptyset$. We claim that $g \notin I$. If $g \in I$, then there exists $n \in \mathbb{N}$ such that $\sqrt{x} \leq nx$, for each $x \in [0,1]$. Hence $1 \leq n\sqrt{x}$, for each $x \in [0,1]$ which is impossible. Therefore I is not a Z°_J-ideal.

(3) Let $X = \mathbb{R}$,

$$I = \{h \in C(X) : [0,1] \cup \{2\} \subseteq Z(h)\} \text{ and } J = \{h \in C(X) : [0,1] \subseteq Z(h)\}.$$

By Example 4.1, I is a Z_J-ideal. Now, let $h,k \in C(X)$ such that $Z(h) = [0,1] \cup \{2\}$ and $Z(k) = [0,1]$. So $Z^{\circ}(h) = Z^{\circ}(k) = (0,1), h \in I, k \in J$ but $k \notin I$. We deduce that I is not a Z°_J-ideal.

(4) Let $X = \mathbb{R}$,

$$I = \{h \in C(X) : [0,\infty) \subseteq Z(h)\} \text{ and } J = \{h \in C(X) : (-\infty,0] \subseteq Z(h)\}.$$

Obviously, I and J are $Z^{\circ}_{C(X)}$-ideals. Put $f(x) = min(|x|,1)$ and $i(x) = 1$ for every $x \in \mathbb{R}$. Since $Z^{\circ}(f) = Z^{\circ}(i) = \emptyset, f \in I \vee J, i \in C(X)$ but $i \notin I \vee J$, then $I \vee J$ is not a $Z^{\circ}_{C(X)}$-ideal.

Theorem 6.3. *Let J be an ideal of $C(X)$ such that J is not a Z°-ideal. Then there exists an ideal H which is a Z°_J-ideal such that $H \subsetneq J$.*

Proof. By hypothesis there exist $f,g \in C(X)$ such that $Z^{\circ}(f) \subseteq Z^{\circ}(g), f \in J$ but $g \notin J$. Put $H = P_g \cap J$. From Corollary 6.1, we derive that H is a Z°_J-ideal. We claim that H is a proper subset of J. By way of contradiction suppose that $H = J$. Then $P_g \cap J = J$, so $P_g \subseteq J$, we get $g \in J$, which is a contradiction. Therefore $H \subsetneq J$. ☐

Theorem 6.4. *Let $I_{z^{\circ}}$ be as in Theorem 3.5. The following statements are equivalent:*
(1) I is a Z°_J-ideal of $C(X)$.
(2) $I_{z^{\circ}} \cap J = I$.
(3) There exists a Z°-ideal H in $C(X)$ such that $H \cap J = I$.

Proof. (1) ⇒ (2) Obviously, $I \subseteq I_{z^{\circ}} \cap J$. Now let $g \in I_{z^{\circ}} \cap J$. Since $g \in I_{z^{\circ}}$, there exists $f \in I$ such that $g \in P_f$. So $P_g \subseteq P_f$. It follows, from Lemma 3.7, $Z^{\circ}(f) \subseteq Z^{\circ}(g)$. Thus $g \in I$.
(2) ⇒ (3) It follows from Theorem 3.5(2).
(3) ⇒ (1) By assumptions $I \subseteq J$. Let $Z^{\circ}(f) \subseteq Z^{\circ}(g), f \in I$ and $g \in J$. We have to prove $g \in I$. Since $f \in H$ and H is a Z°-ideal, we get $g \in H$, hence $g \in I$. ☐

Theorem 6.5. *Let I be a Z°_J-ideal and let $I_{z^{\circ}}$ be as in Theorem 3.5. Then*
(1) If I is a maximal ideal, then either $I = J$ or $J = C(X)$.

(2) If I is a prime ideal, then either $I = J$ or $I = I_{z^\circ}$.
(3) If J is a Z°-ideal, then I is a Z°-ideal.
(4) If $J = C(X)$, then either $I = C(X)$ or $Z^\circ(f) \neq \emptyset$, for all $f \in I$.

Proof. (1) The proof is straightforward.
(2) By Theorem 6.4, we have $I_{z^\circ} \cap J = I$. It follows from Theorem 2.2 that either $I_{z^\circ} \subseteq I$ or $J \subseteq I$. By Theorem 3.5(3) we get either $I = I_{z^\circ}$ or $J = I$.
(3) Let $Z^\circ(f) \subseteq Z^\circ(g)$ such that $f \in I$ and $g \in C(X)$, we have to prove that $g \in I$. Since I is a Z°_J-ideal, we have $I \subseteq J$, hence $f \in J$, that implies $g \in J$, as J is a Z°-ideal. Thus $g \in I$.
(4) Let $f \in I$ such that $Z^\circ(f) = \emptyset$. Then $Z^\circ(f) = Z^\circ(i)$ where $i(x) = 1$, for all $x \in X$. So $i \in I$ implies $I = C(X)$. □

Theorem 6.6. *Let I be a subset of $C(X)$ such that $Z(h)$ be clopen, for every $h \in I$ and J be an ideal of $C(X)$ such that $Ann_A(I) \subseteq J$. Then $Ann_A(I)$ is a Z°_J-ideal.*

Proof. Let $Z^\circ(f) \subseteq Z^\circ(g), f \in Ann_A(I)$ and $g \in J$ but $g \notin Ann_A(I)$. Then there exists $k \in I$ such that $g \wedge k \neq 0$. Thus there exists $x \in X$, such that $(g \wedge k)(x) \neq 0$, so $g(x) \neq 0$ and $k(x) \neq 0$. We deduce $x \notin Z(g)$ and $x \notin Z(k)$, then $x \notin Z^\circ(g)$ implies that $x \notin Z^\circ(f) = Z(f)$. So $(f \wedge k)(x) \neq 0$, implies that $f \notin Ann_A(I)$, which is a contradiction. Therefore $Ann_A(I)$ is a Z°_J-ideal. □

Lemma 6.6. *[5] If $e \in B(C(X))$, then $Z(e)$ is an open subset of X.*

Corollary 6.2. *Let I be a subset of $C(X)$ such that $h \in B(C(X))$, for every $h \in I$ and J be an ideal of $C(X)$ such that $Ann_A(I) \subseteq J$. Then $Ann_A(I)$ is a Z°_J-ideal.*

Proof. It follows from Lemma 6.6 and Theorem 6.6. □

Proposition 6.1. *Let I be a Z°_J-ideal, $P \in Min(I)$. Then P is either a Z°-ideal or $J \subseteq P$.*

Proof. By Theorem 6.4, we get $I_{z^\circ} \cap J = I$. On the other hand $I \subseteq P$, so either $J \subseteq P$ or $I_{z^\circ} \subseteq P$. Suppose $J \nsubseteq P$. Obviously, by Remark 2.2, $P \in Min(I_{z^\circ})$. It follows from Theorem 3.5(2) and Theorem 3.1 that P is a Z°-ideal. □

7 Conclusions and future work

In this paper, Z_J-ideals and Z°_J-ideals in MV-algebras of continuous functions have been introduced: our scope is to characterize ideals of $C(X)$ by means of Z_J-ideals or Z°_J-ideals. We gave some properties of Z_J-ideals and Z°_J-ideals of $C(X)$; e.g., it is proved that if J is not a Z-ideal (Z°-ideal, respectively) then there exists an ideal H which is a Z_J-ideal (Z°_J-ideal, respectively) and $H \subsetneq J$. It is also proved that every element f of a $Z_{C(X)}$-ideal ($Z^{\circ}_{C(X)}$-ideal, respectively) has $Z(f) \neq \emptyset$ ($Z^{\circ}(f) \neq \emptyset$, respectively). Moreover, every minimal prime ideal over a Z_J-ideal (Z°_J-ideal, respectively) is either a Z-ideal (Z°-ideal, respectively) or includes J. We posed the question if the join of two Z_J-ideals is a Z_J-ideal and we solved this problem in the case $X = \mathbb{R}$ or X is a compact space, the general case is still open. Future work could be this line of investigation: the study of this special type of ideals in the MV-algebra of continuous functions $C(X)$ whenever X is an arbitrary topological space. This paper is just the beginning!

We have seen in the introduction that prime spectra of MV-algebras have been characterized in [11]. We intend to continue the study of spectra of MV-algebras, for instance, spectra of MV-algebras belonging to subclasses of MV-algebras. Several classes of MV-algebras can be characterized topologically, that is, by means of topological properties of the spectrum. For instance, local MV-algebras are the MV-algebras whose spectrum has a unique closed point. On the other hand this does not happen for every interesting class of MV-algebras, for instance the class of perfect MV-algebras is not topological. An investigation of topological properties is under investigation, see [2].

References

[1] Giuseppina Gerarda Barbieri, Antonio Di Nola, Giacomo Lenzi, Spectral MV-algebras and equispectrality, submitted.

[2] Giuseppina Gerarda Barbieri, Antonio Di Nola, Giacomo Lenzi, Topological classes of MV-algebras, in preparation.

[3] M. Bedrood, F. Forouzesh, F. Sajadian, *Some results in types of extensions of* $MV-algebras$, Publications de l'lnstitut Mathematique (2019), 161-177.

[4] M. Bedrood, F. Forouzesh, F. Sajadian, *Inverse topology in MV-algebras,* Mathematica Bohemica (2019), 273–285.

[5] M. Bedrood, A.Borumand Saeid, F. Sajadian *Z°-ideals in MV-algebras of continuous functions,* Filomat 36 7 (2022), 2439-2450.

[6] M. Bedrood, G. Lenzi, A.Borumand Saeid, F. Sajadian, *Z°-ideals in MV-algebras,* submitted.

[7] M.W. Bunder, *Ideals in positive implicative BCK-algebras*, Math. Japon. 26 no. 5 (1981), 589–593.

[8] C. C. Chang, *Algebraic analysis of many valued logic,* Trans. Amer. Math. Soc., **88** (1958), 467-490.

[9] C. C. Chang, *A new proof of the completeness of the Łukasiewicz axioms*,Trans. Amer. Math.Soc **93** (1959), 74-80.

[10] R. Cignoli, I. M. L. *D*'Ottaviano, D. Mundici, *Algebraic foundations of many- valued reasoning,* Kluwer Academic, Dordrecht, (2000).

[11] Di Nola, Antonio; Lenzi, Giacomo, The spectrum problem for Abelian ℓ-groups and MV-algebras. Algebra Universalis 81 (2020), no. 3, Paper No. 39, 43 pp.

[12] A. Di Nola, S. Sessa, *On MV-algebras of continuous functions,* in U. Höhle et al. (eds.), Non-Classical Logics and their Applications to Fuzzy Subsets (1995).

[13] L. Gillman, M. Henriksen, Concerning rings of continuous functions, Transactions of the AMS (1954), 340–362.

[14] L. Gillman, M. Jerison, Rings of continuous functions. The University Series in Higher Mathematics D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York (1960).

[15] O. A. Heubo-Kwegna, J. B. Nganou.: Closure Operations on MV-algebras, Fuzzy Sets Syst, 418, (2021) 139–152.

[16] C. S. Hoo, *MV -algebras, ideals and semisimplicity*, Math. Japon. 34, no. 4 (1989), 563–583.

[17] C. S. Hoo, *Implicative and Boolean ideals of MV-algebras*, Math. Japon. 39 (1994), 215–219.

[18] C. S. Hoo, *Fuzzy implicative and Boolean ideals of MV-algebras*, (English summary)Fuzzy Sets and Systems 66 no. 3 (1994), 315–327.

[19] K. Iseki, S. Tanaka, *Ideal theory of BCK-algebras*, Math. Japon. 21(1976), 351–366.

[20] V. Marra, L. Spada, *The dual adjunction between MV-algebras and Tychonoff spaces* Studia Logica 100 no. 1-2 (2012), 253–278.

[21] D. Mundici, *Interpretation of AF C*-algebras in Łukasiewicz sentential calculus* Journal of Functional Analysis 65, (1986), 15–63.

[22] D. Piciu, *Algebras of fuzzy logic,* Ed. Universitaria Craiova (2007).

[23] M. H. Stone, *The Theory of Representation for Boolean Algebras,* Trans. Amer. Math. Soc. 74 1(1936), 37–111.

[24] F. Wehrung, *Spectral spaces of countable Abelian lattice-ordered groups* Trans. Amer. Math. Soc. 371 3 (2019), 2133–2158.

Received 6 December 2022

Dripping the Poison: The Instruments of Bias
A Qualitative Case Study of News Articles in four Languages over nine Years

Sviatlana Höhn*
University of Luxembourg, DCS/SnT, Esch-sur-Alzette, Luxembourg
sviatlana.hoehn@uni.lu

Barbara Lewandowska-Tomaszczyk†
University of Applied Sciences in Konin, Poland

Sjouke Mauw
University of Luxembourg, DCS/SnT, Esch-sur-Alzette, Luxembourg

Dov Gabbay
King's College, London, UK and University of Luxembourg, Esch-sur-Alzette, Luxembourg

Abstract

Recent years have shown the power of dripping propaganda poison in minds of large populations with the goals to earn their support and to justify actions against particular groups or nations. These instruments of dripping the poison have been analysed in the literature on automated bias detection mainly from the lexical perspective. However, a more comprehensive perspective is needed for a more complete understanding and modelling of political biases present in this type of communication within such communities. To achieve our aim of a better perspective, we identified a case of such political propaganda that exerts influence and affects the general perception of one group by another one, which is general enough, well known enough but yet specific enough for our aim, namely Russia's portraying Ukrainians as right radicals. We analysed the

*Supported by FNR Luxembourg INTER-SLANT 13320890
†Supported by EU COST Action LITHME CA19102

instruments of bias in 68 articles in 4 languages reporting an event of fire in the House of Trade Unions in Odessa (Ukraine) on May 2, 2014. The analysis methods are inspired by the concepts of event structure, linguistic analysis and emotion studies. We identified three main dimensions of bias: logic, lexicon and emotionality, which are constitutive of the analysed emotionally, lexically and propositionally biased reports. Moreover, we extracted 28 types of instruments of bias which use the three main dimensions of bias to a different extent. Lexical choices are just one of them. This research puts forward a basic coherent classification of the instruments of bias for further computational and formal modelling.

1 Introduction

Biased political communication and propaganda campaigns shape our understanding of reality, and our understanding of what is right and what is wrong. Despite all research efforts, we still rely on intuitive understanding of what bias actually is. This paper deals with the instruments of creating bias. To identify and fight political and other social biases in media more effectively, we need better models of their effective instruments.

Non-neutral news articles can be symptoms of a bigger, strategic narrative embedded into an overall meta-interpretation - dubbed framing in media science - used to encourage one perspective on an event or its participants. A strategic narrative creates a link between intentions of an actor (e.g. state, power elites) and media content. In some cases, such strategic narratives own qualities of propaganda whose goals are managing collective attitudes and influencing a particular group of people [1]. To implement such strategic narratives, actors use various techniques, such as *nudging* (reshaping narratives and pushing them in the right direction), or the promotion of "alternative truth" which can be a lie, half-truth or truth out of context.

Reshaping narratives does not happen quickly. Usually, news outlets produce a number of slanted messages over time, adding and transforming the information step by step. Slanted messages usually present one group of individuals in a more positive light than another. Groups of individuals can include political parties, ethnic groups or otherwise classified parts of population. Over time, such slanted messages lead to consistent patterns that are called *bias* [2]. We use the metaphor of poison drip for this process: one drop of poison will not kill a person, however, with time, the saturation will increase and harm the physical condition of the person. The same happens with the slanted news: one slanted article that disfavours a group of individuals may not cause any negative effects for that group while, with time, a larger number of similarly slanted articles will reshape the perception of that group in the society, and can in turn, harm the group seriously [3, 4]. Therefore, critical news readers always need to ask themselves whose power is likely to be enhanced by this type of information.

1.1 Bias conceptualization and identification

Research on phenomena currently referred to as *bias* includes works on political manipulation, propaganda, misinformation and media discourse in a larger sense, see for example [5, 6, 7]. Despite critics such as [8, 9], computational approaches keep relying on intuitive understanding of bias without an operational definition of it [10],[11] or use the definition of **linguistic bias** provided by the Oxford Encyclopedia of Communication as "*a systematic asymmetry in word choice that reflects the social-category cognitions that are applied to the described group or individual(s)*" [12]. This systematic asymmetry can be intentional or unintentional, and takes place when we categorise some individuals. After the *category labels* are in place, we communicate stereotype-related information when we speak about the person, with the person, and also being that person (self-categorisation), as explained by [12]. Especially when we speak about the person, the category label determines what we communicate and how we communicate the category-related information. We also formulate stereotype-congruent information in a different way than stereotype-incongruent, and thus it is analysable [12].

The definition of linguistic bias by [12] offers a convenient argumentation for statistical approaches to bias detection that rely on word frequencies and distances between concepts. Such approaches use curated word lists and concept descriptions for personal attributes (most frequently gender) in order to measure distances between other concepts of interest (e.g. engineering vs. nursery) and the personal attribute of interest in a model that represents the semantics of a dataset [13, 14, 15, 16, 17].

Hamborg (2020) [9] criticises insufficient connection between computational approaches to bias research and social sciences. While social sciences do not employ state-of-the-art methods for automated text analysis, computational models of bias are too simplistic and their results do not provide additional insights [9, p.59].

The ways how the Natural Language Processing (NLP) community conducts bias research have their roots in the corpus-linguistic work on bias annotation. One of the earliest bias annotation schemes has been developed in [18] for Wikipedia. The authors collected articles that were flagged as non-neutral and tagged them at the article level, sentence level and word level by marking the neutrality on a five-point scale (0-neutral to 4 clearly non-neutral). In addition, for each annotated Wikipedia entry, the annotators marked the positions in the sentences that indicate non-neutral language with one of five tags: *polar_phrase*, *factive_phrase*, *weasel*, *repetition*, *personal_tone*. This approach to bias annotation suggests that some words are biased due to their semantics e.g., 'murderers'. Newer bias annotation works continue with this tradition [19]. However, for prototypically neutral phrases such as 'St. George's ribbons', bias can be located in the cultural context. The fact that meanings of words change with time has been recognised since the Antiquity and a large part of historical diachronic linguistics has been dealing with these phenomena. More recent research, for instance by [20] shows that the meaning of labels used for a party of a political conflict by the party itself, can be repurposed by another party in a negatively loaded, changed meaning. Prototypically neutral labels can then become negatively or positively marked over time solely by its logical use, without changing the linguistic context.

Research shows that people are more likely to believe something that is closer to their existing knowledge, beliefs and attitudes [21]. Several theoretical approaches make attempts to model human "common-sense" knowledge, including *scripts* by [22], *frames* by [23] and theory of *social practices* by [24]. These different theories have in common that they all describe some basic, recognisable, reproducible and socially acquired structure for sense-making while offering enough flexibility for changes. Social scripts are not universal; there are cultural differences in details [25]. Political bias in news articles can be seen as deliberate manipulations of these sense-making structures. Large-scale attempts to manipulate people's beliefs can be very effective if they take into account cultural, social and personal aspects of their recipients, as the history of Cambridge Analytica shows [26]. In particular, micro-targeting techniques still successfully use personal information about the users taken from social networks to influence their voting and consumption behaviour [27, 28, 29].

Framing in media is not the same as frames introduced by [23]. [2] defines *framing* as *the process of culling a few elements of perceived reality and assembling a narrative that highlights connections among them to promote a particular interpretation.* He uses the term *framing* to explain the basic relation between power, media and bias, and uses the term **content bias** for *"consistent patterns in the framing of mediated communication that promote the influence of one side in conflicts over the use of government power"*. According to [2], framing includes four steps: (1) agenda setting, (2) causal analysis, (3) moral judgement, and (4) remedy promotion. Unlike Minsky's frames introduced above, *framing* activates a particular structure (social script, social practice) *on purpose* in order to promote a very specific understanding or emotional reaction in the recipient. In this way, consistent patterns in the framing will be consistent patterns in activation of particular sense-making structures.

[30] developed a classification scheme with 15 dimensions to analyse framing bias in policy presentations. The dimensions include categories: *Economic*; *Capacity and resources*; *Morality*; *Fairness and equality*; *Legality*; *Policy prescription and evalua-*

tion; *Crime and punishment*; *Security and defense*; *Health and safety*; *Quality of life*; *Cultural identity*; *Public opinion*; *Political*; *External regulation and reputation*; and *Other*. These frame categories classify *how* a policy is presented to public focusing on the *topic* of a frame.

We will look at such manipulations by a bottom bottom-up, qualitative analysis (see Section 2.3 for a detailed explanation of the method) of text data. Looking ahead to the results of this analysis described in Section 3 in detail, we found that category label manipulation and argumentation fallacies are effective means of meaning manipulation, however, other means were also identified.

Category label manipulation is closer to the definition of linguistic bias by [12] and can be explained using the vocabulary of Membership Categorisation Analysis (MCA) [31, 32, 33]. Argumentation fallacies can be understood with the help of argument mining and categorisation of fallacies [34, 35].

MCA reveals how category labels activate our social memories, telling us what to expect from members of those categories. As [32] demonstrate, each social category is stored in our social category cognition as an aggregate of *features* that are tied to category labels at different degrees and include what Sacks called *category-bound activities* [36]:

- *Constitutive* features are type-embedded and criterial for that category, and must be observable;

- *Tied* features are criterial for that category under certain conditions, and will be definitely observable when those conditions occur; can generate category membership under those conditions;

- *Occasioned* features are not criterial but might be made criterial under certain conditions; not sufficient to infer about category membership.

These features are important because the observers (or readers of a news text) categorise people not only via explicit labelling but also via descriptions of their actions or attributes.

In relation to bias, MCA shows that social category label manipulation can be used to justify actions against a particular social category. In relation to the four steps of framing described by Entman [2], category labels and category attributes can be chosen deliberately so that a particular social category is made responsible for some problems chosen for the agenda and the remedy proposed will harm the members of that social category.

Example 1.1 demonstrates how changes in person-reference throughout a text cause changes in meaning: categorising women as engineering students, the speaker further classifies them as feminists (and then kills them).

Example 1.1. *"Women" transformed to "feminists" [37, p. 211]*

*"You're **women**. You're going to be **engineers**. You're all a **bunch of feminists**. I hate **feminists**."*

Example 1.1 is part of a larger MCA-driven study of events referred to as "the Montreal Massacre" in which a gunman shot 13 women to death because they were "all a bunch of feminists" [37].

Similar transformations are reported in literature, such as 'problem pupil' to 'shy boy' [38], 'offender' to 'murder suspect' [37] and 'police' to 'punishers' referring to SS nazi brigade that was acting with particular cruelty on the Belorussian territory during WW II [20]. Such replacements in category labels are also purposefully used in propaganda to justify actions against particular social groups and in support to groups in power positions. In our dataset, we observe lexical replacements such as *Ukrainian activists* to *fascists* which are used to justify the action against the Ukrainian nation.

In addition to the transformations of labels, different types of *fallacies* can be employed to convince or to persuade the reader. Academic literature distinguished between *formal fallacies* (errors in the argument form) [39] and *informal fallacies* (logically unsound arguments) [40]. Goffredo et al [35] annotated the following fallacies to support automated argument classification:

1. *ad hominem* personal attack against the opponent that can be of different types: *general ad hominem*, *tu quoque ad hominem*, *bias ad hominem* and *labeling*;

2. *appeal to emotion* includes appeal to pity, appeal to fear, emotionally loaded language and references to identity of the recipient;

3. *appeal to authority* including popular opinion;

4. *slippery slope* present a catastrophic situation;

5. *false cause* falsify the causal relationships of events;

6. *slogans* often repeated, a brief phrase used as a symbol of connection.

Walton [34] analyses fallacies in everyday conversations, listing six types of fallacies:

1. *ad populum* directed to a popular opinion or sentiment;

2. *ad misericordiam* (pity) relies on the recipient's empathy or sympathy.

3. *ad baculum* (stick) appeals to a threat, force or fear;

4. *ad hominem* is a personal attack against the opponent;

5. *ad ignorantum* is a claim that something is true because it has not been proved false or the other way round;

6. *ad verecundiam* (shame, modesty) - relies on respect for authority.

While appeal to emotion is seen as a separate type of fallacies in [35], Walton [34] argues that all his six types of (fallacious) arguments are appeals directed toward something in the mind of the recipient, such as their state of knowledge, their obligations and emotions. Walton also uses the term *emotional arguments* for his types of fallacies. He argues that arguments based on emotion can be good and reasonable if they contribute to the proper goals of the dialogue. In Aristotele's *Rhetoric*, such arguments were called *persuasive appeals* which reflects their purpose and function.

In relation to (social) scripts, frames and social practice theory, both label manipulation and fallacies can be useful techniques to modify scripts in a desired way, so that a peaceful protest is made a terrorist attack or a fascist mob, as we will observe in our dataset in Section 3.

1.2 Research Objective

The theories and empirical research discussed in Section 1.1 help to understand how biased information can be made believable and fit in the reader's existing knowledge and beliefs. While models of knowledge suggest that argumentation (including fallacies), social category labels and emotions are parts of one system, NLP-based bias detection research mostly analyses word choices and sentiments, for example [14, 41]. Only a few works acknowledge that lexical and argumentation techniques are used together to spread biased political information [42, 43]. However, the 22 categories used for annotation in [42, 43] were taken from literature. A data-driven study is needed to understand whether the 22 categories are sufficient to describe and annotate all possible instruments of bias.

The research objective of our article is to systematise and, if needed, to extend the *instruments of bias* in political news articles. Authors of all articles in our dataset make efforts to convince the reader that their content is true, although their stories about the same event are very different. The authors of this research do not know whether the texts are intentionally biased or not. However, we can analyse which linguistic and other means they use in their attempt to make the reader believe their reports. Our main research question is

"Which instruments of bias are employed in news articles?"

To answer this question, we analyse linguistic strategies used by authors of a collection of articles about the fire in the Trade Unions house in Odessa on May 2, 2014. We pay attention to lexical choices and fallacies already described in literature, and we are especially interested in finding other instruments of bias. Besides the asymmetries in word choice that only reflect bias in lexical choices, we identified propositional and emotional biases. Following the definition of [2], we searched for consistent patterns of bias with the future research objective to build a formal, data-driven model of bias in news articles that goes beyond lexicon-based and statistical approaches most frequent in computational bias modelling.

2 Data and Method

This section starts with a brief explanation of the background of the event chosen for our study. Further, the section describes the composition of the dataset and the analysis method. Our general approach was to analyse and compare reports about the same event from multiple information sources and in multiple languages distributed over multiple years. It allows for a study that is general enough but still specific enough with the aim of creation of a basic classification of the instruments of bias.

2.1 Event Background

On May 2, 2014 in Odessa (Ukraine), six people died in riots between Russia-oriented supporters of Ukrainian federation and West-oriented supporters of Ukrainian unity. Some of the participants entered the House of the Trade Unions that later caught fire. Dozens of people died in the House of the Trade Unions. There are numerous speculations about who set up the fire and why police and firefighters intervened with a delay of about 40 minutes. The events in Odessa have been used by media in different languages to picture Ukrainian people and authorities as nationalists and fascists. More details can be found in [44]. The narratives about this event have a strong relation to the geo-political developments in that region, in the decade following it.

2.2 Data

We collected a dataset of 68 online articles in four languages spoken by authors of this paper at native and near-native level (Russian, English, Polish and German) distributed over nine years (2014-2022)[1]. Russian-language sources include publi-

[1] Download the dataset from Zenodo: https://doi.org/10.5281/zenodo.7322863

cations from Russian, Belorussian and Ukrainian venues. Starting with the data collection in Russian because this was one of the languages spoken in the country where the events had happened, we wanted to see whether articles in other languages would contain some other types of biases than Russian. Therefore, we added also more western perspective by adding Polish, German and English.

The data were collected in May-June 2022 on Google search engine using key phrases in four languages. We started with neutral formulations such as *Odessa fire*, *Fire in Odessa House of Trade Unions*. An initial dataset was retrieved and scanned for additional keywords. New, also non-neutral key phrases such as *Odessa tragedy* and *Odessa massacre* were added to the search. This process was repeated several times until no new articles could be found. All articles were stored locally as PDF documents. A file with metadata accompanies the dataset and contains information about publication date, language and link to the source for each text by the time of the download.

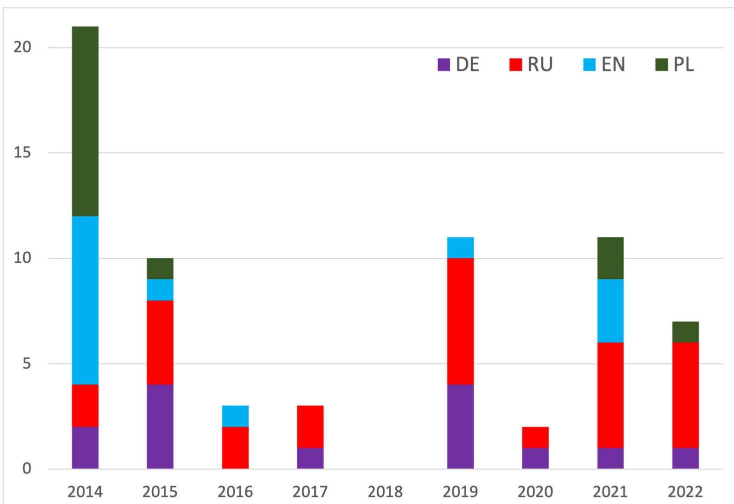


Figure 1: “Odessa Papers” by language and year

Figure 1 provides an overview of the number of articles per year per language. It is worth mentioning that the number of articles published on the chosen topic decreased between 2014 and 2017 going to zero in 2018, which would be a typical curve [44]. However, in 2019 and later the issue has been made relevant again by media, which raises questions, for which purpose it was made and who benefits from it.

While all articles in Russian were clearly written by native speakers, articles in English, Polish and German were sometimes translations from Russian, and the quality of translation ranged from very good to very poor, with lexical errors and ungrammatical constructions. The source languages are probably Russian and Ukrainian. Nevertheless, we have not excluded any texts from the analysis because all of them contributed to some extent to priming and opinion-making at the time of publication.

The texts in our collection represent the type of media text that attempt to persuade the reader. We refrain here from arguing whether the judgment against or inclination toward individuals or groups in our dataset is fair or unfair. A reliable description of an authentic course of events referred to in the analysed internet media texts is not available.

2.3 Method

The dataset is opportunistic: reports about the chosen events differ in quantity over years and language, and, as Section 3 shows, also differ in bias. To select the texts from our dataset for the qualitative study, we manually analysed the event structure in every text, as inspired by [45]. We included all texts that we could find with out search methodology with no exclusion criteria. The majority of the texts addressed the question "Who set up the fire?" at some point. We selected all descriptions in each article that address this question. Our starting hypothesis was that articles which blame Ukrainian pro-unity groups for causing that fire will be biased against them, and all other articles will not contain such biases.

Figure 2 shows the distribution of the answers: some texts blame "Ukrainian nationalists" for the crime (labelled as UKRAINIANS), some other texts claim that the events in Odessa were a "pro-Russian provocation" (labelled as PRO-RUSSIAN), some other texts state that it was an accident and both sides of the conflict contributed to the fire (labelled as ACCIDENT). There are also texts that chose to use grammatical form and vocabulary from which it is not clear who is to blame for the fire e.g. "the house started burning" (labelled as UNCLEAR) or simply do not address this issue (labelled as NO STANCE). While the UNCLEAR class clearly dominated directly after the events, the UKRAINIANS class became dominant in 2021-2022. Note that these numbers characterise the dataset without making conclusions about media in general.

Event texts labelled as ACCIDENT, UNCLEAR or NO STANCE often still exhibit a strong bias against one or the other party of the political conflict. Some of the texts do make attempts to present the information in an objective way. All of the texts make efforts to look credible. To understand the difference in bias formulations, we included two articles classified as UNCLEAR in our qualitative

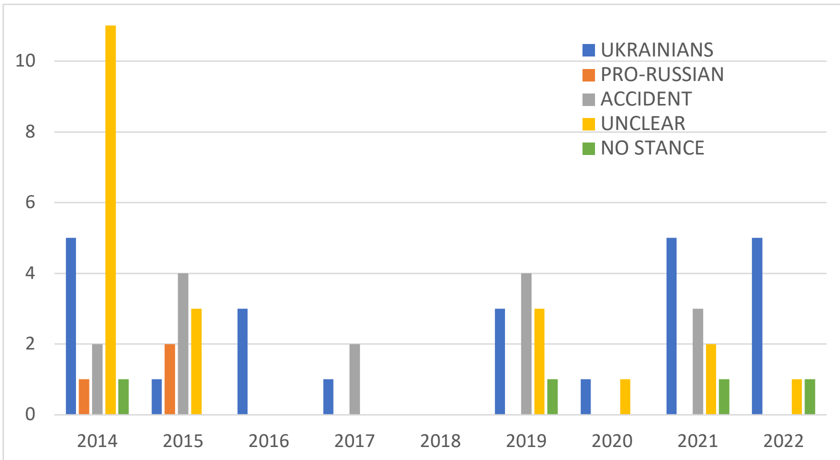


Figure 2: The distribution of classes over years as per answer to the question “Who set up the fire?” in each article.

case study.

	UKRAINIANS				PRO-RUSSIAN				ACCIDENT				UNCLEAR				NO STANCE			
Year	RU	DE	PL	EN	RU	DE	PL	EN	RU	DE	PL	EN	RU	DE	PL	EN	RU	DE	PL	EN
2014	2	1	1	1			1				1	1		1	6	5				1
2015	1				1		1		1	2		1	1	2						
2016	2			1																
2017	1								1	1										
2018																				
2019	3								1	2		1	1	2			1			
2020	1													1						
2021	2	1	1	1					2			1			1	1	1			
2022	3	1	1										1				1			

Figure 3: The detailed distribution of classes over years per language as per answer to the question “Who set up the fire?” in each article.

Figure 3 shows the distribution of classes by language and year. For a micro-analytic, exploratory case study we selected two texts per language labelled as UKRAINIANS, one per language from 2014, and one per language from 2021-2022 because other classes did not have enough articles in all languages. The texts were analysed by two experts in Linguistics experienced in qualitative research (two languages per expert) with the goal to identify all instruments of bias exploited in the texts. The results of the analysis of the experts were then compared.

The number of texts collected and selected for the analysis may look small for some of the readers. Indeed, for statistically-based research methods it would not deliver reliable results. For a qualitative study, the size of the dataset is large enough to draw reliable conclusions. In general, there is no way of predetermining the size of the dataset for qualitative research beforehand, and the size of the dataset always depends on the phenomena in focus and the depth of the analysis [46],

We chose qualitative methods inspired by Conversation Analysis and Membership Categorisation Analysis [33, 32] to analyse how bias works at the micro-analytic level. A similar approach has been successfully applied in [47] for a data-driven modelling of an artificial conversation companion that helps learners of German as a non-native language to practice conversation. The method consists of three steps that can be iteratively repeated and fine-tuned:

1. "Unmotivated looking", that is, careful reading and making notes about data without any preconception about what may be found. In this step, we marked and described all ways that make a (part) of the text look slanted towards one of the parties and against the other.

2. Building collections of similar examples. In this step we systematised all marked text parts with comments from Step 1 by their structural, semantic, syntactic, discourse, logical and/or pragmatic features.

3. Generalizations based on collections from Step 2. In this step we formulated high-level descriptions for the patterns that we identified in earlier steps.

The three steps above were effective in obtaining new insights from the data and helped us to answer the research question.

3 Data Analysis and Results

This section proposed a basic classification of the instruments of bias based on our dataset. Since our research question required exploration without any preconceptions of what might be found, we looked at biases in a very open way: where to find them, how they look like, what they do to meaning, how they work and so on. We found it meaningful to classify our findings as *levels*, *dimensions*, and *instruments*. Levels characterise the structural organisation of bias and their granularity. Dimensions describe qualitative composition of biases and their (supposed) effect on the recipient. Instruments describe the mechanics of bias more specifically.

3.1 Levels of Bias

We observe four granularity levels of bias in the dataset:

1. Bias on the syntax and lexical level (direct concept manipulation), for example, strongly suggestive rhetorical questions, use of additional attributes and descriptions for some categories vs. using just category labels for the other, extreme formulations and strongly negatively/positively marked category labels, attributes and action descriptions. This includes the sentence and concept level annotations as described in academic publications [48, 49].

2. Bias on the local discourse level (within one article) such as presence or absence of a lead, witness evidence, appeal to authority, citations and indirect speech, supporting visual material and picture captions. This includes article, paragraph, sentence and concept-level annotations mentioned in literature, but goes beyond that. For example, citations from witnesses were used to support a thesis put forward by the author, but also to discredit the party represented by the witness.

3. Bias on the global discourse level such as translations, re-publications and repeated occurrence of the same witnesses. This includes bias at source level mentioned in literature, however, we clearly see connections across sources and languages.

4. Bias strengthening over time which includes more and more extreme lexical choices for the propaganda sources, but more or less the same for more neutral venues; certainty in guilt and agency attribution; and objectivity pretence. This level of bias captures the temporal dynamics on the first three levels.

It is important to note that item 4. shows the way how bias becomes an iterative and dynamic process of "dripping the poison" over time. This gradual effect on the mind evolving over time is probably the most poisonous aspect of bias. We were able to extract evidence for bias strengthening over time because the texts in our dataset were published over nine years, and our in-depth analysis revealed changes, for example, in certainty in guilt attribution, more extreme formulations over time, and changes in agency attribution. We can use these insights for a future quantitative study that would measure these phenomena statistically.

3.2 Dimensions of Bias

At the same time, we observe three major dimensions of bias in each text:

1. **Emotions**: this dimension works with labels and visual material that aim at eliciting a particular emotional response from the reader;

2. **Language**: this dimension uses lexical, syntactic and discourse-structure to elicit a particular interpretation of the text.

3. **Logic**: this dimension uses means of argumentation and reasoning to persuade the reader and to promote a particular understanding.

All three are connected with each other. Emotions are expressed in texts via language, including lexical choices, e.g. calling names (murderes, Nazis, etc). Language is used to express emotions and propositional content in texts. Logical argumentation is used to present the validity of the arguments used to justify the reasoning behind particular propositional content.

3.3 Instruments of Bias

From the collections of similar examples, we extracted 28 instruments of bias (Steps 2 and 3 of the method explained in Section 2.3). We analysed and categorised patterns that we found in the texts, and found that some of the patterns are already described in literature as *fallacies*. After the generalisation step we found that the 28 instruments of bias are built on all three dimensions, and the proportion of each dimension in each instrument can vary.

Some of the instruments are particularly well-suited for an example-based discussion because they can be identified at the level of concepts, sentences and paragraphs. These instruments include evaluative lexis, use of metonymy, metaphor and simile, agency, appeal to authority, question answering avoidance, contrast-induced categories and certainty. Some other instrument types become visible only at the higher structural levels, for instance contrast-based identity construction, contrast-based personification, reasoning by stereotypes and objectivity pretence are frequently but not always the result of biases on the local discourse level. Label transformation, contradictions, event structure manipulation and temporal/sequence aspects of certainty typically occur at the local and at the global discourse levels.

Figure 4 places the instruments of bias on a 2D space and shows one possible clustering of the instruments based on their components. Because emotions are part of the resources that can be used for creation of the instruments, the proportion of both emotions and logic in each of the instruments can vary. The occurrences of the instruments at each structural level differs in different texts. In practice, all instruments can occur at all structural levels.

We can categorise the instruments of bias in three major parts:

1. *Fallacies*: are marked red in Figure 4 and are build upon appeals to emotion, attempts to convince the reader by elicitation of particular emotions such as fear, pity and anger. We found eleven fallacies and present them in Table 1.

2. *Categorisation work*: marked blue in Figure 4 and rely on the explicit use of negative and positive connotation for descriptions of the groups and persons. We found five instruments of bias that exploit mostly lexical choices and present them in Table 2.

3. *Others*. These bias instruments are marked green in Figure 4 and explained in Table 3.

We summarise appeals to emotions (various attempts to elicit emotions in readers that are supposed to lead the reader in a particular direction, believe a particular message, trust a particular position etc.) under the general concept of *fallacies* because they match the descriptions of fallacies in literature [35]. It is still to clarify whether only particular types of fallacies are used as instruments of bias, or all of them have this potential.

The instruments categorised as *other* in Table 3 can be grouped by their semantics, such as *contrast-based* instruments. Although we found only listed contrast-based instruments of bias, other datasets can contain other specific types of such instruments of bias. An important finding of this study is that contrast in creation of social categories in news offer effective tools for making the readers believe biased information. Such contrasts (or asymmetries) can appear via a variety of bias instruments, including those that are clearly contrast-based (contrast in personification, contrast in space allocation, contrast-based identity construction and contrast-induced categories). This implies that existing methods of bias annotation that label some parts of text as biased will have difficulties in capturing most of contrast-based biases.

Several instruments are quite complex and can be broken down to more specific descriptions. For the reason of their structural significance and the role they play in the (desired) reading of the texts by potential audience, we explain here two complex instruments more precisely: *emotional persuasion* and *imposed framing*.

Emotional persuasion is an umbrella instrument that involves creation of opposite identities (for example 'us' and 'them', 'good ones' and 'bad ones') by verbal and non-verbal descriptions of their members (attributes and actions) that target opposite emotional reactions in readers, such as empathy and hate (although calling them in general *positive* and *negative* would be simplified). For example, describing members of groups as parents (e.g. father who went to a demonstration with his 17 year old son, a young man who died and left his pregnant wife) persuade the

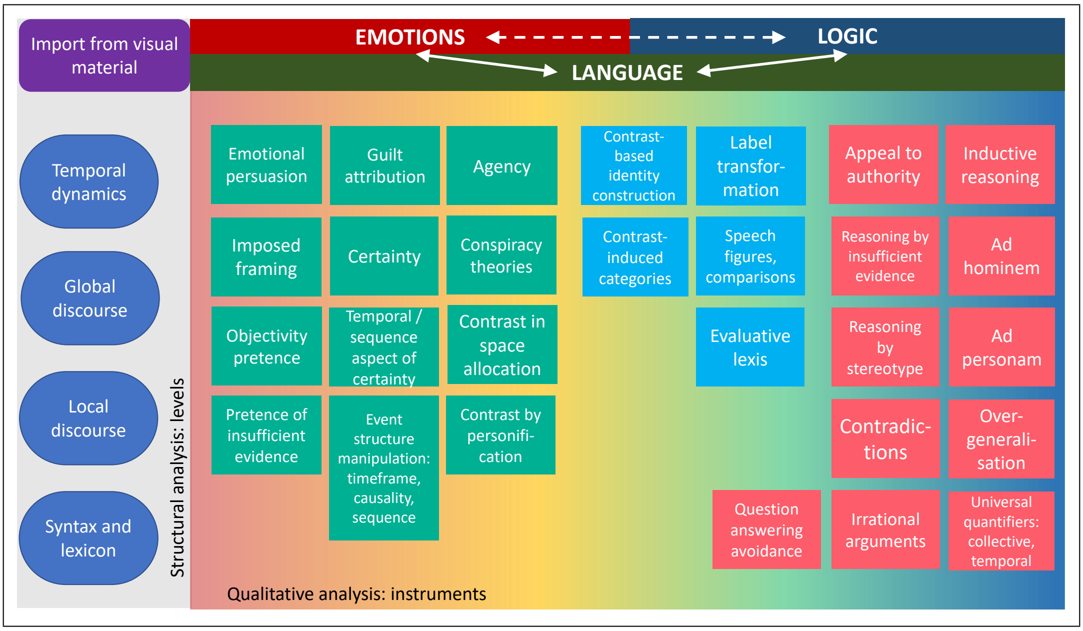


Figure 4: The instruments of bias: three dimensions are present in each instrument to a different extent (represented by the gradient in the background), and all of them can appear at all levels. *Fallacies* in red, *categorisation work* in blue and *others* in green.

reader to feel empathy although such descriptions would not be labelled as biased according to mainstream bias annotation approaches.

Imposed framing is a way of framing a message in a particular way without labelling it specifically. For instance, some articles label the events in Odessa on May 2, 2014 as a *massacre* while other articles use other bias instruments to transmit the message that it was a *massacre* while avoiding such extreme labels in the text.

The instruments of bias in our data are rarely used in isolation, they usually occur in combinations, integrating propositional and emotional bias, with lexical bias used now and then. For instance, Example 3.1 shows how agency, certainty, overgeneralisation and evaluative lexis are used in combination to construct two opposite identities: the murder and the victim.

Example 3.1. *EN301: agency, certainty, evaluative lexis, overgeneralisation.*

"Some 1,000 Ukrainian rightists, led by the notorious Right Sector, surrounded, stormed, and burned the House of Trade Unions in Odessa last Friday, killing 39 pro-Russia demonstrators in the building."

Instrument	Description
Inductive reasoning	Bottom-up reasoning: X caused Y once, then the second time it must be X again
Appeal to authority	A prominent person or politician says something, then it must be true
Reasoning by stereotype	Stereotypical features of social categories are taken as reasons for claims that something was possible or impossible, likely or unlikely.
Reasoning by insufficient evidence	Drawing conclusions from incomplete information, uncertainty while making negative claims about the opponents
Arguments ad hominem	Attack directed towards a person's qualities instead of reacting to arguments
Arguments ad personam	Attack on a person or group unrelated to the argument.
Over-generalisations	Application of single cases of observed qualities to similar other instances
Universal quantifiers	Use of universal quantifiers without evidence or reasoning, only claims that something happens always, and all of them are the same
Contradictions	Close to formal fallacies, clear errors in argumentation
Irrational arguments	Irrelevant information, missing the point
Question answering avoidance	Answers that miss or ignore the question

Table 1: Definitions of instruments of bias based on fallacies including emotional appeals

Example 3.2 illustrates the combined use of question answering avoidance, irrational arguments and emotional persuasion.

Example 3.2. *RU024 (translation): irrational arguments and question answering avoidance, emotional persuasion.*

"Who drove us into this building? - he answers a question with a question. - Yes, the people at Kulikovo Polye had the opportunity to leave. But I personally, as one of the leaders of the movement, could not leave everyone."

The question is not cited in the article, but from the reformulated repetition of the question in Example 3.2, it must have been something like *Who drove you into the building?*, which is already an accusation trying to establish an intent. Instead of

Instrument	Description
Label transformation	Replacement of social category labels within an article or within a sequence of articles, such as football fans → rightists
Speech figures and comparisons	Comparison of events with other known events and making them look very similar by stressing their shared properties
Evaluative lexis	Use of evaluations in descriptions of social categories to mark the good ones and the bad ones
Contrast-based identity construction	Descriptions that make one party look good and the other look bad based on descriptions of personal qualities, group composition, behaviour, category labels
Contrast-induced categories	Categorisation of persons in groups based on their contrasting features (no matter whether optional or mandatory), creation of "us" and 'them"

Table 2: Lexical instruments of bias with definitions

answering the question (question answering avoidance), the interviewee makes clear that all people must have chosen to enter the building, since the people had the opportunity to leave (irrational arguments). The speaker makes his leadership role relevant and positions himself as a responsible leader (irrational arguments), which is supposed to create emotion of sympathy and agreement that the speaker did everything right (emotional persuasion). It needs more efforts to understand that, on the logical level, the speaker replies that no one drove them into the building.

Example 3.3. *RU018 (translation): contrast-based identity construction, reasoning by stereotypes, emotional persuasion, guilt attribution, objectivity pretense.*

"The Kulikovo Polye Square became the center of attraction for anti-Western forces in Odessa long before the change of power in the country. Here, its own 'interest group' was formed, with women and the elderly in the core, in addition to the activists.

[...]

When anti-Maidanists and their opponents, who were chasing them, came running from the side of Grechskevaya Street, the traditional Kulikovo Pole crowd gathered on the steps of the Trade Union House and stayed there even when the tents began to burn. The attackers must have seen whom they were fighting with."

The contrast-based identity construction in Example 3.3 uses description of one party as "anti-Western forces" who had "women and elderly people at its core",

Instrument	Description
Emotional persuasion	Describe people who need protection, experienced injustice or suffered, personification of emotions
Guilt attribution	Blaming one party of a conflict without evidence
Agency	Taking agency out of a politician by claims that they are a puppet of some other actors
Imposed framing	Collection of indices that an event had a particular form without mentioning a specific label
Certainty	Presentation of some relationships with certainty without evidence about the course of events
Temporal/sequence aspect of certainty	Changes in certainty within a text and throughout a number of publications
Conspiracy theories	Claims that everything was planned by power elites or some other groups
Objectivity pretence	Includes detailed description of events, eyewitness interviews, video footage etc.
Pretence of insufficient evidence	Statements such as “we do not know” and “not clear” that contradict other reports
Contrast in space allocation	One party gets a lot of space allocated in an article while the other party is almost not mentioned
Contrast by personification	Presenting one party as persons via interviews, personal stories, and the other as “bulk”
Event structure manipulation	Inconsistencies in descriptions of causality and sequentiality

Table 3: Other instruments of bias with definitions

and the “activists” receive only a secondary role from this description. At the opposite side, the example describes as “their opponents who were chasing them” (guilt attribution) and the “attackers”. The “attackers” are accused of fighting against “the traditional Kulikovo Pole crowd” who were earlier described as “women and elderly people at its core” (guilt attribution). Reasoning by stereotypes suggests that the social category of “women and elderly people” needs protection, but the opposite happened (guilt attribution, emotional persuasion). The entire narrative uses historical links such as “long before the change of power in the country” and detailed descriptions of the event in focus. Both are supposed to create an impression of objective reporting (objectivity pretence).

Lexical choices are made in such a way that the reader must have no doubt what exactly happened, who is responsible and who are the victims. While other articles describe the victims as women and elderly people (Example 3.3), Example 3.4 mentions *young people* (young people are supposed to live and not to die).

Example 3.4. *PL405 (translation): certainty, evaluative lexis, guilt attribution, emotional persuasion and objectivity pretense.*

About 50 young people were murdered in such a bestial way by Ukrainian Nazis, the Kiev Bandera junta

The group accused of crime is mentioned twice using the overgeneralisation (*Kiev*, *Ukrainian*), the actions of the accused party are described in an emotionally loaded way and impose a deliberate, even planned act of crime.

Example 3.5. *EN301: evaluative lexis, guilt attribution, emotional persuasion and label manipulation.*

As the building burned some of the pro-Kiev activists said on Twitter that "Colorado beetles are being roasted up in Odessa," using a derogatory term for the St. George's ribbons worn by many of the anti-Kiev government demonstrators.

Example 3.5 demonstrates a strong lexical and phraseological bias and emotional persuasion by means of direct speech. The citation from the unity-supporters is chosen to present them as merciless and inhuman. The symbol of St. George's ribbons that is actually a historical symbol of the Russian Empire is assigned to "anti-Kiev government demonstrators", which makes the symbol itself something anti-Ukrainian. The St. George's ribbons are "revitalised" in Russia since 2005 as a symbol of the victory in the WW-II, patriotic attitude and Russian spirit[2].

3.4 Language-specific and language-independent findings

We found that some of the instruments of bias are tailored to influence speakers of a particular language by cultural references and culture-specific emotional persuasion. For instance, appeal to authority and guilt attribution in Polish texts refer to the Polish government while in Russian, German and English texts, the Polish government is not mentioned at all. In contrast, Russian texts mention Russian and Ukrainian authorities, and contrasts are created by mentioning the "West" as a collective opposite, and European institutions.

Culture-specific emotional persuasion is visible via mentions of local historical events. For instance, Russian texts compare the events in Odessa with Khatyn[3] while Polish texts make a reference to Auschwitz and Katyn.

[2] https://en.wikipedia.org/wiki/Ribbon_of_Saint_George
[3] https://en.wikipedia.org/wiki/Khatyn_massacre

However, after two experts analysed independently articles in different languages (one expert Polish and English, the other expert Russian and German texts) the identified instruments turned out to overlap to a large extent in all analysed languages.

3.5 Towards a New Definition of Bias

Based on the analysis described in this article, and reflecting on earlier definitions of bias, we came to the following, new definition of bias. First, we see that defining bias solely as asymmetries in lexical choices, as in [12], does not cover all instruments of bias. Second, the definition of bias as consistent patters in framing, as in [2] does not specify the quality of patterns. We see that bias is build upon asymmetries in social category construction created via fallacies, categorisation work and various other instruments in order to influence the perception of a social category. This kind of perception is always linked to normative reasoning ('good ones' vs. 'bad ones', 'us' vs. 'them'), therefore, we can generalise this definition as a difference between an average estimated value V of a category arrived at by a process $P(V)$ and its different value $Bias(V)$ attributed by a particular language user, following a process $BiasP[Bias(V)]$.

4 Contribution and Discussion

This research presents a multilingual qualitative study of instruments of bias and analyses 68 reports about the same event distributed over nine years. This is the first study of this kind, and despite a relatively small number of texts selected for the qualitative analysis, we were able to find more bias instruments than were previously described in academic literature discussed in Section 1.1. Our results contribute to the state of the art in bias conceptualisation, annotation and modelling in several ways.

First, our study extends the understanding of granularity levels in bias annotation. Currently dominant approaches for bias annotation distinguish between high-level annotation (usually article and source) and more fine-grained annotation (usually paragraph, sentence and concept level) [48, 49]. Our research shows that also larger levels of granularity need to be taken into consideration: local discourse, global discourse and temporal dynamics at all granularity levels. This is because biases become visible only in comparison with something unbiased or differently biased. To have this comparison, multiple texts on the level of local discourse (e.g. the same country the same time) can reveal biases that single texts will not (no matter whether annotated by sentence, paragraph or article level). Sometimes it is important to look at the global discourse because local discourses may consistently call an incident a "holocaust" like in our dataset, or they may consistently call a war a "special operation", and biases will not be detectable at a local discourse level. Temporal dynamics in bias (bias strengthening over time, e.g. in certainty, extreme formulations, label replacement) need to be captured, and it can be done by qualitative changes in annotation on the levels of words, sentences and paragraphs, for example by adding types of bias from our identified instruments, and not solely marking them as biased or neutral. Therefore we argue that taking all levels into account when modelling bias will help to identify bias in a more comprehensive way.

Second, our study shows that referential and non-referential functions of language are employed together in bias creation. However, state-of-the-art approaches to bias annotation mentioned in Section 1.1 mainly rely on the referential function of language, and handle language as a socially acquired system of symbols that store meaning [50]. Other important functions of language such as creation of identities, community construction, group formation and social bonding [51, 52, 53] are not taken into account for annotation of bias in linguistic data. However, as our study shows, diverse instruments of bias are used to create opposite identities of "good" and "bad" ones and to elicit a feeling of a group membership in the readers. Nevertheless, the referential function of language is important in the selection and use of the instruments of bias because lexical choices (among other means) activate particular parts of the socially acquired knowledge in readers (see the discussion about scripts, frames and social practices in Section 1.1). In this way, non-referential functions of language are linked to referential functions via linguistic context. Thus, all bias instruments must be annotated as a *system* in order to supply new high-quality labelled data to computational models of bias.

Third, the number of bias instruments discovered in this small-scale study is much larger than those mentioned in state-of-the-art annotation guidelines. For example, [54] recognise that lexical and argumentation techniques are used in mixture for propaganda purposes. Their most recent list of propaganda techniques includes 22 categories obtained from literature and applied to English texts [43]. While appeal to emotion is seen in [43] as only related to pictures, our study shows that appeal to emotion in text is a very frequently used instrument of bias. Transfer (a.k.a. Association) is another technique used in [43] for pictures only, and it aims at eliciting emotional response by comparing one entity with another and borrowing positive or negative properties from the other entity. Our research makes use of MCA methods and shows how such transfers work in text by category label manipulation. However, transfer is only one possible type of label manipulation. We also show that such techniques are tailored to a particular target audience by providing culturally relevant comparisons (e.g. references to locally relevant historical events).

Fourth, this work extends other conceptualisations of framing in political matters. The framing dimensions presented by [30] and discussed in Section 1.1 are also applicable for analysis of non-policy-related reporting, such as reports about political conflicts. Our classification looks at the mechanics of bias and advances our understanding of how frames can be created regardless of their topics.

Fifth, this work extends our understanding how speaking *about*, *as* and *to* members of specific social categories is reflected in biases. [55] distinguish among speaking *about*, *as* and *to* a particular gender in their gender bias classifications. This classification can be also applied to other personal attributes and is supported by linguistic research [12]. In our corpus we can distinguish between Maidan and Anti-Maidan supporters, and we found text passages speaking *from* and *about* those two identities. However, both voices can be used in multiple roles simultaneously, for instance to speak from the category with the purpose to speak about the category, as Example 3.5 shows. Thus, although those voices are analysable different, their function in texts can be ambiguous and manipulative.

Theoretical results show that biases become stronger over time, for example the game theoretic approach validated on a dataset of Telegram messages mainly focusing on dynamics in category labelling [20, 21]. Our research develops a better understanding what precisely changes over time and which instruments of bias are suitable for it. In addition to label manipulation, they also include certainty in reasoning, emotional bias amplification, as well as cross-source and cross-linguistic connections.

Research presented in this article also contributes to a better understanding of identity creation and its (mis)use in biased news articles. In particular, culturally specific stereotypes of gender and local ethnic groups are frequently used in various bias instruments to activate *particular* socially acquired knowledge. We saw in our data that gender stereotypes are used to construct a positive or negative identity of a party of a political conflict (stereotype-based reasoning). This confirms the result presented in [56] that bias form can differ from bias function. In particular, personal attributes such as ethnicity, gender, race, religion, sexuality and appearance are used by the instruments of bias in political reports with the purpose of population manipulation. As discussed in Example 3.3, gender and age stereotypes can be used for contrast-based identity creation, as in this case *women and elderly people* are the category that needs protection, but has been *attacked* instead, according to the presentation in the text.

[57] analyse linguistic features that are associated with differences in racial identity images. The work shows that it is possible to differentiate between a person writing from their own racial identity and the same person writing from a different racial identity. However, the "black language" stereotype turned our to be stronger

than the "white language" stereotype. The word-level analysis showed that the other-race-identity construction employed tokens related to appearance and interests, that were not used in same-race descriptions. Related to our dataset, the "nazi" stereotype as it is built on WW-II heritage in post-soviet countries, seems stronger than the "democratic" stereotype that is being cultivated in the Western countries and in the post-Maidan Ukraine.

Emotions such as pity, disgust, anger and fear play an important role in bias modelling, as our research and earlier studies show. [58] investigate the use of impoliteness and persuasive emotionality in media texts. The authors show how negative emotions are instrumentalised in media, how event descriptions are framed ideologically by using implicit emotional persuasion. Our study confirms the findings and includes the role of lexical choices and logical argumentation into the persuasion model. Playing with negative emotions tells the reader, who is "good" and who is "bad" in the narrative. Because it also has been shown that negative emotions such as anger and sadness lead to increased sharing of political news [59], and because especially the sadness bias supports the news believability, as [60] show, it is important to identify all emotional appeals and emotional persuasion in news articles, and not only those explicitly labelled.

In the particular case of the Odessa fire, the users rely on the emotional framing of the news while searching for truth, and this fact is used as an opportunity for creation of an "alternative reality" [44]. Because research shows cultural differences in emotion clustering (for fear - disgust - anger cluster see [61]) is also important to understand emotions as *culturally acquired social practice* and not a universal concept. Although bias instruments in their principles provide a universal structure for persuasion, their concrete applications in concrete articles are always tailored for a particular speaker of a particular language with a particular cultural background including familiarity with particular stereotypes.

Although this article focuses on only one specific event in a specific geopolitical context, we can draw some generalisations by looking at several other events world-wide mentioned in [26], such as Brexit campaign and US presidential elections in which Donald Trump won. Techniques described by the author also target the recipients at the emotional level, although rationally they make no sense. Cultural tailoring of the information for a particular population happened in the campaigns described by [26] by addressing cognitive biases and stereotyped perception of social groups to manipulate their voting behaviour. Our work discloses how such manipulations are implemented at the operational level.

5 Limitations

Although this research made a contribution to linguistic bias research, it also has several limitations. Qualitative research has been criticised for its subjectivity and missing scalability. This study did not employ any statistically representative datasets, and therefore, this study does not make any statistically-based claims. The strength of this type of research is in its bottom-up nature that allows for discovery of new entities. The input of this research, however, can support future quantitative studies.

This study did not aim at studying receivers' reactions. Reactions of recipients may vary and stay in line with the speaker's biases or be completely opposite. We can only assume that a number of readers got convinced and their beliefs were changed, and the present study partially reflects potential effect. In this way, critics may judge the study as subjective. We mitigate this risk by general-level descriptions of bias instruments, with no relations to particular personal and subjective emotions.

We can also assume that our study does not list all existing instruments of bias, neither can we predict that no new instruments can be produced. What this study does, on the other hand, is to show the variety of instruments in types, levels and quality, and their functioning as a system that involves logic, language and emotions.

6 Conclusions and Future Work

This basic classification of the instruments of bias shows that the issue is very complex, and we need to be cautious about simplifications, especially when making attempts of "debiasing" linguistic corpora, pre-trained language models and measuring bias in downstream tasks. Models of bias in NLP tend to simplify the phenomenon at the moment of its conceptualisation. Although it might be convenient to define discrete classes of bias and detect them based on a set of discrete types of features, our qualitative study suggests that the dimensions of bias may have liquid boundaries and/or operate in combinatorial clusters, hence they need to be studied as a context-bound system, sensitive to world knowledge. The value of logical modelling based on this research is likely to remedy this shortfall.

Our study has implications for the field of AI, and in particular for the NLP and logic communities. For NLP-related approaches for bias annotation, bias classification and bias correction, our study shows that researchers need to look beyond lexical choices, and language models based on distributional semantics allow only partially for insights into logical structure and emotions. Further, our study shows that a separation of logic, language and emotion in the analysis of bias would not reflect a complete picture. As the results of our study indicate, those three dimensions are mutually dependent and interrelated.

This study also has implications for online content moderation and similar applications. We clearly see a need for new models to identify emotional persuasion and fallacies as part of content moderation [62]. This will help to overcome the limitations of currently implemented tools for content moderation relying on lexical choices, such as simple word filters. However, we must be clear about potential misuse of new models for information censorship.

Our research also has implications for policy making. For instance, personal data protection regulation requires websites to obtain consent from users to share their personal data, and the user interfaces for this are purposefully designed to mislead the users. This problem is referred to as *dark patterns*, and users often feel manipulated when using online services [63]. Applying the bias instruments concept and an operationalisation of emotional persuasion can help to formulate better policies for regulation of this issue. Bias awareness can be also made part of digital literacy and information literacy education.

We identified research questions that we can target in our future research. Although it is difficult to calculate objective values for each of the dimensions in each instrument, an approximation can be proposed for computational purposes. For example, different degrees of negative or positive emotions can be calculated based on arousal and valence [64, 65]. Furthermore, new emotion research approaches introduce two further relevant emotion value distinctions: power or emotion control, as well as a degree of emotion unexpectedness [66]. Lexical biases can be analysed further with the help of membership categorisation analysis to tackle the category label manipulation more precisely [32, 67]. Linguistic structure on the other hand will help us to uncover degrees of the explicit and implicit bias profiles [68]. Propositional biases, in turn, can be approached with a new type of logic [69]. Since bias shifts meaning also in time, and logic is in principle capable of doing the same, we need to adapt logic to replicate the mechanics of bias.

We also note in our data analysis that some of the instruments of bias appear in combinatorial clusters. It would be interesting to scrutinize this phenomenon in a wider perspective in diverse contexts in future research. It would be interesting to explore whether grouping texts by their instruments enable further, deeper insights into biased news, misinformation and propaganda poison. Future research could also investigate whether the combinations of bias instruments change over time and how they contribute to temporal manipulation of the categories and event structure.

References

[1] S. Oates, Russian state narrative in the digital age: Rewired propaganda in Russian television news framing of Malaysia Airlines flight 17, in: American Political Science

Association Annual Meeting, 2014.

[2] R. M. Entman, Framing bias: Media in the distribution of power, Journal of communication 57 (1) (2007) 163–173.

[3] H. Seo, H. Ebrahim, Visual propaganda on facebook: A comparative analysis of Syrian conflicts, Media, War & Conflict 9 (3) (2016) 227–251.

[4] O. Ogunyemi, Shaping the perception of African conflicts through framing: A case study of the African diasporic press in the UK, Media, War & Conflict 11 (4) (2018) 421–433.

[5] P. Schlesinger, Media, state and nation: Political violence and collective identities, Sage, 1991.

[6] D. Machin, T. Van Leeuwen, Global media discourse: A critical introduction, Routledge, 2007.

[7] S. Johnson, A. Ensslin, Language in the media: Representations, identities, ideologies, A&C Black, 2007.

[8] S. L. Blodgett, S. Barocas, H. Daumé III, H. Wallach, Language (technology) is power: A critical survey of "bias" in NLP, in: Proc. 58th ACL Meeting, ACL, 2020, pp. 5454–5476.

[9] F. Hamborg, Towards automated frame analysis: Natural language processing techniques to reveal media bias in news articles, Ph.D. thesis, University of Konstanz, Konstanz (2022).

[10] E. Manzoor, N. B. Shah, Uncovering latent biases in text: Method and application to peer review, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, 2021, pp. 4767–4775.

[11] R. Liu, C. Jia, J. Wei, G. Xu, L. Wang, S. Vosoughi, Mitigating political bias in language models through reinforced calibration, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, 2021, pp. 14857–14866.

[12] C. J. Beukeboom, C. Burgers, Linguistic bias, in: Oxford research encyclopedia of communication, 2017.

[13] C. Hube, Bias in Wikipedia, in: Proc. 26th Int. Conf. on World Wide Web Companion (WWW'17), 2017, pp. 717–721.

[14] X. Ferrer Aran, T. van Nuenen, J. Such, N. Criado Pacheco, Discovering and categorising language biases in Reddit, in: Proc. 15th Int. AAAI Conference on Web and Social Media (ICWSM'21), 2021, pp. 140–151.

[15] K. Lazaridou, R. Krestel, Identifying political bias in news articles, Bull. IEEE Tech. Comm. Digit. Libr. 12 (2016).

[16] R. R. Reddy, S. R. Duggenpudi, R. Mamidi, Detecting political bias in news articles using headline attention, in: Proc. 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP, Association for Computational Linguistics, Florence, Italy, 2019, pp. 77–84.

[17] T. Quijote, A. Zamoras, A. Ceniza, Bias detection in Philippine political news articles using SentiWordNet and inverse reinforcement model, in: IOP Conference Series:

Materials Science and Engineering, Vol. 482, IOP Publishing, 2019, p. 012036.

[18] L. Herzig, A. Nunes, B. Snir, An annotation scheme for automated bias detection in wikipedia, in: Proc. 5th Linguistic Annotation Workshop (LAW'11), Association for Computational Linguistics, 2011, pp. 47–55.

[19] T. Spinde, L. Rudnitckaia, K. Sinha, F. Hamborg, B. Gipp, K. Donnay, MBIC–a media bias annotation dataset including annotator characteristics, arXiv preprint arXiv:2105.11910 (2021).

[20] S. Höhn, N. Asher, S. Mauw, Examining linguistic biases in Telegram with a game theoretic analysis, in: Proc. 3rd Multidisciplinary International Symposium on Disinformation in Open Online Media (MISDOOM'21), Vol. 12887 of LNCS, Springer-Verlag, 2021, pp. 16–32.

[21] N. Asher, J. Hunter, When learning becomes impossible, in: 2022 ACM Conference on Fairness, Accountability, and Transparency, 2022, pp. 107–116.

[22] R. C. Schank, R. P. Abelson, Scripts, plans, and knowledge, in: IJCAI, Vol. 75, 1975, pp. 151–157.

[23] M. Minsky, A framework for representing knowledge, MIT, Cambridge, 1974.

[24] A. Reckwitz, Toward a theory of social practices: A development in culturalist theorizing, European Journal of Social Theory 5 (2) (2002) 243–263.

[25] H. Meng, Social script theory and cross-cultural communication, Intercultural Communication Studies 17 (1) (2008) 132–138.

[26] C. Wylie, Mindf*ck: Inside Cambridge Analytica's Plot to Break the World, Profile Books Ltd., 2020.

[27] A. López Ortega, Are microtargeted campaign messages more negative and diverse? An analysis of Facebook ads in European election campaigns, European Political Science 21 (3) (2022) 335–358.

[28] T. Gafus, Microtargeting staatlicher Stellen, AfP 53 (4) (2022) 285–293.

[29] K. Haenschen, The conditional effects of microtargeted Facebook advertisements on voter turnout, Political Behavior (2022) 1–21.

[30] A. E. Boydstun, D. Card, J. Gross, P. Resnick, N. A. Smith, Tracking the development of media frames within and across policy issues (2014). `doi:10.1184/R1/6473780.v1`.

[31] H. Sacks, Lectures on Conversation, Vol. 1 and 2, Blackwell, 1995.

[32] J. Freiberg, P. Freebody, Applying membership categorisation analysis to discourse: When the 'tripwire critique' is not enough, in: T. Le, Q. Le, M. Short (Eds.), Critical discourse analysis: An interdisciplinary perspective, Nova Science Publishers, 2009, pp. 49–64.

[33] E. A. Schegloff, A tutorial on membership categorization, Journal of Pragmatics 39 (3) (2007) 462–482.

[34] D. Walton, The place of emotion in argument, Penn State Press, 1992.

[35] P. Goffredo, S. Haddadan, V. Vorakitphan, E. Cabrio, S. Villata, Fallacious argument classification in political debates, in: L. D. Raedt (Ed.), Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22, International Joint

Conferences on Artificial Intelligence Organization, 2022, pp. 4143–4149, main Track.
[36] H. Sacks, On the analyzability of stories by children, in: Directions in Sociolinguistics. The Ethnography of Communications, Holt, Rinehart and Winston, NY, 1972, pp. 325–345.
[37] P. Eglin, S. Hester, Moral order and the Montreal massacre: A story of membership categorization analysis, Media studies: Ethnomethodological approaches (1999) 195–230.
[38] S. Hester, P. Eglin, Culture in action: Studies in membership categorization analysis, no. 4, University Press of America, 1997.
[39] A. Gibson, G. Rowe, C. Reed, A computational approach to identifying formal fallacy, CMNA VII-Computational Models of Natural Argument (2007).
[40] K. W. Saunders, Informal fallacies in legal argumentation, SCL Rev. 44 (1992) 343.
[41] R. Mao, Q. Liu, K. He, W. Li, E. Cambria, The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection, IEEE Transactions on Affective Computing (2022).
[42] G. Da San Martino, S. Yu, A. Barrón-Cedeno, R. Petrov, P. Nakov, Fine-grained analysis of propaganda in news article, in: Proc. EMNLP-IJCNLP-2019, 2019, pp. 5636–5646.
[43] G. Da San Martino, A. Barrón-Cedeño, H. Wachsmuth, R. Petrov, P. Nakov, Semeval-2020 task 11: Detection of propaganda techniques in news articles, in: Proceedings of the Fourteenth Workshop on Semantic Evaluation, 2020, pp. 1377–1414.
[44] V. Zvereva, Social media users in search of 'facts': the trade union house fire case, Russian Journal of Communication 13 (1) (2021) 11–28.
[45] J. Pustejovsky, The syntax of event structure, cognition 41 (1-3) (1991) 47–81.
[46] J. Sim, B. Saunders, J. Waterfield, T. Kingstone, Can sample size in qualitative research be determined a priori?, International journal of social research methodology 21 (5) (2018) 619–634.
[47] S. Hoehn, Artificial companion for second language conversation, Springer, 2019.
[48] D. Aleksandrova, F. Lareau, P. A. Ménard, Multilingual sentence-level bias detection in Wikipedia, in: Proc. Int. Conf. on Recent Advances in Natural Language Processing (RANLP'19), Varna, Bulgaria, 2019, pp. 42–51.
[49] W.-F. Chen, K. Al-Khatib, H. Wachsmuth, B. Stein, Analyzing political bias and unfairness in news articles at different levels of granularity, arXiv preprint arXiv:2010.10652 (2020).
[50] M. B. Trosdal, Meaning: The referential function of language, Philippine quarterly of culture and society 23 (3/4) (1995) 361–368.
[51] A. Sowińska, T. Dubrovskaya, Discursive construction and transformation of 'us' and 'them' categories in the newspaper coverage on the US anti-ballistic missile system: Polish versus Russian view, Discourse & Communication 6 (4) (2012) 449–468.
[52] K. Whitehead, G. Lerner, When are persons 'white'?: On some practical asymmetries of racial reference in talk-in-interaction, Discourse & Society 20 (2009) 613–641. doi:

10.1177/0957926509106413.

[53] N. Shrikant, "It's like,'I've never met a lesbian before!'": Personal narratives and the construction of diverse female identities in a lesbian counterpublic, Pragmatics 24 (4) (2014) 799–818.

[54] G. Da San Martino, S. Yu, A. Barrón-Cedeño, R. Petrov, P. Nakov, Fine-grained analysis of propaganda in news article, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), ACL, 2019, pp. 5636–5646. doi:10.18653/v1/D19-1565.
URL https://aclanthology.org/D19-1565

[55] E. Dinan, A. Fan, L. Wu, J. Weston, D. Kiela, A. Williams, Multi-dimensional gender bias classification, in: Proc. 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP'20), 2020, pp. 314–331.

[56] S. Hoehn, S. Mauw, N. Asher, BelElect: a new dataset for bias research from a "dark" platform, in: The 16th International AAAI Conference on Web and Social Media (ICWSM'22), 2022, pp. 1–8.

[57] G. Kambhatla, I. Stewart, R. Mihalcea, Surfacing racial stereotypes through identity portrayal, in: 2022 ACM Conference on Fairness, Accountability, and Transparency, 2022, pp. 1604–1615.

[58] B. Lewandowska-Tomaszczyk, P. Pezik, Emergent impoliteness and persuasive emotionality in polish media discourses, Russian Journal of Linguistics 25 (3) (2021) 685–704.

[59] E. De León, D. Trilling, A sadness bias in political news sharing? the role of discrete emotions in the engagement and dissemination of political news on Facebook, Social Media + Society 7 (4) (2021) 1 – 12.

[60] B. Deng, M. Chau, The effect of the expressed anger and sadness on online news believability, Journal of Management Information Systems 38 (4) (2021) 959–988.

[61] B. Lewandowska-Tomaszczyk, P. A.Wilson, Contrasts and analogies in cluster categories of emotion concepts in monolingual and cross-linguistic contexts, Benjamin, 2022, pp. 405 – 437.

[62] T. Gillespie, Content moderation, ai, and the question of scale, Big Data & Society 7 (2) (2020) 2053951720943234.

[63] K. Bongard-Blanchy, A. Rossi, S. Rivas, S. Doublet, V. Koenig, G. Lenzini, "I am definitely manipulated, even when I am aware of it. it's ridiculous!"-dark patterns from the end-user perspective, in: Designing Interactive Systems Conference 2021, 2021, pp. 763–776.

[64] R. A. Calvo, S. Mac Kim, Emotions in text: dimensional and categorical models, Computational Intelligence 29 (3) (2013) 527–543.

[65] T. Eerola, J. K. Vuoskoski, A comparison of the discrete and dimensional models of emotion in music, Psychology of Music 39 (1) (2011) 18–49.

[66] J. J. R. Fontaine, K. R. Scherer, E. B. Roesch, P. C. Ellsworth, The world of emotions is not two-dimensional, Psychological Science 18 (12) (2007) 1050–1057.

[67] P. Eglin, S. Hester, The Montreal massacre: A story of membership categorization analysis, Wilfrid Laurier Univ. Press, 2003.

[68] B. Lewandowska-Tomaszczyk, P. Wilson, Physical and moral disgust in socially believable behaving systems in different cultures., in: A. Esposito, L. Jain (Eds.), Socially Believable Behaving Systems, Berlin: Springer, 2016, pp. 105–132.

[69] D. M. Gabbay, L. Rivlin, Heal2100: Human effective argumentation and logic for the 21st century. the next step in the evolution of logic, IfCoLog Journal of Logics and Their Applications 4 (6) (2017) 1633–1687.

Received 25 November 2022

Height of valuations in non-deterministic semantics for modal logic.

Pawel Pawlowski
CLPS, Ghent University, Belgium
haptism89@gmail.com

Abstract

This paper examines the height of Kearns valuations in the context of non-deterministic semantics for modal logics. The authors first show that the height of these valuations can be reduced to one for logics that extend $\mathbf{H4}^{-}$. This generalizes previous work by [34].

Next, we demonstrate that there is no relationship between the number of non-equivalent modalities that can be defined in a given logic and the height of the valuations that are needed to regain the rule of necessitation. This refutes an implicit conjecture put forward by Omori and Skurt in their 2016 work [34]. The authors also show that the reverse implication is false.

As a byproduct of this study, non-deterministic semantics for several modal logics are provided. In conclusion, this paper makes significant contributions to the study of non-deterministic semantics and the application of Kearns valuations in this context.

1 Introduction

Non-deterministic semantics is a generalization of the traditional semantics used for finitely many-valued logics. The main difference between classical and non-deterministic semantics lies in the way that connectives are interpreted. In deterministic semantics, connectives are interpreted as functions that map from the Cartesian power of the set of values to the set of values. In non-deterministic semantics, connectives are still functions from the same set as in classical semantics, but they are mapped to the power set of the set of values excluding the empty set.

The author of the paper would like to acknowledge that this work has been supported by the BOF (Bijzonder Onderzoeksfonds) funds. Additionally, the author would like to thank to anonymous reviewers for their comments and effort.

This means that the interpretation of a complex formula constructed using a given connective may not be uniquely determined.

By allowing for this flexibility, non-deterministic semantics is more versatile and can be used to characterize a wider range of logics than is possible with deterministic semantics. This added flexibility makes non-deterministic semantics a valuable tool for studying the meaning of complex formulas in modal logics and other many-valued logics.

The origins of non-deterministic semantics can be traced back to the works of Zich in 1938 [41, 37]. Based on these ideas, Rescher [38] developed the first proper non-deterministic semantics to study the natural language conditional. In our paper, we apply non-deterministic semantics to characterize modal logics. The first such application was studied by Ivlev [29, 24]. Ivlev's work focused on semantics for non-normal modal logics, which are not closed under the rule of necessitation. The central concepts of his semantics were quasi-functions and quasi-matrices, which can be seen as semi-formal versions of what is now known as nmatrices. Ivlev stated completeness theorems for a family of non-normal modal logics without providing a proof, and his approach was later developed for other modal logics [30, 27, 28, 25, 26]. According to Ivlev's account, there are four truth-values: necessarily true, contingently true, contingently false, and necessarily false. He also developed two and three-valued modal semantics in his work.

The paper by Kearns in [31] discusses the use of non-deterministic semantics in modal logic, focusing on normal modal logics **T**, **S4**, and **S5**. Unlike Ivlev, Kearns uses a filtration method on the set of valuations, called the mth level valuations, to restrict the admissible valuations. This approach was further studied and generalized in [34, 21, 22].

In [34], the authors extended the set of values by separating the possibility, truth, and necessity of a given proposition, and used the resulting non-deterministic semantics to study **K**, **KD**, and **KTB**. In [21], non-deterministic semantics for **KD**, **KDB**, **KD4**, and **KD45** were presented. These results were simplified and further extended in [23, 35, 36]. Finally, in [22], the authors developed the first-order counterparts of some of these semantics.

The proper formalization and meta-theory for non-deterministic semantics has been studied in [12, 13, 3, 2, 10, 8, 5, 14, 39, 9, 7, 6, 15, 40, 1, 4, 18, 16, 11]. One of the main motivations there was to semantically study a particular type of Gentzen's sequent systems.

One of the other main applications of non-deterministic semantics is paraconsistent logics. The relation between non-deterministic semantics for paraconsistent logics and non-deterministic semantics for modal logic has not yet been studied thoroughly. There might be an interesting connection between these two approaches via translation results as pointed out in [23].[1]

Despite the number of publications, the use of non-deterministic semantics in modal logic has not yet become mainstream. One of the main challenges with these semantics is the lack of a decision procedure for level valuations. In general, it is difficult to determine which starting valuations will not be filtered out at a given level. However, recent work by [23, 32] has constructed decision procedures for some variants of the level valuation technique. These results may be generalized to other cases.

The relationship between level valuations and possible world semantics remains largely unexplored. Additionally, it is not clear what criteria are necessary and sufficient for a given modal logic to be described by non-deterministic semantics. Further research in these areas could help to improve our understanding of non-deterministic semantics in modal logic.

The main aims of this paper are as follows. The first aim is to extend the result presented in [34] regarding the limitations of the height of the m-th level valuation hierarchy needed to regain NEC. In [34], it was shown that for the NEC-free fragment of **S4**, the first-level valuations are sufficient to regain NEC [Theorem 4 in [34]]. We demonstrate that this result holds for a much weaker logic, and that axiom $\mathsf{4}$ is crucial for this theorem.

The second aim of this paper is to address the conjecture mentioned in Remark 41 of [34], which suggests that there may be a relationship between the number of non-equivalent modalities in a given modal logic and the level of the m-th level hierarchy needed for NEC. We show that this claim and its opposite are false by providing suitable counter-examples.

In the next section, we introduce the necessary technical concepts. Then, in the third section, we extend the result from [34] and show that the very weak logic $\mathbf{H4}^-$ requires only a finite number of m-th level valuations to regain the full rule of necessitation. In the fourth section, we address the hypothesis proposed by Omori and Skurt, and provide counter-examples to show that it is false. Finally, in the last section, we summarize the findings of this paper and discuss potential avenues for further research.

[1] Avron [10, 6, 15, 17, 7, 1, 4] constructed a non-deterministic semantics for a number of logics in the so-called family of *logics of formal inconsistency* (**LFI**). These are logics that aim at capturing inconsistencies within the object language. These results were lifted-up to the first-order setting [15, 40, 17, 6, 11]. [33] presented a three-valued non-deterministic semantics for Janusz Ciuciura's logic $\mathbf{mbC^1}$ which is the logic **mbC** formalized in the language without the inconsistency operator.

2 Technical preliminaries

In this paper, we use two modal propositional languages, $\mathcal{L}\Box\Diamond$ and $\mathcal{L}\Box$, which consist of propositional variables (Var) and the connectives $\neg, \to, \Box, \Diamond$ and $\neg, \to, \Box$ respectively. The remaining classical Boolean connectives are treated as abbreviations. In the case of $\mathcal{L}\Box$, we define $\Diamond$ as $\neg\Box\neg\varphi$. We use lowercase Greek letters as metavariables for formulas and uppercase Greek letters for sets of formulas. The central concept in non-deterministic semantics is the non-deterministic matrix (nmatrix for short).

Definition 1 (nmatrix)**.** An *nmatrix* is a triple $\mathtt{M} = \langle \mathtt{Val}, \mathtt{D}, \mathtt{O}\rangle$, where:

- $\mathtt{Val}$ is a non-empty set of truth values.
- $\emptyset \neq \mathtt{D} \subseteq \mathtt{Val}$ is a set of designated values. By $\overline{\mathtt{D}}$ we mean the set of non-designated values i.e $\overline{\mathtt{D}} = \{x \mid x \in \mathtt{Val} \wedge x \notin \mathtt{D}\}$
- $\mathtt{O}$ is a set $\{\overline{\neg}, \overline{\to}, \overline{\Box}, \overline{\Diamond}\}$ [or the set $\{\overline{\neg}, \overline{\to}, \overline{\Box}\}$ in the case of $\mathcal{L}\Box$] of functions $\overline{\circ} : \mathtt{Val}^n \to 2^{\mathtt{Val}} \setminus \{\emptyset\}$ for $\overline{\circ} \in \mathtt{O}$, where n is the arity of the connective. This set provides the interpretation of the connectives of the language.

It is easy to see that the concept of an nmatrix is a generalization of deterministic matrices. In the case of deterministic matrices, the set $\mathtt{O}$ consists of functions that assign a single possible value to each complex formula, given the values of its components. In nmatrices, the functions in $\mathtt{O}$ assign non-empty sets of values to each formula, which may include sets with more than one element. In this case, any valuation picks exactly one value from the sets of possible values assigned to a formula by the functions in $\mathtt{O}$.

The above is a reason why sometimes non-deterministic semantics is called quasi-extensional. The value of a complex formula is not fully-determined by the values of components, but it is severely restricted since the interpretation of the complex formula assigns to it a unique and non-empty set of values.

Definition 2 (Valuation)**.** A *valuation* $v : \mathcal{L}\Box \to \mathtt{Val}$ in an nmatrix $\mathtt{M}$ is a function such that for any connective $\circ^n$ of the language and for any sequence of formulas $\varphi_1, \varphi_2 \dots, \varphi_n$, we have:

$$v(\circ^n(\varphi_1, \varphi_2, \dots, \varphi_n)) \in \overline{\circ^n}(v(\varphi_1), v(\varphi_2), \dots, v(\varphi_n))$$

The modification of the notion of valuation for the language $\mathcal{L}\Box\Diamond$ is as expected.

A *valuation* v *satisfies* φ in $\mathtt{M}$ (notation $v \vDash_{\mathtt{M}} \varphi$) iff $v(\varphi) \in \mathtt{D}$. We say that φ *follows from* Γ (notation $\Gamma \vDash_{\mathtt{M}} \varphi$) iff every valuation that satisfies all elements of Γ

satisfies φ. Sometimes we will refer to $\Gamma \vDash_{\mathtt{M}} \varphi$ as the consequence relation induced by the nmatrix $\mathtt{M}$. We also use the abbreviation $\vDash_{\mathtt{M}_1} \subseteq \vDash_{\mathtt{M}_2}$ to mean that for any Γ, φ, $\Gamma \vDash_{\mathtt{M}_1} \varphi$ implies $\Gamma \vDash_{\mathtt{M}_2}$. An nmatrix $\mathtt{M}$ is k-valued iff $|\mathtt{Val}| = k$. By a partial valuation we mean a restriction of the valuation to a subset closed under sub-formulas.

Definition 3 (Consequence relation and tautology)**.** Let $\mathtt{M}$ be a nmatrix. We say that φ is an $\mathtt{M}$ tautology $[\vDash_{\mathtt{M}} \varphi]$ iff for any $\mathtt{M}$-valuation v, $v(\varphi) \in \mathtt{D}$. We say that $\Gamma \vDash_{\mathtt{M}} \varphi$ iff for any valuation that assigns to every element of Γ a designated value, it assigns a designated value to φ as well. We will refer to $\vDash_{\mathtt{M}}$ as the logic induced by the nmatrix $\mathtt{M}$.

The first system that we are going to introduce is $\mathbf{H4}^-$.[2] It is a very weak modal system in the language $\mathcal{L}_\Box$ that has the following axioms and rules of inference:

1. Propositional tautologies.
2. Axiom 4, $\Box\varphi \to \Box\Box\varphi$.
3. *Modus ponens*

By $\vdash_{\mathbf{H4}^-}$ we denote the theoremhood relation of this system. We use $\vdash^{\mathsf{NEC}}_{\mathbf{H4}^-}$ to denote the theoremhood relation of the closure of $\vdash_{\mathbf{H4}^-}$ under the rule of necessitation. Note that $\mathbf{H4}^-$ is not a normal modal system. It does not even validate the substitution of provably equivalent formulas. This implies that semantics for this system has to be even weaker than neighborhood semantics. Fortunately, this logic has rather simple non-deterministic semantics.

Definition 4 ($\mathbf{H4}^-$)**.** $\mathtt{M}_{\mathbf{H4}^-} = (\mathtt{Val}_{\mathbf{H4}^-}, \mathtt{D}_{\mathbf{H4}^-}, \mathtt{O}_{\mathbf{H4}^-})$ is an nmatrix where:

- $\mathtt{Val}_{\mathbf{H4}^-} = \{\mathtt{T}, \mathtt{t}, \mathtt{f}, \mathtt{F}\}$.
- $\mathtt{D}_{\mathbf{H4}^-} = \{\mathtt{T}, \mathtt{t}\}$.
- and $\mathtt{O}_{\mathbf{H4}^-}$ is defined by the following truth-tables:

We will use $\vDash_{\mathtt{M}_{\mathbf{H4}^-}}$ to denote the consequence relation induced by this matrix.

Theorem 1. $\mathtt{M}_{\mathbf{H4}^-}$ *is sound and complete with respect to* $\vdash_{\mathbf{H4}^-}$.

[2]The system is named after Lloyd Humberstone since he was the first one, as far as we know to study such a weak modal systems.

φ	$\neg$	$\Box\varphi$
T	$\overline{\mathtt{D}}$	T
t	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
f	D	$\overline{\mathtt{D}}$
F	D	T

$\Rightarrow$	T	t	f	F
T	D	D	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
t	D	D	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
f	D	D	D	D
F	D	D	D	D

Table 1: Connectives of **H4**$^-$

Proof. Let us start with the soundness. Observe that the interpretations of Boolean connective behave classically. This means that any classical propositional tautology is a tautology of $\mathtt{M}_{\mathbf{H4}^-}$ and that $\mathtt{M}_{\mathbf{H4}^-}$ is closed under *modus ponens*. To see that $\Box\varphi \to \Box\Box\varphi$ is a tautology, observe that if $v(\Box\varphi) \in \mathtt{D}$, then $v(\Box\Box\varphi) \in \mathtt{D}$.

Let us move to the completeness part. We start with the following lemma:

Lemma 1 (**H4**$^-$ valuation Lemma)**.** *Let* Γ *be an* **H4**$^-$*-relatively maximal set. The following function is a* $v_{\mathbf{H4}^-}$*-valuation:*

$$v_\Gamma(\varphi) = \begin{cases} \mathtt{T} & \text{if } \Gamma \vdash_{\mathbf{H4}^-} \Box\varphi, \Gamma \vdash_{\mathbf{H4}^-} \varphi \\ \mathtt{t} & \text{if } \Gamma \vdash_{\mathbf{H4}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{H4}^-} \varphi \\ \mathtt{F} & \text{if } \Gamma \vdash_{\mathbf{H4}^-} \Box\varphi, \Gamma \vdash_{\mathbf{H4}^-} \neg\varphi \\ \mathtt{f} & \text{if } \Gamma \vdash_{\mathbf{H4}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{H4}^-} \neg\varphi \end{cases}$$

Proof. First observe that the above function is well-defined. The proof proceeds by the induction on the complexity of the formula φ. The base case is straightforward. Assume that the lemma holds for formulas φ and ψ. We will show that it is also the case for $\neg\varphi, \Box\varphi, \varphi \to \psi$.

We start with $\neg\varphi$. Assume $v_\Gamma(\varphi) \in \{\mathtt{T}, \mathtt{t}\}$. We need to show that $v_\Gamma(\neg\neg\varphi) \in \mathtt{D}$. This means that we have to show that $\Gamma \vdash_{\mathbf{H4}^-} \neg\neg\varphi$. By our assumption, we have $\Gamma \vdash_{\mathbf{H4}^-} \varphi$, so by propositional logic $\Gamma \vdash_{\mathbf{H4}^-} \neg\neg\varphi$. The case for $v_\Gamma(\varphi) \in \{\mathtt{F}, \mathtt{f}\}$ is symmetric.

Let us proceed to $\Box\varphi$. Suppose that $v_\Gamma(\varphi) \in \{\mathtt{T}, \mathtt{F}\}$. We need to show that $v_\Gamma(\Box\varphi) = \mathtt{T}$. After unpacking the definitions, this case boils down to showing that from $\Gamma \vdash_{\mathbf{H4}^-}\vdash \Box\varphi$ follows $\Gamma \vdash_{\mathbf{H4}^-}\vdash \Box\Box\varphi$. Clearly it does, since $\Box\varphi \to \Box\Box\varphi$ is one of the axioms. The remaining cases for $\Box$ and all the cases for the implication are straightforward. $\Box$

The completeness proof is by contraposition. We suppose that $\Gamma \nvdash_{\mathbf{H4}^-} \varphi$. So, there is a relatively maximal set Σ that extends Γ. We can use this set to define a valuation function v_Σ by following the construction of the previous lemma. This valuation assigns designated values to all elements of Σ, hence to all elements of Γ, and $v_\Sigma(\varphi) \notin \mathtt{D}$ which shows that $\Gamma \nvDash_{\mathtt{M}_{\mathbf{H4}^-}} \varphi$.

Now, we move to the level valuations in order to regain the rule of necessitation. First, we start with a definition:

Definition 5 (Super-designated)**.** Let $\mathtt{M}$ be an nmatrix. We say that a value a is super-designated iff $a \in \mathtt{D}$ and for any φ and any valuation v, if $v(\varphi) = a$, then $v(\Box\varphi) \in \mathtt{D}$. We'll use $\mathtt{SD}$ to denote the set of all super-designated values of a given nmatrix.

Informally speaking, super-designated values are designated values that preserves the $\Box$ modality. In the case of $\mathtt{M}_{\mathbf{H4}^-}$, there is one such value, $\mathtt{T}$.
In order to get back the rule of necessitation we will recursively define the notion of m-th level valuation. Next, we show that if we restrict the set of admissible valuations to m-th level valuations we regain the NEC.

Definition 6 (m-level valuations)**.** Let $\mathtt{M}$ be a matrix and v a valuation in it. We say that:

1. v is a 0th-level $\mathtt{M}$-valuation, if v is a $\mathtt{M}$-valuation.
2. v is an $m+1$st-level $\mathtt{M}$-valuation if v is mth-level $\mathtt{M}$-valuation and v assigns super designated value(s) to every sentence φ that is mth level tautology (that gets a designated value for any mth-level $\mathtt{M}$-valuation v').
3. We say that v is an $\mathtt{M}^+$-valuation iff v is an mth-level $\mathtt{M}$-valuation for every $m \in \mathtt{N}$.

By $\vDash_{\mathtt{M}^k} \varphi$ we mean that for any k-th level valuation v, $v(\varphi) \in \mathtt{SD}$. By $\vDash_{\mathtt{M}^+} \varphi$ we mean that for any $\mathtt{M}^+$- valuation v, $v(\varphi)$ is super designated. These level valuations allows us to regain NEC. Simply, because we are eliminating valuations that invalidates the rule.
In other words this technique eliminates those valuations under which tautologies are not super-designated. This gives us the following theorem.

Theorem 2 (Completeness of $\mathbf{H4}^+$)**.** $\vdash^{\mathsf{NEC}}_{\mathbf{H4}^-} \varphi$ *iff* $\vDash_{\mathtt{M}_{\mathbf{H4}^-}} \varphi$.

Proof. The proof is a straightforward adaption of the similar proofs in [34, 35]. The proof relies on the observation that v_Γ is not only $\mathbf{H4}$-valuation but also $\mathbf{H4}^+$-valuation. $\Box$

3 Reducing the height of level valuations

In this section, we generalize the observation of [34] labeled in the paper as Theorem 4. According to this theorem, for logics extending the NEC-free fragment of modal logic **S4**, the hierarchy of valuations can be reduced only to the first level. Already at the first level we regain the closure under NEC. We show that the theorem holds even for weaker logics. The crucial property for reducing the height of the hierarchy is the validity of the axiom **4**.

Theorem 3. *Let* $\vDash_{\mathbf{L}}$ *be a logic in* $\mathcal{L}_\Box$ *that extends the logic* $\vDash_{\mathtt{M}_{\mathbf{H4}^-}}$. *For any formula* $\varphi \in \mathcal{L}_\Box$, $\vDash_{\mathbf{L}^1} \varphi$ *iff* $\vDash_{\mathbf{L}^+} \varphi$.

Proof. To show that those systems are equivalent, we only need to show that $\mathbf{L}^1$ is closed under NEC. Suppose that $\vDash_{\mathbf{L}^1} \varphi$. We need to show that $\vDash_{\mathbf{L}^1} \Box\varphi$. From the fact that φ is an $\mathbf{L}^1$-tautology, follows that under all 1st level $\mathbf{L}^1$-valuations v, $v(\varphi) = \mathtt{T}$. Recall that $\mathtt{T}$ is a super-designated value, so $v(\Box\varphi) \in \mathtt{D}$. It is sufficient to show that $v(\Box\varphi) \in \mathtt{SD}$. Suppose for contradiction that $v(\Box\varphi) \in \mathtt{D}$ but $v(\Box\varphi) \notin \mathtt{SD}$. Consider an instance of $\vDash^1_{\mathbf{L}}$-tautology, $\Box\varphi \to \Box\Box\varphi$. According to v, $v(\Box\varphi \to \Box\Box\varphi) \notin \mathtt{D}$ because $v(\Box\varphi)$ is not super-designated, so the antecedent of the implication is designated and the consequent is not. This leads to a contradiction, and we can infer that $v(\Box\varphi) \in \mathtt{SD}$, which ultimately gives us that $\vDash_{\mathbf{L}^1} \Box\varphi$, so $\mathbf{L}^1$ is closed under NEC.

The finiteness of the m-th level valuations hierarchy is essential for two reasons. The first reason is that this semantics is not effective despite the logic being decidable. Knowing that the validity of $\Box\varphi \to \Box\Box\varphi$ reduces the height of the hierarchy to one, can be used to simplify the search for an effective procedure of deciding whether something is a tautology for this logic.

The second reason has do to with making the rule of NEC modular. What we actually doing while we move up within the hierarchy, we make sure that all tautologies of the previous levels get only the super-designated values. The valuations that do not do that are simply removed. So at each level we regain the partial rule of NEC that says if φ is a tautology of the previous level, $\Box\varphi$ is a tautology at the current level. Ultimately this shows that the standard axiomatization of the logic **S4** is not a minimal one. One could simply restrict the rule of necessitation to the following: if φ is provable without NEC, $\Box\varphi$ is provable. The resulting logic, by our theorem is equivalent to **S4**.

4 Omori-Skurt hypothesis

As we already have mentioned the critical property for reducing the height of the level valuations is the validity of 4. In [34] in Remark 41 the authors say:

> We have established that we do not need the whole hierarchy of valuations for **S4** and **S5**. One of the obvious properties of **S4** and **S5** is that there are only finitely many iterated modalities, and this is not the case in other modal logics we handle in this paper. There might be a deeper relation between the iterated modalities and the 'height' of the hierarchy, but we will leave this topic for further investigation.

The passage suggests that there is a relationship between the finiteness of the hierarchy and the number of non-equivalent modalities in a given logic. Based on this, if the hierarchy needed for NEC is finite, then the logic can only distinguish between finitely many non-equivalent modalities. It is natural to wonder whether the converse is also true, namely, if a logic only requires a finite hierarchy to regain NEC, then it has finitely many non-equivalent modalities.

Our main theorem shows that the remark from [34] is not correct. Logic $\mathbf{H4}^-$ is a sublogic of $\mathbf{K4}^-$, and the latter has infinitely many non-equivalent modalities. This shows that the relationship between the finiteness of the hierarchy and the number of non-equivalent modalities in a given logic is not as simple as suggested in the passage.

Fact 1. *Modal logic* **K4** *(in our notation* $\mathbf{K4}^+$*) has infinitely many non-equivalent modalities.*

Proof. See [19].

The reverse implication is also false. To see that we will define a non-deterministic semantics for logic $\mathbf{K5}^-$ and we will show that $\mathbf{K5}^-$ has finitely many non-equivalent modalities.

We start with the theoremhood relation of $\vdash_{\mathbf{K5}^-}$ which is given by the following axioms and rules:

Definition 7. By $\vdash_{\mathbf{K5}^-}$ we mean theorem-hood relation given by the following axioms and rules:

1. Propositional tautologies.
2. Axiom K, $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
3. Axiom 5, $\Diamond\varphi \to \Box\Diamond\varphi$.
4. Dual axioms $\Box\varphi \to \neg\Diamond\neg\varphi, \neg\Diamond\neg\varphi \to \Box\varphi, \neg\Box\neg\varphi \to \Diamond\varphi, \Diamond\varphi \to \neg\Box\neg\varphi$.
5. *Modus ponens*: if $\varphi \to \psi$ and φ, then ψ.

As in previous cases we use $\vdash^{\mathtt{NEC}}_{\mathbf{K5}^-}$ we denote the closure of the theoremhood relation under the rule of necessitation. The next step is to define the non-deterministic semantics sound and complete for $\mathbf{K5}^-$. Since in this case the axiomatization uses both modal operators, we will formulate the nmatrix in the language $\mathcal{L}_{\Box\Diamond}$. This also explains why we axiomatized $\mathbf{K5}^-$ by explicitly stating the Dual axioms.

Definition 8 ($\mathbf{K5}^-$). $\mathtt{M}_{\mathbf{K5}^-} = (\mathtt{Val}_{\mathbf{K5}^-}, \mathtt{D}_{\mathbf{K5}^-}, \mathtt{O}_{\mathbf{K5}^-})$ is an nmatrix where:

- $\mathtt{Val}_{\mathbf{K5}^-} = \{\mathtt{T}_\diamond, \mathtt{T}, \mathtt{t}_\diamond, \mathtt{t}, \mathtt{f}, \mathtt{f}_\diamond, \mathtt{F}, \mathtt{F}_\diamond\}$.
- $\mathtt{D}_{\mathbf{K5}^-} = \{\mathtt{T}_\diamond, \mathtt{T}, \mathtt{t}_\diamond, \mathtt{t}\}$.
- $\mathtt{O}_{\mathbf{K5}^-}$ is given by:

φ	$\overline{\neg}\varphi$	$\overline{\Box}\varphi$	$\overline{\Diamond}\varphi$
$\mathtt{T}_\diamond$	$\mathtt{f}$	$\mathtt{D}$	$\{\mathtt{T}_\diamond, \mathtt{T}\}$
$\mathtt{T}$	$\mathtt{F}$	$\mathtt{D}$	$\overline{\mathtt{D}}$
$\mathtt{t}_\diamond$	$\mathtt{f}_\diamond$	$\overline{\mathtt{D}}$	$\{\mathtt{T}_\diamond, \mathtt{T}\}$
$\mathtt{t}$	$\mathtt{F}_\diamond$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
$\mathtt{F}_\diamond$	$\mathtt{t}$	$\mathtt{D}$	$\{\mathtt{T}_\diamond, \mathtt{T}\}$
$\mathtt{F}$	$\mathtt{T}$	$\mathtt{D}$	$\overline{\mathtt{D}}$
$\mathtt{f}_\diamond$	$\mathtt{t}_\diamond$	$\overline{\mathtt{D}}$	$\{\mathtt{T}_\diamond, \mathtt{T}\}$
$\mathtt{f}$	$\mathtt{T}_\diamond$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$

Table 2: Table for unary connectives

and the implication is given by:

$\overline{\to}$	$\mathtt{T}_\diamond$	$\mathtt{T}$	$\mathtt{t}_\diamond$	$\mathtt{t}$	$\mathtt{F}_\diamond$	$\mathtt{F}$	$\mathtt{f}_\diamond$	$\mathtt{f}$
$\mathtt{T}_\diamond$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\{\mathtt{f}_\diamond, \mathtt{f}\}$	$\{\mathtt{f}_\diamond, \mathtt{f}\}$
$\mathtt{T}$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\{\mathtt{f}_\diamond, \mathtt{f}\}$	$\{\mathtt{f}_\diamond, \mathtt{f}\}$
$\mathtt{t}_\diamond$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
$\mathtt{t}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$	$\overline{\mathtt{D}}$
$\mathtt{F}_\diamond$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$
$\mathtt{F}$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\mathtt{D}$	$\mathtt{D}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$	$\{\mathtt{t}_\diamond, \mathtt{t}\}$
$\mathtt{f}_\diamond$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$
$\mathtt{f}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$	$\mathtt{D}$

Table 3: Table for implication

Theorem 4. $\mathtt{M}_{\mathbf{K5}^-}$ *is sound and complete with respect to* $\vdash_{\mathbf{K5}^-}$.

Proof. Soundness is proved by the induction. We need to show that all axioms are valid in $\mathtt{M}_{\mathbf{K5}^-}$ and that the nmatrix is closed under the rule of *modus ponens*. This is straightforward. For completeness, we start with the following lemma:

Lemma 2 ($\mathbf{K5}^-$ valuation Lemma)**.** *Let Γ be a $\mathbf{K5}^-$-relatively maximal set. The following function is a $v_{\mathbf{K5}^-}$-valuation:*

$$v_\Gamma(\varphi) = \begin{cases} \mathtt{T}_\Diamond & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \varphi, \Gamma \vdash_{\mathbf{K5}^-} \Diamond\varphi \\ \mathtt{T} & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\Diamond\varphi \\ \mathtt{t}_\Diamond & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \varphi, \Gamma \vdash_{\mathbf{K5}^-} \Diamond\varphi \\ \mathtt{t} & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\Diamond\varphi \\ \mathtt{F} & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\Diamond\varphi \\ \mathtt{f}_\Diamond & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\varphi, \Gamma \vdash_{\mathbf{K5}^-} \Diamond\varphi \\ \mathtt{F}_\Diamond & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\varphi, \Gamma \vdash_{\mathbf{K5}^-} \Diamond\varphi \\ \mathtt{f} & \textit{if } \Gamma \vdash_{\mathbf{K5}^-} \neg\Box\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\varphi, \Gamma \vdash_{\mathbf{K5}^-} \neg\Diamond\varphi \end{cases}$$

Proof. Similarly to the previous case, observe that this function is well-defined. The proof is by induction on complexity. The base case is straightforward. Assume that the lemma works for φ, ψ. We use $\Gamma \vdash_{\mathbf{K5}^-} \{\varphi_1, \varphi_2, \varphi_3\}$ to mean that $\Gamma \vdash_{\mathbf{K5}^-} \varphi_1$, and $\Gamma \vdash_{\mathbf{K5}^-} \varphi_2$, and $\Gamma \vdash_{\mathbf{K5}^-} \varphi_3$. Then we can use the following table to summarizes all the cases:

- **Negation**:

Assumption	Unpacked	Wts	Unpacked	Justification
$v_\Gamma(\varphi) = \mathtt{T}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\varphi, \varphi, \Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{f}$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\neg\varphi, \neg\neg\varphi, \neg\Diamond\neg\varphi\}$	Dual; DN; Dual
$v_\Gamma(\varphi) = \mathtt{T}$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\varphi, \varphi, \neg\Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{F}$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\neg\varphi, \neg\neg\varphi, \neg\Diamond\neg\varphi\}$	Dual, DN; DN; Dual
$v_\Gamma(\varphi) = \mathtt{t}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\varphi, \varphi, \Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{f}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\neg\varphi, \neg\neg\varphi, \Diamond\neg\varphi\}$	Dual;DN ; Dual
$v_\Gamma(\varphi) = \mathtt{t}$,	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\varphi, \varphi, \neg\Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{F}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\neg\varphi, \neg\neg\varphi, \Diamond\neg\varphi\}$	Dual, DN; DN; Dual, DN
$v_\Gamma(\varphi) = \mathtt{F}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\varphi, \neg\varphi, \Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{t}$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\neg\varphi, \neg\varphi, \neg\Diamond\neg\varphi\}$	Dual; Dual
$v_\Gamma(\varphi) = \mathtt{F}$,	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\varphi, \neg\varphi, \neg\Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{T}$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\neg\varphi, \neg\varphi, \neg\Diamond\neg\varphi\}$	Dual; Dual
$v_\Gamma(\varphi) = \mathtt{f}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\varphi, \neg\varphi, \Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{t}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\neg\varphi, \neg\varphi, \Diamond\neg\varphi\}$	Dual; Dual
$v_\Gamma(\varphi) = \mathtt{f}$,	$\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\varphi, \neg\varphi, \neg\Diamond\varphi\}$	$v_\Gamma(\neg\varphi) = \mathtt{T}_\Diamond$	$\Gamma \vdash_{\mathbf{K5}^-} \{\Box\neg\varphi, \neg\varphi, \Diamond\neg\varphi\}$	Dual; Dual

Table 4: The case of negation

  To clarify the notation we will run the first case from the table explicitly. By assumption we have $v_\Gamma(\varphi) = \mathtt{T}_\Diamond$. After unpacking we get $\Gamma \vdash_{\mathbf{K5}^-} \Box\varphi$, $\Gamma \vdash_{\mathbf{K5}^-} \varphi$, and $\Gamma \vdash_{\mathbf{K5}^-} \Diamond\varphi$. We need to show that $\Gamma \vdash_{\mathbf{K5}^-} \{\neg\Box\neg\varphi, \neg\neg\varphi, \neg\Diamond\neg\varphi\}$. We get it by the law of double negation and Dual axioms in a straightforward manner.

- **Implication**: We need to go through 16 cases. For the positive cases (where the conditional's value is exactly $\{\mathtt{t}_\Diamond, \mathtt{t}\}$) we need to show that $\Gamma \vdash_{\mathbf{K5}^-}$

$\neg\Box(\varphi \to \psi)$, and $\Gamma \vdash_{\mathbf{K5^-}} \varphi \to \psi$. For negative cases (where the conditional's value is exactly $\{\mathtt{f}_\Diamond, \mathtt{f}\}$) we need to show that $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box(\varphi \to \psi)$, and $\Gamma \vdash_{\mathbf{K5^-}} \neg(\varphi \to \psi)$. Let us start with the negative cases and assume that $v_\Gamma(\varphi) \in \{\mathtt{T}_\Diamond, \mathtt{T}\}$ and $v_\Gamma(\psi) \in \{\mathtt{f}_\Diamond, \mathtt{f}\}$. After unpacking the definitions we get $\Gamma \vdash_{\mathbf{K5^-}} \Box\varphi$ and $\Gamma \vdash_{\mathbf{K5^-}} \varphi$. For ψ, we know that $\Gamma \vdash_{\mathbf{K5^-}} \neg\psi$ and $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box\psi$. By classical propositional logic we get $\Gamma \vdash_{\mathbf{K5^-}} \neg(\varphi \to \psi)$ and $\Gamma \vdash_{\mathbf{K5^-}} \neg(\Box\varphi \to \Box\psi)$, which together with the axiom K by *modus tollens* results in $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box(\varphi \to \psi)$. The positive cases we will split into two sub-cases. Let us start with the case where $v_\Gamma(\varphi) \in \{\mathtt{T}_\Diamond, \mathtt{T}\}$ and $v_\Gamma(\psi) \in \{\mathtt{t}_\Diamond, \mathtt{t}\}$. Hence $\Gamma \vdash_{\mathbf{K5^-}} \Box\varphi$, $\Gamma \vdash_{\mathbf{K5^-}} \varphi$, $\Gamma \vdash_{\mathbf{K5^-}} \psi$, and $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box\psi$. We get by propositional logic $\Gamma \vdash_{\mathbf{K5^-}} \varphi \to \psi$ and as in the previous case we get $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box(\varphi \to \psi)$ by the same reasoning. The last case is $v_\Gamma(\varphi) \in \{\mathtt{F}_\Diamond, \mathtt{F}\}$ and $v_\Gamma(\psi) \in \{\mathtt{t}_\Diamond, \mathtt{t}\}$. So, $\Gamma \vdash_{\mathbf{K5^-}} \neg\varphi$, $\Gamma \vdash_{\mathbf{K5^-}} \Box\varphi$, $\Gamma \vdash_{\mathbf{K5^-}} \psi$, and $\Gamma \vdash_{\mathbf{K5^-}} \neg\Box\psi$ and the reasoning is similar to the previous two cases.

- $\Box$ **modality**. This is straightforward.
- $\Diamond$ **modality**. We focus on four problematic cases, namely cases where we have $v_\Gamma(\Diamond\varphi) \in \{\mathtt{T}_\Diamond, \mathtt{t}_\Diamond, \mathtt{f}_\Diamond, \mathtt{F}_\Diamond\}$. For each of them we need to show that $\Gamma \vdash_{\mathbf{K5^-}} \Diamond\varphi$ and $\Gamma \vdash_{\mathbf{K5^-}} \Box\Diamond\varphi$. From the assumption it follows that $\Gamma \vdash_{\mathbf{K5^-}} \Diamond\varphi$. Together with the validity of the axiom $\mathtt{5}$, by *modus ponens* we get $\Gamma \vdash_{\mathbf{K5^-}} \Box\Diamond\varphi$.

This ends the proof of the lemma.

$\Box$

The rest of the proof follows the standard scheme for proving completeness.

Theorem 5 ($\mathbf{K5^+}$ completeness)**.** $\vdash^{\mathtt{NEC}}_{\mathbf{K5}} \varphi$ *iff* $\vDash_{\mathtt{M}_{\mathbf{K5^-}}} \varphi$.

Clearly by taking into account only $\mathbf{K5^+}$-valuations we regain the rule of necessitation. Yet, it is impossible to reduce the height of the hierarchy for the following reason. Consider a sequence of formulas $p \vee \neg p, \Box(p \vee \neg p), \Box\Box(p \vee \neg p), \ldots$. The first formula in the sequence is a 0th level $\mathbf{K5^-}$-tautology and kth formula in the sequence is a $\mathbf{K5^{k+1}}$-tautology. One look at the truth-table for modality makes it clear that this sequence needs an infinite number of levels. On the other hand, the number of non-equivalent modalities in $\mathbf{K5^+}$ is finite as proven in [20] as theorem 4.23 on page 150. This means that the reverse implication is false.

5 Conclusions

In this paper, we have generalized one of the theorems of [34]. We have shown that the key property needed to prove this theorem is the validity of axiom $\mathtt{4}$. We have also demonstrated that the relation between the height of the hierarchy needed for NEC is not related to the number of non-equivalent modalities in a given logic. Future work in this area could focus on finding the necessary and sufficient conditions under which the hierarchy becomes finite. Our conjecture is that the height of the valuation depends directly on the validity of axioms of the form $\underbrace{\Box \dots \Box\varphi}_{n} \to \varphi$, which would be a natural generalization of our main theorem.

References

[1] Arnon Avron. “5-valued non-deterministic semantics for the basic paraconsistent logic mCi.” In: *Studies in Logic, Grammar and Rhetoric* 14 (2008), pp. 127–136.

[2] Arnon Avron. “A non-deterministic view on non-classical negations.” In: *Studia Logica* 80.2-3 (2005), pp. 159–194. DOI: 10.1007/s11225-005-8468-5.

[3] Arnon Avron. “Logical Non-determinism as a Tool for Logical Modularity: An Introduction.” In: *We Will Show Them!(1)*. 2005, pp. 105–124.

[4] Arnon Avron. “Modular semantics for some basic logics of formal inconsistency.” In: *The many sides of logic* 21 (2009), pp. 15–26.

[5] Arnon Avron. “Non-deterministic matrices and modular semantics of rules.” In: *Logica universalis*. Springer, 2005, pp. 149–167.

[6] Arnon Avron. “Non-deterministic semantics for families of paraconsistent logics.” In: *Handbook of Paraconsistency* 9 (2007), pp. 285–320.

[7] Arnon Avron. “Non-deterministic semantics for logics with a consistency operator.” In: *International Journal of Approximate Reasoning* 45.2 (2007), pp. 271–287. DOI: 10.1016/j.ijar.2006.06.011.

[8] Arnon Avron. “Non-deterministic semantics for paraconsistent C-systems.” In: *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer. 2005, pp. 625–637. DOI: 10.1007/11518655_53.

[9] Arnon Avron, Jonathan Ben-Naim, and Beata Konikowska. “Cut-Free Ordinary Sequent Calculi for Logics Having Generalized Finite-Valued Semantics.” In: *Logica Universalis* 1.1 (Jan. 2007), pp. 41–70. DOI: 10.1007/s11787-006-0003-6.

[10] Arnon Avron and Beata Konikowska. “Multi-valued calculi for logics based on non-determinism.” In: *Logic Journal of IGPL* 13.4 (2005), pp. 365–387. DOI: 10.1093/jigpal/jzi030.

[11] Arnon Avron, Beata Konikowska, and Anna Zamansky. “Cut-free sequent calculi for C-systems with generalized finite-valued semantics.” In: *Journal of Logic and Computation* 23.3 (2013), pp. 517–540.

[12] Arnon Avron and Iddo Lev. “Canonical propositional Gentzen-type systems.” In: *Automated Reasoning: First International Joint Conference, IJCAR 2001 Siena, Italy, June 18-23, 2001 Proceedings (Lecture Notes in Computer Science)*. Springer-Verlag, 2001, pp. 529–544. DOI: 10.1007/3-540-45744-5_45.

[13] Arnon Avron and Iddo Lev. “Non-deterministic matrices.” In: *Proceedings. 34th International Symposium on Multiple-Valued Logic*. IEEE. 2004, pp. 282–287.

[14] Arnon Avron and Iddo Lev. "Non-deterministic Multiple-valued Structures." In: *Journal of Logic and Computation* 15.3 (2005), pp. 241–261. DOI: 10.1093/logcom/exi001. eprint: http://logcom.oxfordjournals.org/content/15/3/241.full.pdf+html. URL: http://logcom.oxfordjournals.org/content/15/3/241.abstract.

[15] Arnon Avron and Anna Zamansky. "Many-Valued Non-deterministic Semantics for First-Order Logics of Formal (In)consistency." In: *Algebraic and Proof-theoretic Aspects of Non-classical Logics: Papers in Honor of Daniele Mundici on the Occasion of His 60th birthday*. Ed. by Stefano Aguzzoli et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 1–24. ISBN: 978-3-540-75939-3. DOI: 10.1007/978-3-540-75939-3_1. URL: https://doi.org/10.1007/978-3-540-75939-3_1.

[16] Arnon Avron and Anna Zamansky. "Non-deterministic multi-valued logics-A tutorial." In: *2010 40th IEEE International Symposium on Multiple-Valued Logic*. IEEE. 2010, pp. 53–60. DOI: 10.1109/ismvl.2010.18.

[17] Arnon Avron and Anna Zamansky. "Non-deterministic multi-valued matrices for first-order logics of formal inconsistency." In: *37th International Symposium on Multiple-Valued Logic (ISMVL'07)*. IEEE. 2007, pp. 14–14. DOI: 10.1109/ismvl.2007.38.

[18] Arnon Avron and Yoni Zohar. "Non-deterministic matrices in action: Expansions, refinements, and rexpansions." In: *2017 IEEE 47th international symposium on multiple-valued logic (ISMVL)*. IEEE. 2017, pp. 118–123. DOI: 10.1109/ismvl.2017.16.

[19] Fabio Bellissima. "Infinite sets of nonequivalent modalities." In: *Notre Dame Journal of Formal Logic* 30.4 (Sept. 1989). DOI: 10.1305/ndjfl/1093635240.

[20] Brian Chellas. *Modal logic: an introduction*. Cambridge England New York: Cambridge University Press, 1980. ISBN: 0521224764. DOI: 10.1017/cbo9780511621192.

[21] Marcelo E Coniglio, Fariñas Del Cerro Luis, and Marques Peron Newton. "Modal logic with non-deterministic semantics: Part I—Propositional case." In: *Logic Journal of the IGPL* (Oct. 2019). jzz027. ISSN: 1367-0751. DOI: 10.1093/jigpal/jzz027. eprint: https://academic.oup.com/jigpal/advance-article-pdf/doi/10.1093/jigpal/jzz027/30101966/jzz027.pdf. URL: https://doi.org/10.1093/jigpal/jzz027.

[22] Marcelo E Coniglio, Fariñas Del Cerro Luis, and Marques Peron Newton. "Modal Logic With Non-Deterministic Semantics: Part II—Quantified Case." In: *Logic Journal of the IGPL* (June 2021). jzab020. ISSN: 1367-0751. DOI: 10.1093/jigpal/jzab020. eprint: https://academic.oup.com/jigpal/advance-article-pdf/doi/10.1093/jigpal/jzab020/38609545/jzab020.pdf. URL: https://doi.org/10.1093/jigpal/jzab020.

[23] Lukas Grätz. *Truth Tables for Modal Logics T and S4, by using Three-Valued Non-Deterministic Level Semantics.* Accepted for publication. Submitted. 2021.

[24] Yuriy Ivlev. "A semantics for modal calculi." In: *Bulletin of the Section of Logic* 17.3/4 (1988), pp. 114–121.

[25] Yuriy Ivlev. "Generalization of Kalmar's method for quasi-matrix logic." In: *Логические исследования* 19.1 (2013).

[26] Yuriy Ivlev. "Деонтическая квазиматричная логика. Логика норм (Deontic quasi-matrix logic)." In: *Гуманитарный вестник* 7 (69) (2018).

[27] Yuriy Ivlev. "Основные области приложения квазиматричной логики (Quasi-matrix logic)." In: *Логические исследования* 9 (2002).

[28] Yuriy Ivlev. "Основы логической теории аргументации (Fundamentals of the logical theory of argumentation)." In: *Логические исследования* 10 (2003).

[29] Yuriy Ivlev. "Табличное построение пропозициональной модальной логики (Tabular construction of the propositional modal logic)." In: *Вестник Московского университета. Серия «Философия* 6 (1973), pp. 51–61.

[30] Yury Ivlev. "Квазиматричная логика-основа теории фактических (физических) модальностей (Quasi-matrix logic - the basis of theories of actual (physical) modality." In: *Логические исследования* 8 (2001).

[31] John T Kearns. "Modal semantics without possible worlds." In: *The Journal of Symbolic Logic* 46.1 (1981), pp. 77–86. DOI: 10.2307/2273259.

[32] Ori Lahav and Yoni Zohar. "Effective Semantics for the Modal Logics K and KT via Non-deterministic Matrices." In: *Automated Reasoning.* Ed. by Jasmin Blanchette, Laura Kovács, and Dirk Pattinson. Cham: Springer International Publishing, 2022, pp. 468–485. ISBN: 978-3-031-10769-6.

[33] Hitoshi Omori. "A Note on Ciuciura's mbC1." In: *Bulletin of the Section of Logic* 48.3 (2019), pp. 161–171. DOI: 10.18778/0138-0680.48.3.01.

[34] Hitoshi Omori and Daniel Skurt. "More modal semantics without possible worlds." In: *IfCoLog Journal of Logics and their Applications* 3.5 (2016), pp. 815–845.

[35] Pawel Pawlowski and Elio La Rosa. "Modular non-deterministic semantics for T, TB, S4, S5 and more." In: *Journal Of Logic And Computation* (2021).

[36] Pawel Pawlowski and Elio La Rosa. "Yes Folks, Well-known Modal Logics Are at Most 8-valued." In: *Under review* (Under review).

[37] K. Reach. "Review of O. V. Zich. Výrokový počet s komplexními hodnotami (Sentential calculus with complex values)." In: *Journal of Symbolic Logic* 4.4 (Dec. 1939), pp. 165–166. DOI: 10.2307/2268724.

[38] Nicholas Rescher. "Quasi-truth-functional systems of propositional logic." In: *The Journal of Symbolic Logic* 27.1 (1962), pp. 1–10. DOI: 10.2307/2963674.

[39] Anna Zamansky and Arnon Avron. "Cut-elimination and quantification in canonical systems." In: *Studia Logica* 82.1 (2006), pp. 157–176. DOI: 10.1007/s11225-006-6611-6.

[40] Anna Zamansky and Arnon Avron. "Non-Deterministic Semantics for First-Order Paraconsistent Logics." In: *KR.* 2006, pp. 431–439.

[41] Otakar V. Zich. "Výrokový počet s komplexními hodnotami." In: *Česká mysl* 34 (1938), pp. 189–196.

Received 24 January 2022

Optical Microscopes Reveal the Hyperreal World

Riccardo Dossena
Liceo Scientifico G. Novello, 26845 Codogno and Computational Philosophy Lab, University of Pavia, 27100 Pavia, Italy
riccardo.dossena@unipv.it

Lorenzo Magnani
Department of Humanities, Philosophy Section and Computational Philosophy Lab, University of Pavia, 27100 Pavia, Italy
lmagnani@unipv.it

Abstract

Many fundamental concepts of the calculus are difficult to grasp, and they may appear epistemologically unjustified. Diagrams allow us to overcome the difficulty in constructing representations of mathematical critical situations and objects. In this article we describe some examples of optical diagrams as a particular kind of what we call *epistemic mediators* able to perform the explanatory abductive task of providing a better understanding of the calculus, through the nonstandard analysis introduced by Abraham Robinson. In particular, we will describe the infinitesimal microscopes that make possible to visualize what the graph of a function looks like in an infinitesimal neighborhood of one of its points and, if desired, to develop consequent reasoning. We will propose a rigorous mathematical formalization of infinitesimal microscopes and show their proper, although more informal, use in teaching situations.

1 Introduction

Nonstandard analysis is a mathematical theory that can reformulate differential and integral calculus through the use of infinitesimal and infinite numbers. It was introduced in the 1960s by Abraham Robinson [18], thanks to the results of mathematical

Some themes of this article are excerpted from the article by R. Dossena *Il mondo iperreale attraverso i microscopi ottici*, published by Matematicamente.it in 2017 (in Italian).

logic, particularly model theory, which enabled the construction of an "extended" field of the real numbers. The techniques of nonstandard analysis allow the same results as classical analysis to be achieved, and can also be taught to beginning students.

Infinitesimal microscopes and infinite telescopes are graphical expedients that play a prominent role in nonstandard analysis-based education, as evidenced by H. J. Keisler's *Elementary Calculus* manuals [11], the volumes in the *Il professor Apotema insegna...* series by Goldoni, in particular [6, 7, 8], and the textbook by B. Stecca and D. Zambelli [19] (inspired by Professor Apotema's books). Unlike Keisler's text, in which the use of optical instruments is aimed primarily at understanding concepts, Goldoni's books also have several applications aimed at discovering solving methods. If in [11] a rather informal discussion is carried out, in [12] definitions are presented more rigorous, originally suggested by K. D. Stroyan (see handbook [20]). Later, D. Tall [21] proposed a slight modification of the definitions, which proved useful in fruitfully completing the formalization process. In the work of L. Magnani and R. Dossena [16, 3], continued in the same vein – also synthetical summarized in the following section of this article – some epistemological aspects were explored in depth, with particular reference to the classification of optical instruments as *epistemic mediators* in the context of *manipulative abduction*, and further applications were technically developed (see, for example, the articles by J. Baire and V. Henry [1, 2]).

A by-product of this intervention is also to offer the teacher leaning toward a nonstandard teaching approach the opportunity to exploit the potential of optical instruments with greater awareness, which we believe can be gained through the analysis of precise definitions and results. The work we propose does not plan to be brought back to a classroom as it is, but should remain, so to speak, "behind the scenes": didactic optical instruments may certainly be presented intuitively and informally, but it is up to the teacher to keep the intuition on the track of correctness. In what follows we will limit ourselves to the description of infinitesimal microscopes and assume that the reader is familiar with the fundamental properties of hyperreal numbers.

2 Mathematical Diagrams: Their Explanatory and Abductive Roles

The term "abduction" was created more than a century ago by the eminent American philosopher Charles Sanders Peirce to describe the cognitive activities that include the production and assessment of explanatory hypotheses. We argue that abductive reasoning accounts for a significant portion of scientific thinking, and that model-based and manipulative abduction can explain some of the functions that diagrams play in mathematical reasoning. Moreover, mathematical and geometrical reasoning, according to Peirce, "consists in constructing a diagram according to a general precept, in observing certain relations between parts of that diagram not explicitly required by the precept, showing that these relations will hold for all such diagrams, and in formulating this conclusion in general terms. All valid necessary reasoning is in fact thus diagrammatic" [17, 1.54].

What exactly is abduction? Abduction is the process of reasoning in which explanatory hypotheses are produced and assessed. It is the process of inferring certain facts and/or laws and hypotheses that render some sentences reasonable and that explain or uncover some (eventually new) phenomena or observations. In [13] it is contended that the term abduction has two basic epistemological meanings: (1) abduction that only creates "plausible" hypotheses ("selected" or "creative") and (2) abduction as inference "to the best explanation", which also evaluates hypotheses. To better clarify the distinction between selective and creative abduction, for example in the field of medicine, the discovery of a new illness and the symptoms it produces might be seen as the consequence of a creative abductive inference. As a result, "creative" abduction encompasses the entire field of scientific knowledge expansion. In medical diagnosis, however, this is unimportant because the aim is to "select" from a collection of pre-stored diagnostic items. Both conclusions – selective and creative – can be classified as ampliative, since the reasoning involved in both circumstances amplifies, or goes beyond, the information included in the premises.

It is worth noting that abductive thinking is frequently linked – so to speak – to the exploitation of the environment, thus it is not only the result of internal cognitive human endowments. In these circumstances, we are dealing with a type of "hypothesizing through doing", in which new and unspoken knowledge gets obtained through manipulations of external objects (known as *epistemic mediators*).

In turn, the notion of the so-called *manipulative abduction* encompasses a substantial part of scientists' reasoning in which action plays a central role and the consequences of that action are implicit and difficult to elicit. For instance, action can give otherwise inaccessible knowledge that allows the agent to solve issues by initiating and carrying out a proper abductive process of hypothesis generation or selection: the act of manipulating diagrams in mathematics is a paradigmatic form of manipulative abduction. Indeed, in the case of manipulative abduction, we face a kind of online thinking: it is a real-world example of distributed cognition in action. The concept of manipulative abduction, which also considers the external aspect of abductive reasoning following an eco-cognitive view, describes, for example, a significant portion of scientific thinking in which the role of action and of externalized models (such as diagrams and artifacts) and various tools is fundamental and in which the characteristics of this action are implicit and difficult to be extracted. By initiating and carrying through an appropriate abductive process of production and/or selection of hypotheses, action can provide otherwise unavailable knowledge that helps the agent to solve problems of various kinds. As we said above, manipulative abduction occurs when we are "thinking through doing" and not only, in a pragmatic sense, about doing [15, chapter one]. We face an abductive/adaptive process produced in the dynamical inner/outer coupling where internal elements are mixed with external cognitive delegations [23].

All inference, Peirce further observes, is a type of sign activity, where the term sign comprises "feeling, image, conception, and other representation" [17, 5.283], and, in a Kantian lexicon, all synthetic forms of cognition. That is, visualizations, diagrams, icons, simulations, analogies, and many other non-propositional components of cognition, both internal and external, account for a significant portion of cognitive effort that can consequently be called model-based. Of course, model-based reasoning takes on a new creative significance when it is incorporated in abductive processes, allowing us to identify what has been called "model-based abduction".

Peirce provides an intriguing example of a basic model-based abduction including sense activity and the manipulation of the environment: "A man can distinguish different textures of cloth by feeling: but not immediately, for he requires to move fingers over the cloth, which shows that he is obliged to compare sensations of one instant with those of another" [17, 5.221]. To summarize, manipulative abduction occurs when we think "through" doing rather than just thinking "about" doing in a pragmatic way, such as for example in the case of planning.

It is worth noting that model-based and manipulative abductions are both prevalent in mathematics. Geometrical constructions, for example, provide situations that are curious and "at the limit". These are intrinsically dynamic and artificial, and provide a variety of contingent epistemic acting options, such as looking at things from different angles, comparing subsequent appearances, dismissing, choosing, reordering, and assessing. Furthermore, they exhibit some of the characteristics listed below, which are typical of so-called abductive epistemic mediators [13]: task simplification and the ability to get visual information that would otherwise be unavailable.

Epistemic mediators have a number of unique characteristics (the first three, for example, may be found in geometrical constructions):

(1) action can provide a simplification of the cognitive task and a redistribution of effort over time, when we favor the manipulation of concrete entities in order to grasp the knowledge of structures which are otherwise too abstract, or when we face excessive and unmanageable information;

(2) in the context of incomplete or inconsistent information – not just from a "perceptual" standpoint – or a decreased capacity to act upon the world, specific actions can be useful: they are employed to get new data in order to restore coherence and enhance insufficient knowledge;

(3) action allows us to create exterior artifactual models of task processes instead of the corresponding internal ones, which are appropriate to adjust the environment to the demands of a cognitive agent;

(4) action as management of sense-data shows how we may manipulate the location of our bodies (and/or external objects) and how we can use various types of prostheses (technology devices and interfaces) to obtain many new types of stimulation: action gives tactile and visual information that might otherwise be inaccessible (surgery is a good example in this case).

Because they may be modified, diagrams play a significant part in abduction. In mathematics, diagrams serve a variety of functions in a common abductive manner and two of them are crucial: they give an intuitive and mathematical explanation that might aid comprehension of concepts that are difficult to grasp or that look unclear and/or unjustified epistemologically. However, many external representations, both in terms of diagrams and symbols, are used in the creation of mathematical concepts. Microscopes and "microscopes pointed in microscopes" (that look at infinitesimally small details), telescopes (that look at infinity), windows (that look at a specific situation), play a mirror role (to represent externally crude mental internal models), but can also play an unveiling role (to help create new and interesting mathematical concepts, theories, and structures). These are all examples of diagrams that play an optical role.

Optical diagrams can also offer a critical explanatory (and didactic) function in reducing impediments and obscurities, as well as improving mathematical understanding of critical circumstances. They make it easier to create new internal representations and attain new symbolic-propositional goals. The amazing relevance of optical diagrams in the field of calculus in the interaction of standard/nonstandard analysis will be stressed in the results discussed in the following sections. Some of them may also serve as unveiling diagrams, shedding fresh light on mathematical structures: it is possible to hypothesize that these diagrams may lead to more fascinating creative outcomes. The optical and unveiling diagrammatic representation of mathematical structures is in turn capable to trigger perceptual procedures (for instance, determining how a real function appears in its points and/or to infinity; determining how to truly achieve its limits).

We stated above that in mathematics diagrams play various roles in a typical abductive way. Now we can add that:

- they are epistemic mediators capable of performing a variety of abductive tasks in so far as
- they are external representations that are dedicated to generating explanatory abductive outcomes, as we will illustrate below in this article.

We will present in the following sections some kinds of diagrams (microscopes pointed in microscopes), which provide very suggestive mental representations of the concept of tangent line at the infinitesimally small regions. They help create new previously unknown concepts, as illustrated in the case of the discovery of the non-Euclidean geometry in [14].

3 The Real Line and Standard Microscopes

The points of a line, according to Hilbert's traditional axiomatic approach to Euclidean geometry, are in biunivocal correspondence with the set $\mathbb{R}$ of real numbers. Thus, we can talk about the *real line* whose points, although arranged in a "dense" manner, are well distinct. For example, the points corresponding to the numbers 2 and 2.1 (henceforth simply the points 2 and 2.1) occupy different positions on the line, in the same way as the points 3 and 3.0001. But what about their actual *visibility*? If we were to draw a straight line on a graph paper, assuming 1 cm as the unit of measure, we would clearly be able to distinguish the first pair of points with the naked eye, but the same would not be true for the second pair: the two points would, in fact, appear confusingly overlapping and appearing to be in the same position. In general terms, we can say that numbers such as a and $a + \frac{1}{n}$, with n "very large", although they correspond to different points on the line, are "too close" to be viewed separately. To overcome this we should use some sort of magnifying glass that "separates" the two numbers and allows us to show that they are indeed distinct. The situation can be represented (somewhat appealingly) as in Figure 1: within the circle representing the magnifying glass, the numbers 3 and 3.0001 appear distinct, but, on the other hand, it is no longer practically possible to visualize 2, which on this scale is too far away and goes out of the field of view (this circumstance is jokingly signaled by a sign).

The mathematical equivalent of such an instrument can be realized by means of the $\mu\colon \mathbb{R} \to \mathbb{R}$ transformation defined by

$$\mu(x) = \frac{x-a}{\frac{1}{n}} = n(x-a)$$

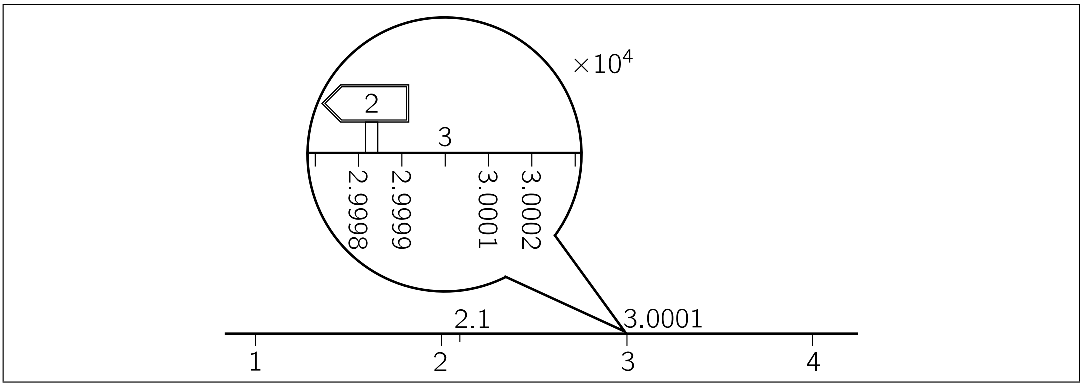


Figure 1: Through the lens the numbers 3 and 3.0001 can be seen well separated.

which we call a *standard microscope* (Figure 2). The μ function allows the view to be enlarged by means of the circle shown, which contains part of its range. The

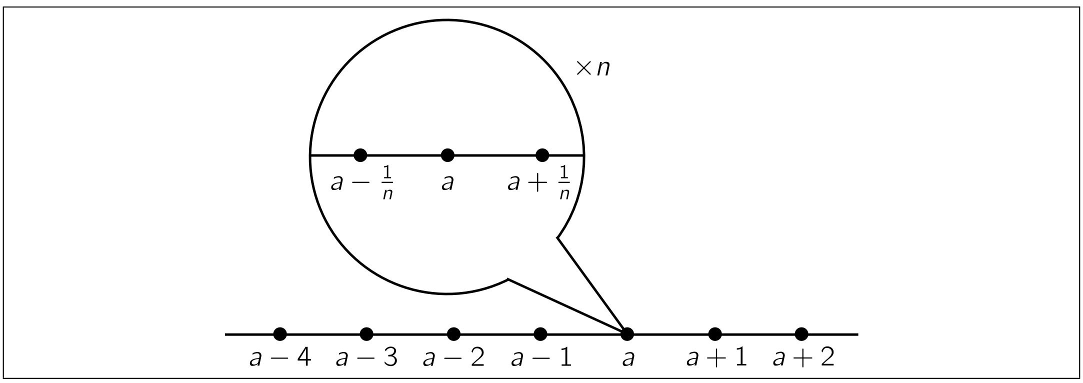


Figure 2: Standard microscope that "separates" a from $a + \frac{1}{n}$.

number n represents the *magnification factor* of the transformation, since it allows us to clearly distinguish a, $a + \frac{1}{n}$, $a - \frac{1}{n}$, $a + \frac{2}{n}$, etc. after matching them with the distinct numbers 0, 1, -1, 2, etc. The conclusive step consists in identifying each x with its image $\mu(x)$[1].

This same idea will enable us to make the infinite hyperreal world visible as well.

[1]This identification, as Tall [22] suggests, is analogous to what occurs in the making of a geographic map: the location of a place (e.g., "Rome") is denoted on the map by its name (Rome).

4 The Hyperreal Line and Infinitesimal Microscopes

The set $\mathbb{R}^*$ of hyperreal numbers, with the usual operations and relations, is a non-Archimedean ordered field containing thus, in addition to the real numbers, infinitely small numbers, i.e., ε numbers such that $0 < |\varepsilon| < \frac{1}{n}$ for every $n \in \mathbb{N}$, and infinitely large numbers, i.e., H numbers such that $|H| > n$ for every $n \in \mathbb{N}$[2]. It can be seen further that every finite hyperreal number (i.e., such that it is in absolute value less than some natural number) is always of the form $c+\varepsilon$, with $c \in \mathbb{R}$ and ε infinitesimal. Well, can our geometric intuition of the line correspond to this structure?

The fact that every real number can be assigned a point on the geometric line responds to the need to conceive of segments of assigned lengths. Conversely, the fact that real numbers exhaust all points on a line is instead a consequence of the unconditional acceptance of Hilbert's axiomatic approach, which influences and shapes our intuitive view of the line, but which could be assumed not so restrictively, at least as far as the Archimedean axiom is concerned. In fact, it would be similarly legitimate to think of a geometric straight line with more elements than just the points that correspond to real numbers (with "infinitely far" and "infinitely close" points) and to reserve Archimedes' axiom only for the points that correspond, as Leibniz put it, to "assignable" quantities (in fact, to standard real numbers). From now on, this will be the perspective from which we will look at the geometric straight line. In other words, the straight line, unless otherwise advised, will mean for us the *hyperreal straight line*.

If on the straight line we cannot, at the ordinary scale, distinguish the points 3 and $3 + 10^{-4}$, a fortiori we cannot distinguish 3 and $3 + \varepsilon$ (ε infinitesimal) since $|\varepsilon| < 10^{-4}$ (and indeed $|\varepsilon| < 10^{-n}$ for any n). The argument generalizes easily: on the hyperreal line numbers at infinitesimal distance can never appear distinct, even after magnification by any factor n, and all elements of the monad of a real number *are not viewable separately* from it, just as if the whole monad "collapsed" on it.

Ultimately, *on the straight line it is not possible to distinguish between the numbers that differ by an infinitesimal because they appear superimposed on the one real number to which they are infinitely close.* Consequently, since what can be seen (with or without standard enlargements) of any finite hyperreal number is nothing more than *its standard part*, the visual image of the hyperreal line turns out to be identical (at least as far as its finite part) to that of the real line. However, since $\mathbb{R}^*$ is an ordered field, if $\varepsilon > 0$ is an infinitesimal and c a real number, the points c, $c + \varepsilon$ e $c - \varepsilon$, albeit not visually, are distinct elements such that $c - \varepsilon < c < c + \varepsilon$:

[2] There are several ways to construct a hyperreal field. A classic one can be found in [10], where such a field is constructed as an ultrapower, and where the reader can find an explanation of all the necessary logical tools.

To represent this situation graphically on the line, one can proceed simply as in Figure 3.

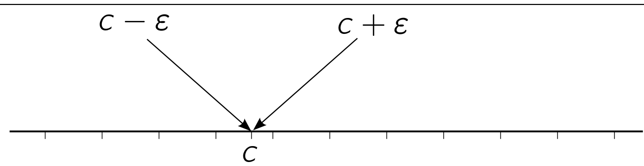
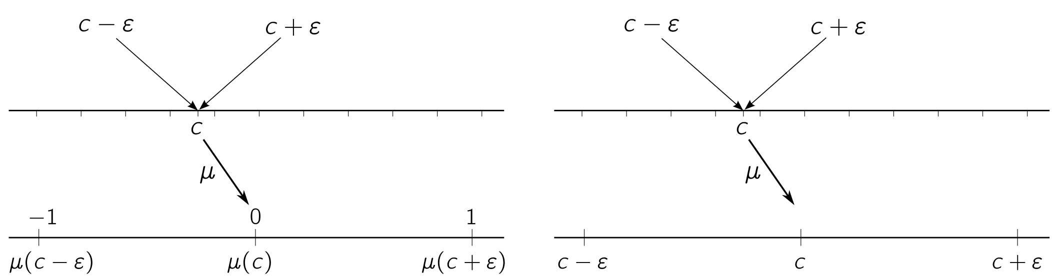


Figure 3: The points c, $c-\varepsilon$ and $c+\varepsilon$ ($c \in \mathbb{R}$ and ε infinitesimal) are visually indistinguishable on the line, but we have that $c-\varepsilon < c < c+\varepsilon$.

To make the difference between c and $c+\varepsilon$ visible, we imitate what we did in the standard case by introducing the transformation $\mu : \mathbb{R}^* \to \mathbb{R}^*$ defined by

$$\mu(x) = \frac{x-c}{\varepsilon}$$

and identifying, as usual, each x with its image $\mu(x)$. We call this transformation an *infinitesimal* or *non standard microscope*. Since $\mu(c) = 0$, $\mu(c+\varepsilon) = 1$ and



(a) The μ transformation. (b) The identification of x with $\mu(x)$.

Figure 4: The μ transformation allows the numbers $c-\varepsilon$ and $c+\varepsilon$ to be represented on the real line, distinguishing them from c.

$\mu(c-\varepsilon) = -1$, in the microscope image the points c, $c+\varepsilon$ and $c-\varepsilon$ appear quite distinct, although they have an infinitesimal distance (Figure 4).

Nevertheless, some clarifications are necessary. If we calculate the image via μ of the non-real hyperreal number $c+\varepsilon^2$, which is also infinitely close to c, we find

$$\mu(c+\varepsilon^2) = \frac{c+\varepsilon^2-c}{\varepsilon} = \varepsilon \in \mathbb{R}^* \setminus \mathbb{R}$$

that is still a non-real hyperreal number, which we cannot distinguish from the only real number to which it is infinitely close, which in this case is 0. Consequently, c

and $c+\varepsilon^2$ result indistinguishable. It is convenient to make explicit this step that assigns the position of the point in the lens image by considering *the standard part of* μ, i.e., by applying the function

$$\bar{\mu}(x) = \mathrm{st}(\mu(x)) = \mathrm{st}\left(\frac{x-c}{\varepsilon}\right)$$

which we call an *optical microscope*. Thus $\bar{\mu}(c) = 0$, $\bar{\mu}(c+\varepsilon) = 1$, $\bar{\mu}(c-\varepsilon) = -1$ and again $\bar{\mu}(c+\varepsilon^2) = 0$, so that the number $c+\varepsilon^2$ is actually seen superimposed on c (Figure 5). The new function considered, therefore, is not injective, but this fact can be easily interpreted: the microscope used is not "powerful enough" to be able to separate these numbers due to the fact that $c+\varepsilon^2$ is much closer to c than $c+\varepsilon$ is. We will see more about why in a moment.

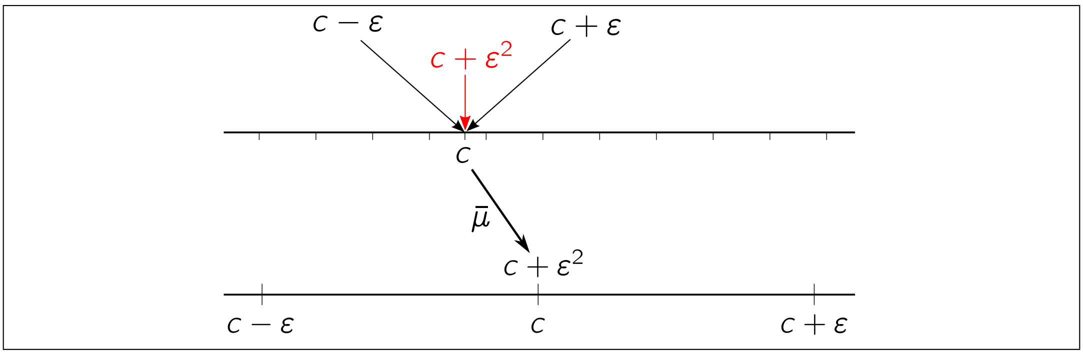


Figure 5: The numbers c and $c+\varepsilon^2$ are still found to be not separated by $\bar{\mu}$.

Let's take stock and put in order the ideas set forth so far, starting with the formal definition of an optical microscope[3].

Definition 1. *Let $c \in \mathbb{R}^*$ and $\varepsilon > 0$ be an infinitesimal. The function*

$$\mu\colon \mathbb{R}^* \to \mathbb{R}^*, \quad \mu(x) = \frac{x-c}{\varepsilon}$$

is called the ε-lens pointed at c. The field of view of the lens is the set

$$C_\mu = \{x \in \mathbb{R}^* \mid \mu(x) \text{ is finite}\}.$$

[3]The use of the standard part is Tall's proposed modification [21, 22] to the original definition discussed in the introduction. This choice remains consistent with Keisler's approach [12, p. 65 ss., online ed. p. 35 ss.], but, as will become clear below, is particularly advantageous in the application of optical tools to graphs of curves in the plane, since it allows for a better explication of the steps.

Considering the standard part of μ, we obtain the function

$$\bar{\mu}\colon C_\mu \to \mathbb{R}, \quad \bar{\mu}(x) = \mathrm{st}\left(\frac{x-c}{\varepsilon}\right)$$

called the optical ε-lens pointed in c (or optical microscope).

The adjective "optical" refers to the fact that one has moved to the standard part of the lens, which allows one to explicitly assign the position in the image. The field of view is the domain of the corresponding optical lens and represents the set of numbers that then appear in the final image, where in fact only a part of it is reproduced, however, enclosed in a circle centered generally (but not always) at the point where the lens is applied.

Even for standard microscopes, it is useful to consider optical lenses by switching to the standard part of μ. In this way they, too, can fit into the definition 1 if we replace ε with $\frac{1}{n}$: the advantage is that the function thus obtained remains defined even for the nonstandard numbers included in the field of view, consistent with the fact that no standard lens can distinguish infinitesimal details for any magnification factor. For example, in Figure 6, the standard microscope magnifies by a factor of $n \in \mathbb{N}$ and sends the infinitesimal ε to 0, while the nonstandard microscope, magnifying by a factor of infinity $\frac{1}{\varepsilon}$, manages to distinguish it from 0. In fact,

$$\bar{\mu}_{\mathrm{st}}(\varepsilon) = \mathrm{st}\left(\frac{\varepsilon - 0}{\frac{1}{n}}\right) = 0 \qquad \bar{\mu}_{\text{non-st}}(\varepsilon) = \mathrm{st}\left(\frac{\varepsilon - 0}{\varepsilon}\right) = 1.$$

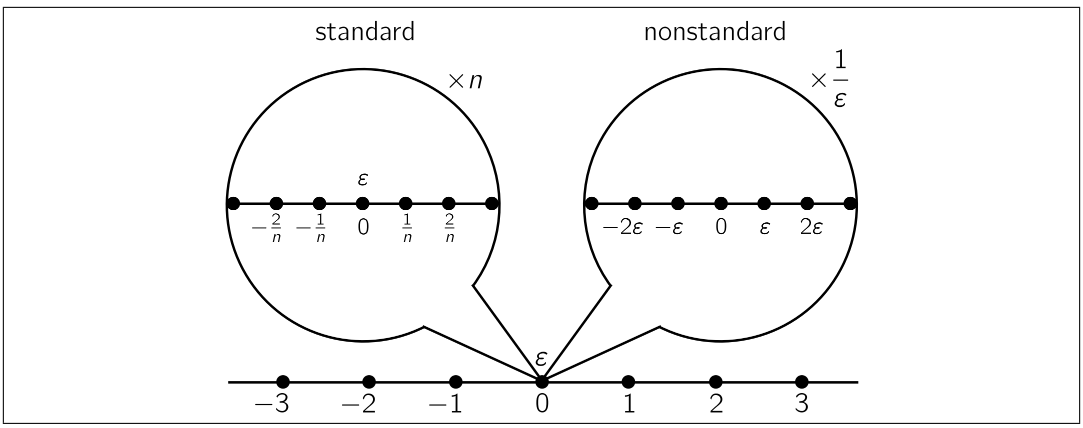


Figure 6: Standard and nonstandard microscope.

In what follows we will simply say "lenses" or "microscopes" instead of "optical lenses" or "optical microscopes", if the distinction is not necessary: the context will make it clear whether or not we are referring to the final result of the visualization.

It is now necessary to introduce a relationship between infinitesimals, which is fundamental to establishing the effective *power* of an infinitesimal microscope and understanding of what level of detail it can reveal.

Definition 2. *Given two nonzero infinitesimals ε and δ, we say that*

i) ε is of higher order than δ if $\dfrac{\varepsilon}{\delta}$ is infinitesimal. We write in that case $\varepsilon = o(\delta)$;

ii) ε is of the same order as δ if $\dfrac{\varepsilon}{\delta}$ is finite not infinitesimal;

iii) ε is of lower order than δ if $\dfrac{\varepsilon}{\delta}$ is infinite. We write in that case $\delta = o(\varepsilon)$.

Intuitively, if ε is of higher order than δ it means that ε is also infinitesimal compared to δ: at a magnification where δ and 0 are visible and well separated, ε still remains superimposed on 0. This intuition is perfectly reflected in optical microscopes: in Figure 5, the numbers $c + \varepsilon^2$ and c do not turn out to be distinct because the microscope can separate only numbers that differ by an infinitesimal of the same order as ε, while ε^2 is an infinitesimal of higher order than ε (in symbols, $\varepsilon^2 = o(\varepsilon)$).

These considerations lead to the question of how exactly to characterize the details that can be revealed by a ε-lens pointed at $c \in \mathbb{R}^*$. The answer is that one can distinguish separate from c effectively only numbers of the type

$$c + \lambda$$

where λ is an infinitesimal of the same order as ε, as can be easily verified[4]. If λ were of higher order than ε, the number $c + \lambda$ would not be distinguishable from c (infinitesimal details of higher order would be too small to be seen); if λ were of lower order, on the other hand, the image of $c + \lambda$ would be infinite and thus the number $c + \lambda$ would be out of sight.

Even more generally, we can say that *two points in the field of view of a ε-lens that differ by an infinitesimal of higher order than ε appear equal through it.*

Figure 7 shows two nonstandard microscopes pointed at 0 with magnifications of different orders: the ε-lens succeeds in separating ε from 0, from which, however, it cannot separate ε^2; the ε^2-lens succeeds in separating ε^2 from 0, but it can no

[4]To get an idea of the situation, consider numbers of the type $c+r\varepsilon$ with r varying in $\mathbb{R}$. In fact, $r\varepsilon$ is an infinitesimal of the same order as ε, and the image of $c + r\varepsilon$ through the optical ε-lens is just r. Since for visualization purposes only the standard part matters, we see that these numbers are already sufficient to complete the lens image, meaning that every other number in the monad of c that falls within the field of view would still have the same image as $c + r\varepsilon$ for some $r \in \mathbb{R}$.

longer visualize ε, which has moved out of its field of view because it is "too far away": although it is (in absolute) infinitesimal, in comparison to ε^2 it behaves as if it were infinite[5].

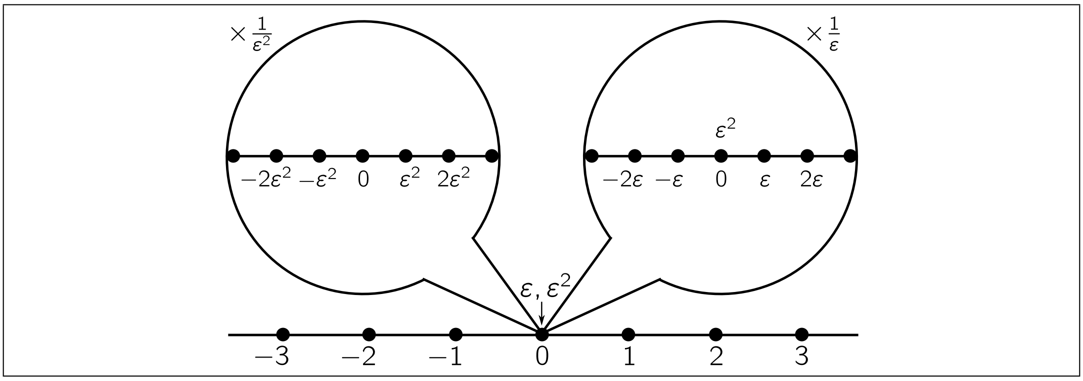


Figure 7: Nonstandard microscopes with magnifications of different orders.

5 Microscopes Pointed in Microscopes

Sometimes it can be advantageous – also from the didactic point of view – to apply a microscope in the image of another microscope, rather than directly applying a more powerful lens, in order to make explicit and clarify steps in a process or explanation.

Definition 3. *A microscope pointed in a microscope is an optical lens applied at a point on another non-optical lens.*

Applying one microscope to another is equivalent to composing two functions: technically, the first applied lens μ_1 must not be optical (otherwise it would lose all infinitesimal details), while the second $\bar{\mu}_2$, applied in the image of the first, is. This results in the composite function $\bar{\mu}_2 \circ \mu_1$. In the graphical transposition, however, the first lens μ_1 can also be calculated as optical in order to be represented in a simple way, as shown in Figure 8.

[5]This is precisely the intuitive meaning of $\varepsilon^2 = o(\varepsilon)$. Note also the analogy with the visualization through a standard microscope: as we have seen above, focusing at standard magnification on a line detail is bound to take some numbers out of the field of view and still fail to distinguish finer details. In the example in Figure 1, the resolution allows us to distinguish the numbers 3 and $3 + n \times 10^{-4}$, but 2 (and a fortiori 1 and 10) disappears from the field of view, and we certainly will not distinguish $3 + 10^{-8}$ nor $3 - 2 \times 10^{-10}$. This natural limitation found in the effectiveness of the standard microscope is analogous to the limitations of the nonstandard microscope in visualizing infinitesimals of different order.

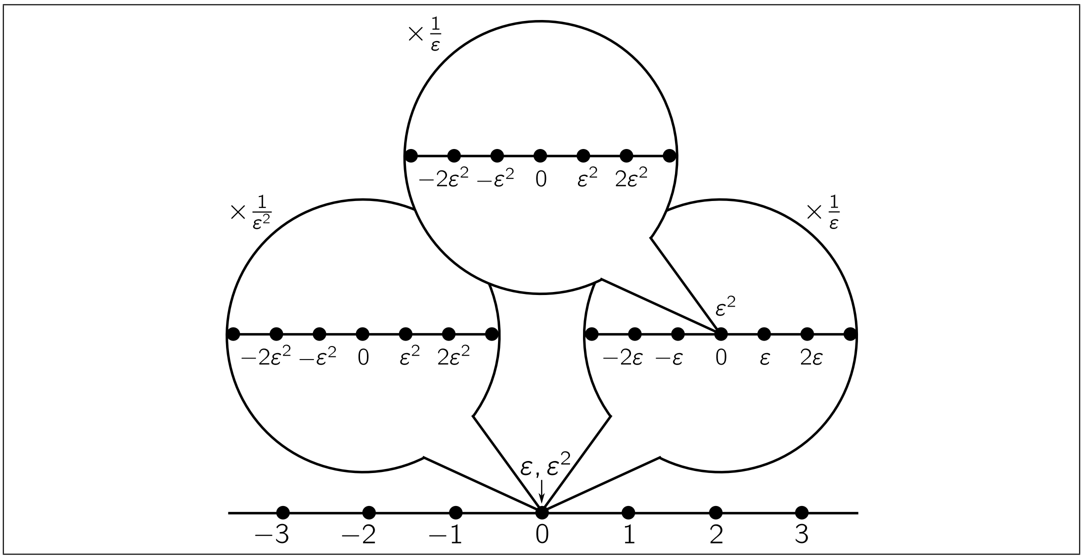


Figure 8: Nonstandard microscope pointed into another microscope.

As can be seen, pointing one microscope inside another is equivalent to applying a single microscope whose scale factor is the product of those of the other two microscopes: the final image is exactly the same. Let us fix this fact in the following lemma.

Lemma. *The image of a ε-lens applied to a point of another ε-lens coincides with the image of a ε^2-lens applied to the same point.*

Proof of the Lemma. Let $c \in \mathbb{R}^*$. Let us apply a ε-lens in the image, produced by another ε-lens, of the point $c+\lambda$, where λ is an infinitesimal of the same order of ε (or $\lambda = 0$). The first ε-lens and the second (pointed in the first) are defined respectively by

$$\mu_1(x) = \frac{x-c}{\varepsilon} \qquad \text{and} \qquad \mu_2(x) = \frac{x-\mu_1(c+\lambda)}{\varepsilon}$$

and their composition gives

$$\mu_2(\mu_1(x)) = \frac{\mu_1(x)-\mu_1(c+\lambda)}{\varepsilon} = \frac{\dfrac{x-c}{\varepsilon} - \dfrac{\cancel{c}+\lambda-\cancel{c}}{\varepsilon}}{\varepsilon} = \frac{\dfrac{x-(c+\lambda)}{\varepsilon}}{\varepsilon} = \frac{x-(c+\lambda)}{\varepsilon^2}$$

representing a ε^2-lens pointed at $c+\lambda$ (Figure 9). □

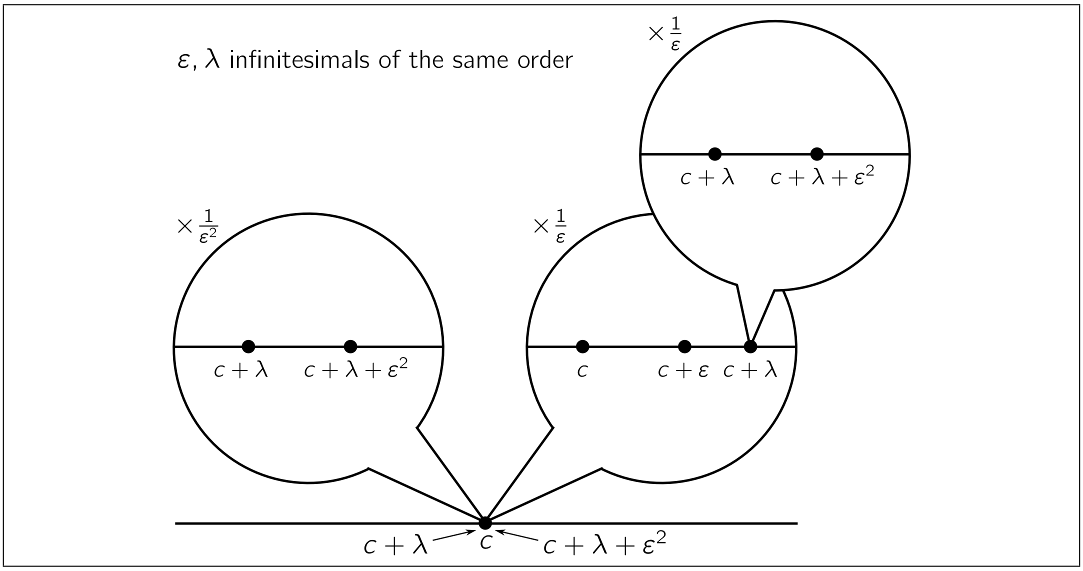


Figure 9: Equivalence of microscopes.

6 Differentiable Functions and Microscopes in Two Dimensions

The multidimensional generalization of the microscopes considered so far is straightforward: just apply a lens to each coordinate. Of interest is the two-dimensional case, which is particularly appropriate for the local description of the graph of a function in the $\mathbb{R}^2$ plane.

Definition 4. *Let $(\alpha, \beta) \in \mathbb{R}^{*2}$ and $\varepsilon > 0$ be an infinitesimal. The function*

$$\mu\colon \mathbb{R}^{*2} \to \mathbb{R}^{*2}, \quad \mu(x, y) = \left(\frac{x-\alpha}{\varepsilon}, \frac{y-\beta}{\varepsilon}\right)$$

is called the ε-lens pointed at (α, β). The field of view of the lens is the set

$$C_\mu = \{(x, y) \in \mathbb{R}^{*2} \mid \mu(x, y) \text{ is finite}\}.$$

Considering the standard part of μ, we obtain the function

$$\bar{\mu}\colon C_\mu \to \mathbb{R}^2, \quad \bar{\mu}(x, y) = \operatorname{st}\left(\frac{x-\alpha}{\varepsilon}, \frac{x-\beta}{\varepsilon}\right)$$

called the optical ε-lens pointed in (α, β) (or optical microscope).

Let us see how to use these tools to explore the world of functions. In what follows, as is usually the case, we will avoid explicit reference to the hyperreal natural extension of the functions involved. Given a function $f\colon \mathbb{R} \to \mathbb{R}$ differentiable at $a \in \mathbb{R}$, we denote by dx a positive infinitesimal increment of a. The goal is to see what the graph of the curve $y = f(x)$ looks like in an infinitesimal neighborhood of the point $(a, f(a))$, for which we will use a microscope. It is well known that the increase in the dependent variable[6]

$$dy = f(a + dx) - f(a)$$

and the increment along the tangent line at the considered point

$$f'(a)dx$$

differ by an infinitesimal of higher order with respect to dx, that is, we have

$$f(a + dx) = f(a) + f'(a)dx + \varepsilon dx$$

where ε is an infinitesimal (which depends on a and dx). We point an optical dx-lens at $(a, f(a))$, that is, we apply the function

$$\bar{\mu}(x, y) = \mathrm{st}\left(\frac{x - a}{dx}, \frac{y - f(a)}{dx}\right).$$

Then we will have

$$(a, f(a)) \mapsto (0, 0)$$

while for the point on the graph corresponding to the dx increment

$$(a + dx, f(a + dx)) \mapsto \mathrm{st}(1, f'(a) + \varepsilon) = (1, f'(a)).$$

To understand what happens to all points on the graph infinitely close to $(a, f(a))$, let's consider another infinitesimal increment λ and see where the microscope image of the corresponding point lies

$$(a + \lambda, f(a + \lambda)).$$

Through the optical dx-lens, it appears as

$$\left(\mathrm{st}\left(\frac{\lambda}{dx}\right), \mathrm{st}\left(\frac{f'(a)\lambda + \lambda\varepsilon}{dx}\right)\right) = \left(\mathrm{st}\left(\frac{\lambda}{dx}\right), \mathrm{st}\left(\frac{f'(a)\lambda}{dx} + \frac{\lambda\varepsilon}{dx}\right)\right).$$

[6] We prefer to adopt the differential symbol dy to denote the increment of the function rather than that of the tangent line, in the manner of Robinson [18, p. 79, ed. it. p. 111] and according to the recommendations of Goldoni [9].

Now, if λ has the same order as dx (and we know that only these can be involved in visualization with an optical dx-lens[7]), then λ/dx is finite and $\lambda\varepsilon/dx$ is infinitesimal. Then

$$(a+\lambda, f(a+\lambda)) \mapsto \left(\mathrm{st}\left(\frac{\lambda}{dx}\right), f'(a)\,\mathrm{st}\left(\frac{\lambda}{dx}\right)\right)$$

and putting $t = \mathrm{st}(\lambda/dx)$, we have that the points of $y = f(x)$ that fall within the field of view of the lens are sent into the parametric curve

$$(t, f'(a)t)$$

as t varies. This means that through the lens the graph of the function is seen as a straight line with a slope exactly the derivative at the given point. Of course, this line is also the tangent to the graph of the function at the point itself. We can then state that *the graph of a real function f differentiable in $a \in \mathbb{R}$ and the tangent line at $(a, f(a))$ appear indistinguishable in an infinitesimal neighborhood of $(a, f(a))$ when viewed through an optical lens*[8] (Figure 10).

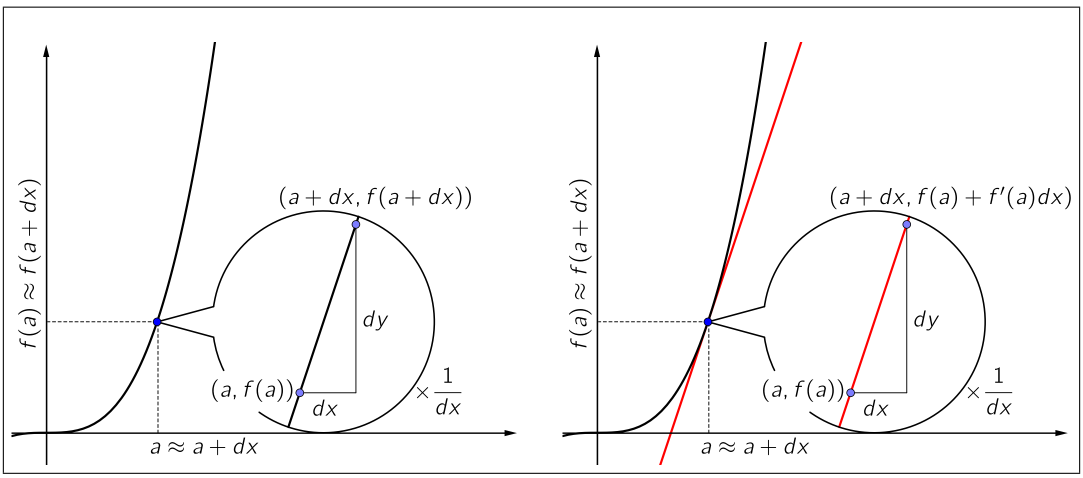


Figure 10: Through an optical microscope, the curve and the tangent line are indistinguishable.

[7]Similar to before, it is sufficient to refer to numbers of the type $r\,dx$, with $r \in \mathbb{R}$.

[8]We note that the preceding discussion was made possible by Tall's proposal to consider the standard parts in optical lenses, particularly for writing in parametric form the equation of the line within the lens image. We will also proceed similarly in the next section.

7 Differentiable Functions and Microscopes Pointed in Microscopes

Let us now see how microscopes pointed in microscopes can be use to study other properties of real functions.

Let f be a real function differentiable twice in a. Let us point in $(a+dx, f(a+dx))$, inside the dx-lens considered above, another dx-lens. Thanks to the Lemma, this can be accomplished by pointing directly into $(a+dx, f(a+dx))$ the dx^2-lens defined by

$$(x,y) \mapsto \mathrm{st}\left(\frac{x-(a+dx)}{dx^2}, \frac{y-f(a+dx)}{dx^2}\right).$$

We write the second-order Taylor expansion[9] for $f(a+dx)$

$$f(a+dx) = f(a) + f'(a)dx + \frac{1}{2}f''(a)dx^2 + \varepsilon_1 dx^2$$

where ε_1 is an infinitesimal.

We give $a+dx$ an additional increment λ of the same order as dx^2 and see, using Taylor's formula again, what is the image of

$$(a+dx+\lambda, f(a+dx+\lambda)).$$

One has

$$f(a+dx+\lambda) = f(a) + f'(a)(dx+\lambda) + \frac{1}{2}f''(a)(dx+\lambda)^2 + \varepsilon_2(dx+\lambda)^2$$

so

$$\begin{aligned}(a+dx+\lambda, f(a+dx+\lambda)) &\mapsto \mathrm{st}\left(\frac{\lambda}{dx^2}, \frac{f(a+dx+\lambda)-f(a+dx)}{dx}\right) = \\ &= \mathrm{st}\left(\frac{\lambda}{dx^2}, \frac{f'(a)\lambda + \frac{1}{2}f''(a)\lambda^2 + f''(a)dx\lambda + \varepsilon_2 dx^2 + \varepsilon_2\lambda^2 + 2\varepsilon_2 dx\lambda - \varepsilon_1 dx^2}{dx^2}\right)\end{aligned}$$

and taking the standard parts, the optical lens gives

$$\left(\mathrm{st}\left(\frac{\lambda}{dx^2}\right), f'(a)\,\mathrm{st}\left(\frac{\lambda}{dx^2}\right)\right).$$

Along the tangent line, the point corresponding to the same increment λ will be

$$(a+dx+\lambda, f(a)+f'(a)(dx+\lambda)).$$

[9] See for example [20, p. 88] or [5, p. 102].

We calculate its image

$$
\begin{aligned}
&\mathrm{st}\left(\frac{\lambda}{dx^2}, \frac{f'(a)(dx+\lambda) - f'(a)dx - \frac{1}{2}f''(a)dx^2 - \varepsilon_1 dx^2}{dx^2}\right) = \\
&= \mathrm{st}\left(\frac{\lambda}{dx^2}, \frac{\lambda f'(a) - \frac{1}{2}f''(a)dx^2 - \varepsilon_1 dx^2}{dx^2}\right) = \mathrm{st}\left(\frac{\lambda}{dx^2}, f'(a)\frac{\lambda}{dx^2} - \frac{1}{2}f''(a) - \varepsilon_1\right)
\end{aligned}
$$

and we make the standard parts explicit

$$\left(\mathrm{st}\left(\frac{\lambda}{dx^2}\right), f'(a)\,\mathrm{st}\left(\frac{\lambda}{dx^2}\right) - \frac{1}{2}f''(a)\right).$$

Ultimately, assigning to $a + dx$ infinitesimal increments λ of the same order as dx^2 and putting $t = \mathrm{st}(\lambda/dx^2)$, through a dx-lens pointed at $(a + dx, f(a + dx))$ in the image of another dx-lens pointed at $(a, f(a))$ or, which is the same, through a dx^2-lens pointed directly at $(a + dx, f(a + dx))$, the graph of the function and that of the tangent appear as the following parametric curves

$$
\begin{aligned}
\text{function} \quad &\rightarrow \quad (t, f'(a)t) \\
\text{tangent} \quad &\rightarrow \quad \left(t, f'(a)t - \frac{1}{2}f''(a)\right)
\end{aligned}
$$

that is, they are seen as two parallel lines of slope $f'(a)$. The situation is illustrated in Figure 11. Furthermore, looking at the sign of the second derivative $f''(a)$, we find that

$f''(a) > 0 \Rightarrow f$ is convex, that is, the tangent line lies below the graph of f;

$f''(a) < 0 \Rightarrow f$ is concave, that is, the tangent line lies above the graph of f.

This particular visualization, obtained through the double microscope, is taken from [11, p. 57] and [12, p. 67, online ed. p. 37], in which, however, the explanations remain at a more intuitive than formal level. Our discussion nevertheless established their correctness and consistency and also allowed further observations on the convexity/concavity of a function.

8 A Simple Application

At the beginning of the volume [6] devoted to hyperreal numbers, G. Goldoni shows how it is possible to solve, by reasoning with infinitesimals and using microscopes in

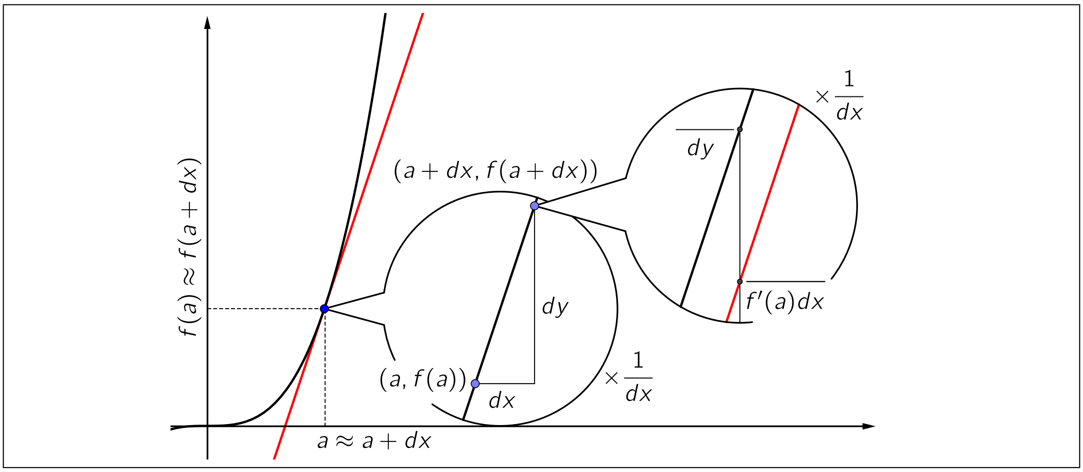


Figure 11: Microscope pointed into another microscope for analysis of a twice-differentiable function.

an intuitive way, without going too deeply into differential calculus, the problem of finding the abscissa of the vertex of a parabola with an axis parallel to the y-axis. In the same spirit, we would like to show how the optical tools we have described, applied to the graph of a sufficiently regular function, in addition to providing a very suggestive of what happens in the infinitesimal neighborhood of a point, can be exploited informally to solve certain classes of problems.

We propose below an application that is intended only as an example and does not claim to pose as a substitute method for the study of derivatives. It involves dealing with the following problem: *to find at what point the graph of a third-degree polynomial function changes concavity, that is, where it has an inflection point.* Let us first ask how a function should behave in an infinitesimal neighborhood of such a point. In light of what we have just learned, at a point where the concavity turns upward or downward, the graph of the tangent line displayed in the second microscope appears below or above that of the function, respectively. But an inflection point is a special point at which the graph *has no* concavity. How does this fact translate to optical visualization? It is reasonable to think that through the second lens the graph of the tangent and that of the function should still appear indistinguishable since this is the only way that the point does not fall into one of the two types of concavity (Figure 12). This is actually confirmed by our previous calculations: if f is a function differentiable twice at a point a where it has an inflection, we have $f''(a) = 0$, and the parametric equation of the tangent line goes to coincide with that of the function[10].

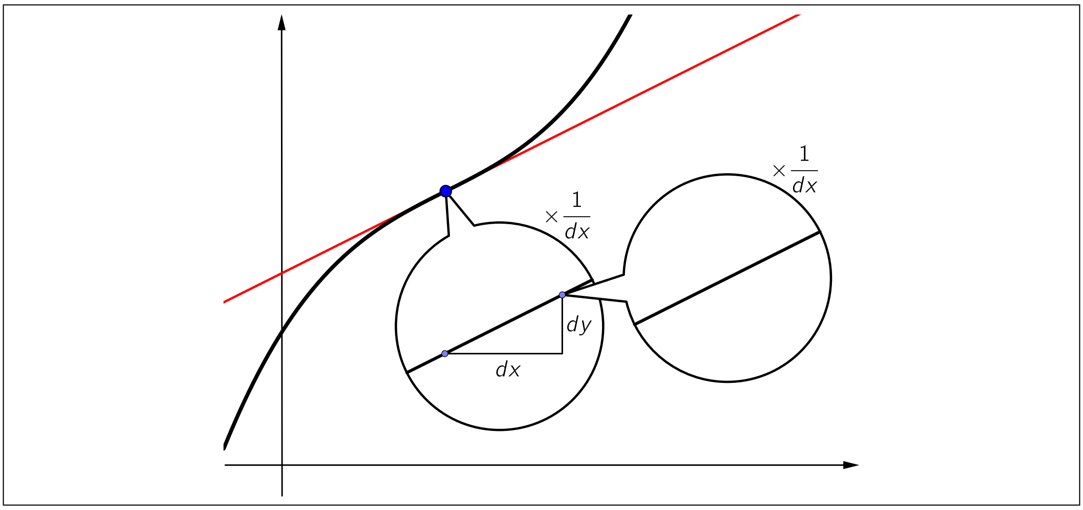


Figure 12: Microscope in a microscope pointed in an inflection point, where there is no concavity.

Consider the function whose graph is the curve of equation $y = x^3 - 2x^2$. We want to look for points where there is change in concavity, which will therefore be those for which a second microscope displays a single straight line (i.e., where it does not detect vertical distance between the curve and the tangent). Given an infinitesimal increment dx of the independent variable, we calculate the corresponding increment dy of the dependent variable

$$\begin{aligned} dy &= (x+dx)^3 - 2(x+dx)^2 - (x^3 - 2x^2) = \\ &= \cancel{x^3} + 3x^2dx + 3xdx^2 + dx^3 - \cancel{2x^2} - 4xdx - 2dx^2 - \cancel{x^3} + \cancel{2x^2} = \\ &= (3x^2 - 4x)dx + (3x-2)dx^2 + dx^3. \end{aligned}$$

Note that the written equality is nothing more than Taylor's formula of order 2, in which three infinitesimal addends of the same order of dx, dx^2 and dx^3, respectively, appear. The first addend (of the same order as dx) represents the increment that is detected by the first microscope (by a factor of $1/dx$), while the second is the increment detected by the second microscope in the image of the first, i.e., the vertical distance between the curve and the tangent line (higher-order infinitesimals

[10] We take this opportunity to reiterate once again that the goal of formalizing optical instruments is precisely to ensure the correctness of intuition-inspired visualization.

are neglected instead). In summary

$$dy = \underbrace{(3x^2 - 4x)dx}_{\substack{\text{increment detected}\\\text{by the 1° microscope}}} + \underbrace{(3x-2)dx^2}_{\substack{\text{increment detected}\\\text{by the 2° microscope}}} + \underbrace{dx^3}_{\substack{\text{infinitesimal of}\\\text{higher order}\\\text{(neglected)}}}$$

and the point sought will be the one for which the dx^2-order increment detected by the second microscope is zero. This is realized if

$$3x - 2 = 0 \quad \Rightarrow \quad x = \frac{2}{3}$$

which gives precisely the inflection point of the given function[11] (Figure 13).

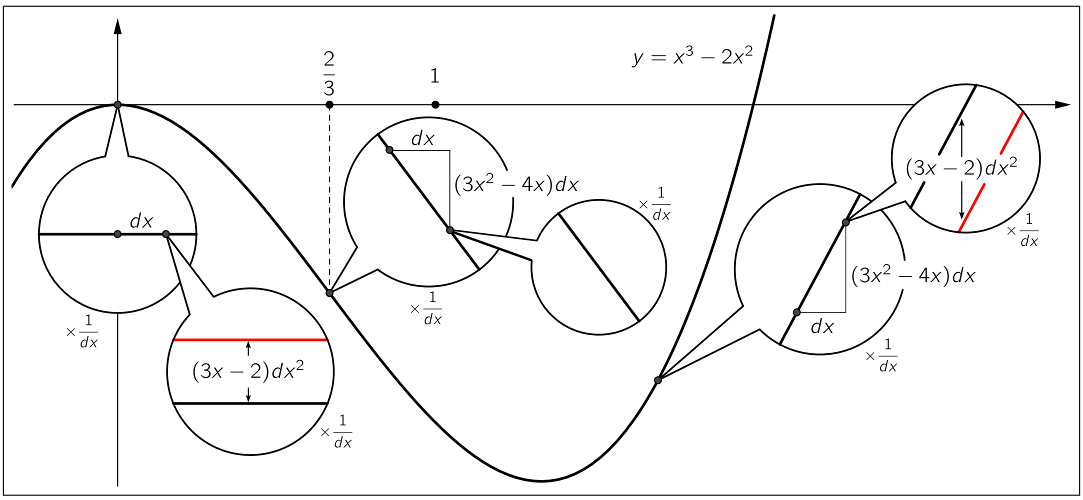


Figure 13: Use of microscopes to search for inflection points of $y = x^3 - 2x^2$.

9 Conclusion

In the first part of this article we have introduced the concept of manipulative abduction, which is widespread in cognitive behaviors that aim to create accounts of new communicable experiences so that, for example in the case of various kinds of scientific reasoning, the abductive process concerning the formation and evaluation of a

[11]Instead, canceling the term of order dx finds the points of maximum and minimum, that is, the points for which the first microscope detects no increase in the function, which thus appears indistinguishable from the horizontal tangent, exactly as in [6].

hypothesis occurs by resorting to a basically extra-theoretical and extra-sentential dimension: in this perspective manipulative abduction represents a kind of redistribution of the epistemic and cognitive effort to manage objects and information that cannot be immediately represented or found "internally" [15, Ch. 1]. An example of manipulative abduction is exactly the case of the human use of the construction of external models in a neural engineering laboratory or, as in our present case, in mathematics, exploiting external diagrams and written proofs. From a general point of view, in these cases the external tools and representations are useful to make observations and "experiments" to transform one cognitive state into another to discover new properties of the target systems/theories or to furnish new explanations and description that can be useful from either an epistemological or a didactic perspective. Manipulative abduction also refers to those more unplanned and unconscious action-based cognitive processes which we have earlier characterized as forms of "thinking through doing".

Hence, manipulative abduction is a kind of abduction, usually model-based,[12] that exploits external models endowed with delegated (and often implicit) cognitive roles and attributes. An example of manipulative abduction can be seen in the case of elementary geometrical reasoning, which takes advantage of diagrams: we can say that

(1) the model (diagram) is external and the strategy that organizes the manipulations is unknown a priori;

(2) the result achieved is new (if we, for instance, refer to the constructions of the first creators of geometry), and adds properties not contained before in the concept.

    Of course in the case in which we are using diagrams to demonstrate already known theorems or to the aim of furnishing a better clarification of the concepts at stake (for instance in didactic settings), the strategy of manipulations is not necessarily unknown and the result is not new.

In sum, the optical diagrams and the applications that we have described show that these pictorial tools enable better abductive learning and understanding of concepts related to calculus. We also believe that they can be used in various mathematical and scientific areas. Some such examples might be

[12] Visual thinking is surely the kind of model-based cognition more extendedly studied in the the epistemology of mathematics. An impressive and rich compendium is provided by Gianquinto [4], which illustrates the relationships between visual thinking, formal and non-formal proofs, and their reliability, visual thinking in discovering strategies, and a priori and a posteriori roles of visual experience.

(1) in non-Euclidean geometry, exploring the behavior at infinity of two parallel or hyperparallel lines in hyperbolic geometry;

(2) in thermodynamics, exploring what happens between two infinitely close equilibrium points of a curve representing a quasi-static process.

These topics may represent future fields of research.

The use of infinitesimal microscopes, their explanations, and consequent methods find their natural place in nonstandard analysis, indeed one might say they are a part of it. We believe that this aspect can contribute to placing nonstandard analysis a step higher than classical analysis, if not in terms of mathematical results, certainly at the level of epistemological deepness, cognitive meaning, and didactic methodology.

Funding

Research for this article was supported by the PRIN 2017 Research 20173YP4N3—MIUR, Ministry of University and Research, Rome, Italy.

References

[1] J. Bair and V. Henry. Implicit differentiation with microscopes. *The Mathematical Intelligencer*, 32(1):53–55, 2010. `http://orbi.ulg.ac.be/bitstream/2268/13591/1/ImplicitDiff.pdf`.

[2] J. Bair and V. Henry. Osculating circle with microscopes within microscopes. *Foundations of Science*, 18(2):319–325, 2013.

[3] R. Dossena and L. Magnani. Mathematics through diagrams: Microscopes in non-standard and smooth analysis. In L. Magnani and P. Li, eds., *Model-Based Reasoning in Science, Technology, and Medicine*, pp. 193–213. Springer Berlin Heidelberg, 2007.

[4] M. Gianquinto. The epistemology of visual thinking in mathematics. In E. N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2020.

[5] R. Golblatt. *Lectures on the Hyperreals. An Introduction to Nonstandard Analysis.* Springer, New York, 1998.

[6] G. Goldoni. *Il professor Apotema insegna... i numeri iperreali.* Ilmiolibro.it, 2011. Second edition, 2014. `https://independent.academia.edu/GoldoniGiorgio`.

[7] G. Goldoni. *Il professor Apotema insegna... il calcolo delle differenze e il calcolo differenziale.* Ilmiolibro.it, 2011. Second edition, 2014. `https://independent.academia.edu/GoldoniGiorgio`.

[8] G. Goldoni. *Il professor Apotema insegna... il calcolo delle somme e il calcolo integrale*. Ilmiolibro.it, 2012. Second edtion, 2014. https://independent.academia.edu/GoldoniGiorgio.

[9] G. Goldoni. *Il differenziale alla Robinson in una e più variabili*, pages 21–31. Matematicamente.it, 2014. Atti del convegno, Vicenza, 7 novembre 2014. https://independent.academia.edu/GoldoniGiorgio.

[10] A. E. Hurd and P. A. Loeb. *An Introduction of Nonstandard Real Analysis*. Academic Press, San Diego, 1985.

[11] H. J. Keisler. *Elementary Calculus*. Prindle, Weber and Schmidt, Boston, 1976. https://www.math.wisc.edu/~keisler/calc.html.

[12] H. J. Keisler. *Foundations of Infinitesimal Calculus*. Prindle, Weber and Schmidt, Boston, 1976. On-line Edition, 2007. https://www.math.wisc.edu/~keisler/foundations.html.

[13] L. Magnani. *Abduction, Reason, and Science. Processes of Discovery and Explanation*. Kluwer Academic/Plenum Publishers, New York, 2001.

[14] L. Magnani. Epistemic mediators and model-based discovery in science. In L. Magnani and N. J. Nersessian, eds., *Model-Based Reasoning: Science, Technology, Values*, pp. 305–329. Kluwer Academic/Plenum Publishers, New York, 2002.

[15] L. Magnani. *Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Springer, Heidelberg/Berlin, 2009.

[16] L. Magnani and R. Dossena. Perceiving the infinite and the infinitesimal world: unveiling and optical diagrams in mathematics. *Foundations of Science*, 10(1):7–23, 2005.

[17] Peirce, C. S., 1931-1958. *Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge, MA, vols. 1-6, Hartshorne, Hartshorne, C. and Weiss, P., eds.; vols. 7-8, Burks, A. W., ed.

[18] A. Robinson. *Non-Standard Analysis*. North Holland, Amsterdam, 1966. Revised Edition, Princeton University Press, 1996.

[19] B. Stecca and D. Zambelli. *Analisi non standard*. Online handbook in Italian. http://nsa.readthedocs.org.

[20] K. D. Stroyan. *Mathematical Background: Foundations of Infinitesimal Calculus*. Academic Press, 1997. http://homepage.math.uiowa.edu/~stroyan/InfsmlCalculus/FoundInfsmlCalc.pdf.

[21] D. Tall. Elementary axioms and pictures for infinitesimal calculus. *Bulletin of the IMA*, 18:43–48, 1982. http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1982b-axioms-for-infls.pdf.

[22] D. Tall. Natural and formal infinities. *Educational Studies in Mathematics*, 48(2):199–238, 2001. http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot2001p-esm-infinity.pdf.

[23] M. Wheeler. Is language an ultimate artifact?. *Language Sciences*, 26: 693–715, 2004.

Received 31 October 2022

On Decidable Extensions of Propositional Dynamic Logic with Converse

Daniel Rogozin
University College London
d.rogozin@ucl.ac.uk

Ilya Shapirovsky
New Mexico State University
ilshapir@nmsu.edu

Abstract

We describe a family of decidable propositional dynamic logics, where atomic modalities satisfy some extra conditions (for example, given by axioms of the logics K5, S5, or K45 for different atomic modalities). It follows from recent results [11], [12] that if a modal logic L admits a special type of filtration (so-called definable filtration), then its enrichments with modalities for the transitive closure and converse relations also admit definable filtration. We use these results to show that if logics $L_1, \dots, L_n$ admit definable filtration, then the propositional dynamic logic with converse extended by the fusion $L_1 * \dots * L_n$ has the finite model property.

Keywords Propositional Dynamic Logic with Converse, definable filtration, fusion of modal logics, finite model property, decidability

1 Introduction

The Propositional Dynamic Logic with Converse is known to be complete with respect to its standard finite models, and hence is decidable [15]. We generalize this result for a family of normal extensions of this logic.

Let **CPDL**(A) be the propositional dynamic logic with converse modalities, where A indicates the set of atomic modalities. For a set of modal formulas Ψ in the language of A, let **CPDL**(A) $+ \Psi$ be the normal extension of **CPDL**(A) with Ψ.

In [11] and [12], it was shown that if a modal logic L admits a special type of filtration (so-called definable filtration), then its enrichments with modalities for the transitive closure and converse relations also admit definable filtration. In particular, it follows that if a logic L admits definable filtration, then $\mathbf{CPDL}(\mathrm{A}) + L$ has the finite model property.

We will be interested in the case when $\mathbf{CPDL}(\mathrm{A})$ is extended by a fusion of logics $L = L_1 * \ldots * L_n$. For example, $\mathbf{CPDL}(\Diamond_1, \Diamond_2, \Diamond_3) + \mathrm{K5} * \mathrm{K45} * \mathrm{K4}$ is the extension of $\mathbf{CPDL}(\mathrm{A})$, where the first and the second atomic modalities satisfy the principle $\Diamond p \to \Box\Diamond p$, the second and the third satisfy $\Diamond\Diamond p \to \Diamond p$. We show in Theorem 10 that if the logics L_i admit definable filtration, then their fusion admits definable filtration as well. It follows that in this case $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ has the finite model property, and, if all L_i are finitely axiomatizable, $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ is decidable (Corollary 13). Consequently, we have the following decidability result (Corollary 17): if each L_i is

- one of the logics

  $$\mathrm{K},\ \mathrm{T},\ \mathrm{K4},\ \mathrm{S4},\ \mathrm{K} + \{\Diamond^m p \to \Diamond p\}\ (m \geq 1),$$

  or an extension of any of these logics with a variable-free formula,

- locally tabular (e.g., K5, K45, S5, the difference logic), or

- a *stable logic* (defined in [1]), or

- axiomatizable by canonical MFP-modal formulas (defined in [12]),

then $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ has the finite model property; if also all L_i are finitely axiomatizable, then $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ is decidable. Some particular instances of this fact (in the language without converse modalities) were known before: for the case when each L_i is a stable logic, it was announced in [10]; the case when each L_i is axiomatizable by canonical MFP-modal formulas follows from [12, Corollary 4.13].

The paper is organized as follows. Section 2 provides basic syntactic and semantic definitions. Section 3 is an exposition of necessary transfer results from [11] and [12]. Main results (Theorem 10, Corollary 13, and Corollary 17) are given in Section 4.

A preliminary report on some results of this paper was given in [16].

2 Syntactic and semantic preliminaries

We assume that the reader is familiar with basic notions of modal logic [3, 4, 9]. Below we briefly recall some of them and fix notation.

Normal logics and Kripke semantics. Fix a set $\mathrm{PV} = \{p_i \mid i < \omega\}$ of *propositional variables*. For a set A, the *set of modal* A*-formulas* Fm(A) is built from propositional variables using Boolean connectives $\bot, \to$ and unary connectives $\langle a \rangle$ for $a \in \mathrm{A}$ (*modalities*). Other connectives are defined in the standard way, in particular $[a]$ abbreviates $\neg\langle a \rangle\neg$. Sometimes we write $\Diamond_a$ for $\langle a \rangle$ and $\Box_a$ for $[a]$. If A is a singleton $\{a\}$, we write $\Diamond$ and $\Box$ for $\langle a \rangle$ and $[a]$, respectively.

A *(normal) modal* A*-logic* is a set of formulas $L \subseteq \mathrm{Fm}(\mathrm{A})$ such that:

1. L contains all Boolean tautologies;
2. For all $a \in \mathrm{A}$, $\langle a \rangle\bot \leftrightarrow \bot \in L$ and $\langle a \rangle(p \vee q) \leftrightarrow \langle a \rangle p \vee \langle a \rangle q \in L$;
3. L is closed under the rules of Modus Ponens, uniform substitution, and *monotonicity*: $\varphi \to \psi \in L$ implies $\langle a \rangle\varphi \to \langle a \rangle\psi \in L$ for all $a \in \mathrm{A}$.

For an A-logic L and a set Ψ of A-formulas, $L + \Psi$ is the smallest modal A-logic that contains $L \cup \Psi$. As usual, the smallest unimodal logic is denoted by K.

An A-frame is a structure $F = (W, (R_a)_{a\in\mathrm{A}})$, where each R_a is a binary relation on W. A model on an A-frame is a structure $M = (F, \vartheta)$, where $\vartheta : \mathrm{PV} \to \mathcal{P}(W)$, where $\mathcal{P}(W)$ is the set of all subsets of W. The truth definition is standard:

- $M, x \models p_i$ iff $x \in \vartheta(p_i)$;
- $M, x \not\models \bot$;
- $M, x \models \varphi \to \psi$ iff $M, x \not\models \varphi$ or $M, x \models \psi$;
- $M, x \models \langle a \rangle\varphi$ iff there exists y such that $xR_a y$ and $M, y \models \varphi$.

We set $M \models \varphi$ iff $M, x \models \varphi$ for all x in M, and $F \vDash \varphi$ iff $M \vDash \varphi$ for all M based on F; $Log(F)$ is the set $\{\varphi \in \mathrm{Fm}(\mathrm{A}) \mid F \vDash \varphi\}$. For a class $\mathcal{F}$ of frames, $Log(\mathcal{F}) = \bigcap\{Log(F) \mid F \in \mathcal{F}\}$. A logic L is *Kripke complete* iff it is characterized by a class $\mathcal{F}$ of frames, that is $L = Log(\mathcal{F})$. A logic L has the *finite model property* iff it is characterized by a class of finite models, or equivalently, by a class of finite frames (see, e.g., [3, Theorem 3.28]).

For a logic L, Mod(L) is the class of models such that $M \models L$, i.e., $M \models \varphi$ for all $\varphi \in L$.

Propositional Dynamic Logics. Let A be finite. The set Prog(A) (*"programs"*) is generated by the following grammar:

$$e ::= a \mid (e \cup e) \mid (e \circ e) \mid e^+ \qquad \text{for } a \in \mathrm{A}$$

Remark 1. Our language of programs is test-free.

Definition 2. A *normal propositional dynamic* A*-logic* is a normal Prog(A)-logic that contains the following formulas for all $e, f \in \mathrm{Prog}(\mathrm{A})$:

A1 $\langle e \cup f\rangle p \leftrightarrow \langle e\rangle p \vee \langle f\rangle p$,

A2 $\langle e \circ f\rangle p \leftrightarrow \langle e\rangle\langle f\rangle p$,

A3 $\langle e\rangle p \to \langle e^+\rangle p$,

A4 $\langle e\rangle\langle e^+\rangle p \to \langle e^+\rangle p$,

A5 $\langle e^+\rangle p \to \langle e\rangle p \vee \langle e^+\rangle(\neg p \wedge \langle e\rangle p)$.

The least normal propositional dynamic A-logic is denoted by **PDL**(A).

We also consider dynamic logics with converse modalities. The set $\mathrm{Prog}_t(\mathrm{A})$ is given by the following grammar:

$$e ::= a \mid (e \cup e) \mid (e \circ e) \mid e^+ \mid e^{-1} \qquad \text{for} \quad a \in \mathrm{A}$$

A *normal propositional dynamic* A*-logic with converse modalities* is a normal $\mathrm{Prog}_t(\mathrm{A})$-logic that contains the formulas **A1**–**A5** and the formulas

A6 $p \to [e]\langle e^{-1}\rangle p$

A7 $p \to [e^{-1}]\langle e\rangle p$

for all $e, f \in \mathrm{Prog}_t(\mathrm{A})$. The smallest dynamic A-logic with converses is denoted by **CPDL**(A).

The validity of formulas **A1-A7** in a frame $(W, (R_e)_{e \in \mathrm{Prog}_t(\mathrm{A})})$ is equivalent to the following identities:

$$R_{(e \circ f)} = R_e \circ R_f,\ R_{(e \cup f)} = R_e \cup R_f,\ R_{e^+} = (R_e)^+, \tag{1}$$
$$R_{e^{-1}} = (R_e)^{-1}, \tag{2}$$

where R^+ denotes the transitive closure of R, R^{-1} the converse of R; models based of such frames are called *standard*; see, e.g., [9, Chapter 10]. It is known that **CPDL**(A) is complete with respect to its standard finite models [15]. Our aim is to prove this for a family of extensions of **CPDL**(A).

3 Filtrations and decidable extensions of dynamic logic

3.1 Logics that admit definable filtration

For a model $M = (W, (R_a)_{a\in \mathrm{A}}, \vartheta)$ and a set of formulas Γ, put

$$x \sim_\Gamma y \quad \text{iff} \quad \forall\psi \in \Gamma\ (M, x \models \psi \Leftrightarrow M, y \models \psi).$$

The equivalence $\sim_\Gamma$ is said to be *induced by* Γ *in* M.

For $\varphi \in \mathrm{Fm}(\mathrm{A})$, let $\mathrm{Sub}(\varphi)$ be the set of all subformulas of φ. A set Γ of formulas is Sub-*closed*, if $\varphi \in \Gamma$ implies $\mathrm{Sub}(\varphi) \subseteq \Gamma$.

Definition 3. Let Γ be a Sub-closed set of formulas. A Γ-*filtration* of a model $M = (W, (R_a)_{a\in \mathrm{A}}, \vartheta)$ is a model $\widehat{M} = (\widehat{W}, (\widehat{R}_a)_{a\in \mathrm{A}}, \widehat{\theta})$ s.t.

1. $\widehat{W} = W/{\sim}$ for some equivalence relation $\sim$ such that $\sim \subseteq \sim_\Gamma$, i.e.,

$$x \sim y \quad \text{implies} \quad \forall\psi \in \Gamma\ (M, x \models \psi \Leftrightarrow M, y \models \psi).$$

2. $\widehat{M}, [x] \models p$ iff $M, x \models p$, for all $p \in \Gamma$. Here $[x]$ is the class of x modulo $\sim$.

3. For all $a \in \mathrm{A}$, we have $(R_a)_\sim \subseteq \widehat{R}_a \subseteq (R_a)^\Gamma_\sim$, where

$$\begin{aligned}
&[x]\,(R_a)_\sim\,[y] \quad \text{iff} \quad \exists x' \sim x\ \exists y' \sim y\ (x'\,R_a\,y'),\\
&[x]\,(R_a)^\Gamma_\sim\,[y] \quad \text{iff} \quad \forall\psi\ (\langle a\rangle\psi \in \Gamma\ \&\ M, y \models \psi \Rightarrow M, x \models \langle a\rangle\psi).
\end{aligned}$$

The relations $(R_a)_\sim$ and $(R_a)^\Gamma_\sim$ on $\widehat{W}$ are called the *minimal* and the *maximal filtered relations*, respectively.

If $\sim\, = \,\sim_\Delta$ for some finite set of formulas $\Delta \supseteq \Gamma$, then $\widehat{M}$ is called a *definable* Γ-*filtration* of the model M. If $\sim\, = \,\sim_\Gamma$, the filtration $\widehat{M}$ is said to be *strict*.

The following fact is standard:

Lemma 4 (Filtration lemma)**.** *Suppose that* Γ *is a finite* Sub-*closed set of formulas and* $\widehat{M}$ *is a* Γ-*filtration of a model* M*. Then, for all points* $x \in W$ *and all formulas* $\varphi \in \Gamma$*, we have:*

$$M, x \models \varphi \textit{ iff } \widehat{M}, [x] \models \varphi.$$

Proof. Straightforward induction on φ. □

Definition 5. We say that a class $\mathcal{M}$ of Kripke models *admits definable (strict) filtration* iff for any $M \in \mathcal{M}$ and for any finite Sub-closed set of formulas Γ, there exists a finite model in $\mathcal{M}$ that is a definable (strict) Γ-filtration of M. A logic *admits definable (strict) filtration* iff the class $\mathrm{Mod}(L)$ of its models does.

It is immediate from the Filtration lemma that if a logic admits filtration, then it has the finite model property.

Strict filtrations are the most widespread in the literature; for example, it is well-known that the logics K, T, K4, S4, S5 admit strict filtration, see e.g., [4]. Constructions where the initial equivalence is refined were also used since the late 1960s [17], [8], and later, see, e.g., [18]. Refining the initial equivalence makes the filtration method much more flexible. For example, it is not difficult to see that the logic $\mathrm{K5} = \mathrm{K} + \{\Diamond p \to \Box\Diamond p\}$ does not admit strict filtration. However, K5 admits definable filtration, see, e.g., [4, Theorem 5.35]. Another explanation is that K5 is locally tabular [14], and every locally tabular logic admits definable filtration, see Section 4.3 for details.

Notice that if a logic L admits definable filtration, then its extension with a variable-free formula φ admits definable filtration as well (for a given $L+\{\varphi\}$-model M and Γ, consider a $\Gamma \cup \mathrm{Sub}(\varphi)$-filtration).

3.2 Transferring admissibility of definable filtration

In [11] and [12], definable filtrations were used to obtain transfer results for logics enriched with modalities for the transitive closure and converse relations.

Let $e \in \mathrm{A}$. For an A-logic L, let L_e^+ be the extension of the logic L with axioms **A3**, **A4**, and **A5**, and let L_e^{C} be the extension of L with the axioms **A6** and **A7**.

For an A-model $M = (W, (R_a)_{a\in\mathrm{A}}, \vartheta)$, let M_e^{C} be its expansion with the converse of R_e:

$$M_e^{\mathrm{C}} = (W, (R_a)_{a\in\mathrm{A}}, R_e^{-1}, \vartheta);$$

similarly, M_e^+ denotes the expansion of M with the transitive closure of R_e:

$$M_e^+ = (W, (R_a)_{a\in\mathrm{A}}, R_e^+, \vartheta).$$

It is straightforward from (1) and (2) that if M is an L-model, then M_e^+ is a model of L_e^+, and M_e^{C} is a model of L_e^{C}.

Assume that a logic L admits definable filtration. In [12, Theorem 3.9], it was shown that in this case the logic L_e^+ admits definable filtration as well. This crucial result implied that $\mathbf{PDL}(\mathrm{A})+L$ has the finite model property, and if also L is finitely axiomatizable, then $\mathbf{PDL}(\mathrm{A}) + L$ is decidable [12, Theorem 4.6].

If follows from [11, Theorem 2.4] that if L admits definable filtration, then so does L_e^{C}.

Remark 6. Theorem [11, Theorem 2.4] was formulated for frames, not for models; however, the definable filtrations given in the proof of this theorem work for models without any modification.

Theorem 7 ([12],[11]). *Let* B *be a subset of a finite set* A*. If a* B*-logic* L *admits definable filtration, then* $\mathbf{CPDL}(\mathrm{A}) + L$ *has the finite model property. If also* L *is finitely axiomatizable, then* $\mathbf{CPDL}(\mathrm{A}) + L$ *is decidable.*

4 Filtrations for fusions

4.1 Fusions

Let $L_1, \ldots, L_n$ be logics in languages that have mutually disjoint sets of modalities. The *fusion* $L_1 * \ldots * L_n$ is the smallest logic that contains $L_1, \ldots, L_n$. We adopt the following convention: for logics $L_1, \ldots, L_n$ in the same language, we also write $L_1 * \ldots * L_n$ assuming that we "shift" modalities; e.g., $\mathrm{K5} * \mathrm{K5}$ denotes the bimodal logic given by the two axioms $\Diamond_i p \to \Box_i \Diamond_i p$, $i = 1, 2$.

It is known that the fusion of consistent modal logics is a conservative extension of its components [20]. Also, the fusion operation preserves Kripke completeness, decidability, and the finite model property [13, 6, 21].

In [12], it was noted that if canonical logics $L_1, \ldots, L_n$ admit strict filtration, then the fusion $L = L_1 * \ldots * L_n$ admits strict filtration; it follows from Theorem 7 that $\mathbf{CPDL}(\mathrm{A}) + L$ has the finite model property for the case of such L.

Example 8. *The logic* $\mathbf{CPDL}(\Diamond_1, \Diamond_2)+$ S4 $*$ S5 *has the finite model property and decidable.*

It does not cover many important examples where logics L_i do not admit strict filtration (like in the case of the logic $\mathrm{K5} * \mathrm{K5}$). We will show below that the admissibility of definable filtration is preserved under the operation of fusion, that extends applications of Theorem 7 significantly.

4.2 Main result

Recall that a set of formulas Ψ is *valid* in a modal algebra B, in symbols $B \vDash \Psi$, iff $\varphi = 1$ holds in B for every $\varphi \in \Psi$.

For a model $M = (W, (R_a)_{a \in \mathrm{A}}, \vartheta)$ and an A-formula φ, put $\varphi_M = \{x \mid M, x \vDash \varphi\}$. Let $D(M) = \{\varphi_M \mid \varphi \in \mathrm{Fm}(\mathrm{A})\}$ be the set of definable subsets of M, considered as a Boolean subalgebra of the powerset algebra $\mathcal{P}(W)$, and let $Alg(M)$ be the modal algebra $(D(M), (f_a)_{a \in \mathrm{A}})$, where $f_a(V) = R_a^{-1}[V]$ for $V \subseteq W$. The following fact is standard: if L is a logic, then

$$M \vDash L \text{ iff } Alg(M) \vDash L \tag{3}$$

("if" is trivial, "only if" follows from the fact that logics are closed under substitutions). If $M' = (W, (R_a)_{a\in \mathrm{A}}, \vartheta')$ is a model such that $\vartheta'(p) \in D(M)$ for all variables p, then it follows from (3) that

$$\text{if } M \vDash L, \text{ then } M' \vDash L; \tag{4}$$

indeed, $Alg(M')$ is a subalgebra of $Alg(M)$.

Proposition 9. *Let Γ be a* Sub*-closed set of formulas, $M = (W, (R_a)_{a\in\mathrm{A}}, \vartheta)$ a model. If $\widehat{M} = (W/{\approx}, (\widehat{R}_a)_{a\in\mathrm{A}}, \widehat{\theta})$ is a Γ-filtration of M for some equivalence $\approx$, then for every equivalence $\sim$ finer than $\approx$ there exists a Γ-filtration $\widehat{M}'$ of M such that $W/{\sim}$ is the carrier of $\widehat{M}'$ and*

$$\widehat{M} \vDash \varphi \textit{ iff } \widehat{M}' \vDash \varphi \tag{5}$$

for every $\varphi \in \mathrm{Fm}(\mathrm{A})$.

Proof. Since $\sim \subseteq \approx$, for every $u \in W/{\sim}$ there exists a unique element of $\widehat{M}$ that contains u; we denote it by $u^{\approx}$. The binary relations $\widehat{R}'_a$ in $\widehat{M}'$ and the valuation $\widehat{\theta}'$ are defined as follows:

$$\widehat{R}'_a = \{(u, v) \mid (u^{\approx}, v^{\approx}) \in \widehat{R}_a\};$$
$$\widehat{\theta}'(p) = \{u \in W/{\sim} \mid u^{\approx} \in \widehat{\theta}(p)\} \text{ for } p \in \mathrm{PV}\,.$$

It is straightforward that the map $u \mapsto u\Delta$ is a p-morphism of a model $\widehat{M}'$ onto $\widehat{M}$. By the p-morphism lemma (see, e.g., [9, Section 1]), we have

$$\widehat{M}', u \vDash \varphi \text{ iff } \widehat{M}, u^{\approx} \vDash \varphi. \tag{6}$$

Now (5) follows.

Trivially, $\sim \subseteq \sim_{\Gamma}$. The second filtration condition follows from the definition of $\widehat{\theta}'$. Let $a \in \mathrm{A}$. For $x \in W$, let $[x]_{\approx}$ and $[x]_{\sim}$ be the classes of x modulo $\approx$ and $\sim$, respectively. If xR_ay, then $[x]_{\approx}\widehat{R}_a[y]_{\approx}$, because $\widehat{M}$ is a filtration of M; now $[x]_{\sim}\widehat{R}'_a[y]_{\sim}$ by the definition of $\widehat{R}'_a$. That $\widehat{R}'_a$ is contained in the maximal filtered relation follows from (6). ☐

The following is a generalization of [12, Theorem 4.8].

Theorem 10. *If logics L_1 and L_2 admit definable filtration, so does $L_1 * L_2$.*

Proof. Let A and B be alphabets of modalities of the logics L_1 and L_2, respectively. Without loss of generality we may assume that A and B are disjoint.

Consider an $L_1 * L_2$-model $M = (W, (R_a)_{a\in \mathrm{A}}, (R_b)_{b\in \mathrm{B}}, \vartheta)$, and a finite Sub-closed set of formulas $\Gamma \subset \mathrm{Fm}(\mathrm{A}\cup\mathrm{B})$. Consider a set of fresh variables $V = \{q_\varphi \mid \varphi \in \Gamma\}$, and define a valuation η in W as follows: for $q_\varphi \in V$, let $\eta(q_\varphi) = \{x \mid M, x \vDash \varphi\}$; otherwise, put $\eta(q) = \varnothing$. Let $M_V = (W, (R_a)_{a\in \mathrm{A}}, (R_b)_{b\in \mathrm{B}}, \eta)$. We have:

$$D(M_V) \subseteq D(M), \tag{7}$$

and by (4),

$$M_V \vDash L_1 * L_2. \tag{8}$$

Consider the A- and B-reducts of M_V:

$$M_\mathrm{A} = (W, (R_a)_{a\in \mathrm{A}}, \eta), \quad M_\mathrm{B} = (W, (R_b)_{b\in \mathrm{B}}, \eta).$$

It follows from (8) that

$$M_\mathrm{A} \vDash L_1, \quad M_\mathrm{B} \vDash L_2. \tag{9}$$

Consider the following sets of formulas:

$$\Gamma_\mathrm{A} = V \cup \{\langle a\rangle q_\varphi \mid \langle a\rangle\varphi \in \Gamma \;\&\; a \in \mathrm{A}\}, \quad \Gamma_\mathrm{B} = V \cup \{\langle b\rangle q_\varphi \mid \langle b\rangle\varphi \in \Gamma \;\&\; b \in \mathrm{B}\}.$$

Since logics L_1 and L_2 admit definable filtration, there are finite sets Δ_A and Δ_B of formulas, and models $\widehat{M}_\mathrm{A}$, $\widehat{M}_\mathrm{B}$ such that

$$
\begin{array}{lll}
\widehat{M}_\mathrm{A} \vDash L_1, & \widehat{M}_\mathrm{B} \vDash L_2, & (10)\\
\Gamma_\mathrm{A} \subseteq \Delta_\mathrm{A} \subset \mathrm{Fm}(\mathrm{A}), & \Gamma_\mathrm{B} \subseteq \Delta_\mathrm{B} \subset \mathrm{Fm}(\mathrm{B}), & (11)\\
\widehat{M}_\mathrm{A} \text{ is a } \Gamma_\mathrm{A}\text{-filtration of } M_\mathrm{A}, & \widehat{M}_\mathrm{B} \text{ is a } \Gamma_\mathrm{B}\text{-filtration of } M_\mathrm{B}, & (12)\\
\text{the carrier of } \widehat{M}_\mathrm{A} \text{ is } W/{\sim_\mathrm{A}}, & \text{the carrier of } \widehat{M}_\mathrm{B} \text{ is } W/{\sim_\mathrm{B}}, & (13)
\end{array}
$$

where $\sim_\mathrm{A}$ is the equivalence on W induced by Δ_A in M_A, and $\sim_\mathrm{B}$ is the equivalence on W induced by Δ_B in M_B. Let $\sim$ be the equivalence $\sim_\mathrm{A} \cap \sim_\mathrm{B}$. By Proposition 9 and (10), there are models $\widehat{M}'_\mathrm{A}$ and $\widehat{M}'_\mathrm{B}$ whose carrier is $W/{\sim}$ such that

$$
\begin{array}{lll}
\widehat{M}'_\mathrm{A} \vDash L_1, & \widehat{M}'_\mathrm{B} \vDash L_2, & (14)\\
\widehat{M}'_\mathrm{A} \text{ is a } \Gamma_\mathrm{A}\text{-filtration of } M_\mathrm{A}, & \widehat{M}'_\mathrm{B} \text{ is a } \Gamma_\mathrm{B}\text{-filtration of } M_\mathrm{B}. & (15)
\end{array}
$$

Notice that Γ_A and Γ_B contain the same variables, namely V. The value of any variable in V is the same in $\widehat{M}'_\mathrm{A}$ as in $\widehat{M}'_\mathrm{B}$. Also, we can assume that the values of variables not in V are empty in these models: making them empty does not affect

(14) by (4), and (15) by the definition of filtration. Consequently, we can assume that $\widehat{M}'_{\mathrm{A}}$ and $\widehat{M}'_{\mathrm{B}}$ have the same valuation:

$$\widehat{M}'_{\mathrm{A}} = (W/{\sim}, (\widehat{R}_a)_{a\in\mathrm{A}}, \widehat{\eta}), \quad \widehat{M}'_{\mathrm{B}} = (W/{\sim}, (\widehat{R}_b)_{b\in\mathrm{B}}, \widehat{\eta}). \tag{16}$$

By (15), the model

$$\widehat{M}_V = (W/{\sim}, (\widehat{R}_a)_{a\in\mathrm{A}}, (\widehat{R}_b)_{b\in\mathrm{B}}, \widehat{\eta}) \text{ is a } (\Gamma_{\mathrm{A}} \cup \Gamma_{\mathrm{B}})\text{-filtration of } M_V. \tag{17}$$

By (14),

$$\widehat{M}_V \vDash L_1 * L_2. \tag{18}$$

Finally, let $\widehat{M} = (W/{\sim}, (\widehat{R}_a)_{a\in\mathrm{A}}, (\widehat{R}_b)_{b\in\mathrm{B}}, \widehat{\theta})$, where $\widehat{\theta}(p) = \widehat{\eta}(q_p)$ for $p \in \Gamma$, and $\widehat{\theta}(p) = \varnothing$ otherwise. By (18) and (4),

$$\widehat{M} \vDash L_1 * L_2.$$

Let us show that $\widehat{M}$ is a definable Γ-filtration of M.

First, observe that $\sim$ is induced in M_V by the set $\Delta_{\mathrm{A}} \cup \Delta_{\mathrm{B}}$, and so it is induced in M by a set of formulas according to (7). Since $V \subseteq \Delta_{\mathrm{A}} \cup \Delta_{\mathrm{B}}$, the equivalence $\sim$ refines the equivalence $\sim_\Gamma$ induced in M by Γ.

Let $c \in \mathrm{A} \cup \mathrm{B}$. That $\widehat{R}_c$ contains the corresponding minimal filtered relation follows from (17). Let us show that $\widehat{R}_c$ is contained in the maximal filtered relation $(R_c)^{\Gamma}_{\sim}$. Notice that by the definition of η, for every $\varphi \in \Gamma$, $z \in W$,

$$M_V, z \vDash q_\varphi \text{ iff } M, z \vDash \varphi, \text{ and hence } M_V, z \vDash \langle c\rangle q_\varphi \text{ iff } M, z \vDash \langle c\rangle\varphi. \tag{19}$$

Consider $\sim$-classes $[x]$, $[y]$ of $x, y \in W$, and assume that $\langle c\rangle\varphi \in \Gamma$ and $M, y \vDash \varphi$. By (19), $M_V, y \vDash q_\varphi$. We have $\langle c\rangle q_\varphi \in \Gamma_{\mathrm{A}} \cup \Gamma_{\mathrm{B}}$, so by (17), $M_V, x \vDash \langle c\rangle q_\varphi$. By (19) again, $M, x \vDash \langle c\rangle\varphi$. □

Example 11. *By the above theorem,* K5 $*$ K5 *admits definable filtration. Consequently, the logic* $\mathbf{CPDL}(\Diamond_1, \Diamond_2)+$ K5 $*$ K5 *has the finite model property and decidable.*

Remark 12. Dynamic logics based on atomic modalities satisfying K5 are considered in the context of epistemic logic and logical investigation of game theory, see, e.g., [7] (in this context, the axiom $\Diamond p \to \Box\Diamond p$ is usually addressed as *negative introspection*).

From Theorems 7 and 10, we obtain:

Corollary 13. *Let* A *be a finite set,* $L_1, \dots, L_n$ *be logics such that* $L_1 * \dots * L_n \subseteq \mathrm{Fm}(\mathrm{A})$. *If* $L_1, \dots, L_n$ *admit definable filtration, then* $\mathbf{CPDL}(\mathrm{A}) + L_1 * \dots * L_n$ *has the finite model property. If also* $L_1, \dots, L_n$ *are finitely axiomatizable, then* $\mathbf{CPDL}(\mathrm{A}) + L_1 * \dots * L_n$ *is decidable.*

4.3 Examples

As we mentioned, for the logics K, T, K4, S4, S5, as well as for many others, strict filtrations are well-known, see e.g., [4, Chapter 5]. In fact, there is a continuum of modal logics that admit strict filtration. In [1], a family of modal logics called *stable* was introduced. Logics T or $\mathrm{K} + \{\Diamond\top\}$ are examples of stable logics. Every stable logic admits strict filtration, which follows from [1, Theorem 7.8], and there are continuum many stable logics [2, Theorem 6.7].

Remark 14. Stable logics were also used to construct decidable extensions of **PDL**. Namely, in [10], it was announced that extensions of **PDL** with axioms of stable logics have the finite model property.

Another class of logics that admit strict filtration are logics given by canonical *MFP-modal formulas* introduced in [12, Section 4.2].

There are logics that do not admit strict filtration, but admit definable filtrations. Consider the family of logics $\mathrm{K} + \{\Diamond^m p \to \Diamond p\}$ for $m \geq 3$. These logics are Kripke complete, and their frames are characterized by the conditions

$$\forall x\, \forall y\, (xR^m y \Rightarrow xRy); \tag{20}$$

moreover, all these logics admit definable filtration [8, Theorem 8]: for a given Γ and a model, the required filtration can be built by letting $\Delta = \{\Diamond^i\varphi \mid \varphi \in \Gamma \;\&\; i \leq m-2\}$. However, these logics do not admit strict filtration. We will illustrate it with the case when $m = 3$, one can generalize it for any $m \geq 3$.

Example 15. $L = \mathrm{K} + \{\Diamond\Diamond\Diamond p \to \Diamond p\}$ *does not admit strict filtration.*

Proof. Consider a five-element model $M = (W, R, \vartheta)$, where the binary relation is defined by the following figure

$$\begin{tikzcd} x \arrow[r] & y & & \\ & y' \arrow[r] & z \arrow[r] & u \end{tikzcd}$$

(R is assumed to be irreflexive), and

$$\vartheta(p) = \{x\}, \quad \vartheta(q) = \{y, y'\}, \quad \vartheta(r) = \{u\}.$$

By (20), the frame of M validates $\Diamond\Diamond\Diamond p \to \Diamond p$, and so M is a model of the logic L. Let $\Gamma = \{p, q, r, \Diamond r\}$. Assume that $\widehat{M} = (W/{\sim_\Gamma}, \widehat{R}, \widehat{\vartheta})$ is a Γ-filtration of M and show that $\widehat{M}$ is not an L-model. Notice that y and y' are $\sim_\Gamma$-equivalent, and hence the quotient $W/{\sim_\Gamma}$ consists of four elements $[x], [y](= [y']), [z], [u]$. Since $\widehat{R}$ contains the minimal filtered relation, we have $[x]\widehat{R}[y]\widehat{R}[z]\widehat{R}[u]$. For the sake of contradiction, assume that $\widehat{M} \vDash L$. We have $\widehat{M}, [u] \vDash r$, and so $\widehat{M}, [x] \vDash \Diamond\Diamond\Diamond r$. Then $\widehat{M}, [x] \vDash \Diamond r$ by assumption. Since $\Diamond r \in \Gamma$ and $\widehat{M}$ is a Γ-filtration of M, we have $M, x \vDash \Diamond r$, which contradicts the definition of M. Hence $\widehat{M}$ is not an L-model. $\square$

A continuum of logics that admit definable filtration are locally tabular logics. Recall that a logic L is *locally tabular*, if, for every finite k, L contains only a finite number of pairwise nonequivalent formulas in a given k variables. Well-known examples of locally tabular modal logics are K5 [14] and so its extensions (e.g., K45, S5), or the *difference logic* $\mathrm{K} + \{p \to \Box\Diamond p,\ \Diamond\Diamond p \to \Diamond p \vee p\}$ [5].

Let $M = (W, (R_a)_{a\in \mathrm{A}}, \theta)$ be a model of a locally tabular logic L, $\Gamma \subset \mathrm{Fm}(\mathrm{A})$ a finite Sub-closed set of formulas. Let V be the set of all variables occurring in Γ, and let Δ be the set of all A-formulas with variables in V. Let $F_L\langle V\rangle$ be the canonical frame of L built from maximal L-consistent subsets of Δ; the canonical relations are defined in the standard way. Consider the maximal Δ-filtration $\widehat{M}$ of M with the carrier $W/{\sim_\Delta}$; in [19], such filtrations are called *canonical*. Since L is locally tabular, $\widehat{M}$ is finite. The frame $\widehat{F}$ of $\widehat{M}$ is isomorphic to a generated subframe of $F_L\langle V\rangle$, see, e.g., [19] for details. Since L is locally tabular, $F_L\langle V\rangle$ is finite, and so $F_L\langle V\rangle \vDash L$. It follows that $\widehat{M} \vDash L$, as required. Hence, we have

Theorem 16 (Corollary from [19])**.** *If L is locally tabular, then L admits definable filtration.*

Putting the above examples together, we obtain the following instance of Corollary 13.

Corollary 17. *Let* A *be a finite set, $L_1, \ldots, L_n$ be logics such that $L_1 * \ldots * L_n \subseteq$* Fm(A)*. If each L_i is*

- *one of the logics*

$$\mathrm{K},\ \mathrm{T},\ \mathrm{K4},\ \mathrm{S4},\ \mathrm{K} + \{\Diamond^m p \to \Diamond p\}\ (m \geq 1),$$

  or an extension of any of these logics with a variable-free formula,
- *locally tabular (e.g.,* K5, K45, S5*, the difference logic), or*
- *a stable logic, or*
- *axiomatizable by canonical MFP-modal formulas,*

then $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ *has the finite model property. If also all L_i are finitely axiomatizable, then* $\mathbf{CPDL}(\mathrm{A}) + L_1 * \ldots * L_n$ *is decidable.*

Acknowledgement

The authors wishes to thank Nick Bezhanishvili for valuable discussions. The authors are also grateful to the referees for their comments on an earlier version of this paper.

References

[1] Guram Bezhanishvili, Nick Bezhanishvili, and Rosalie Iemhoff. Stable canonical rules. *The Journal of Symbolic Logic*, 81(1):284–315, 2016.

[2] Guram Bezhanishvili, Nick Bezhanishvili, and Julia Ilin. Stable modal logics. *The Review of Symbolic Logic*, 11(3):436–469, 2018.

[3] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2002.

[4] Alexander Chagrov and Michael Zakharyaschev. *Modal Logic*, volume 35 of *Oxford Logic Guides*. Oxford University Press, 1997.

[5] Leo Esakia. Weak transitivity-restitution. *Logical Studies*, 8:244–255, 2001.

[6] Kit Fine and Gerhard Schurz. Transfer theorems for multimodal logics. *Logic and reality: essays on the legacy of Arthur Prior*, pages 169–213, 1996.

[7] Melvin Fitting. Reasoning about games. *Studia Logica*, 99(1):143, Aug 2011.

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

[9] Robert Goldblatt. *Logics of Time and Computation*. Number 7 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 2 edition, 1992.

[10] Julia Ilin. Stable canonical rules and formulas for PDL. In *Advances in Modal Logic, Short Papers*, 2016.

[11] Stanislav Kikot, Ilya Shapirovsky, and Evgeny Zolin. Filtration safe operations on frames. In Rajeev Goré, Barteld P. Kooi, and Agi Kurucz, editors, *Advances in Modal Logic*, number 10, pages 333–352. College Publications, 2014.

[12] Stanislav Kikot, Ilya Shapirovsky, and Evgeny Zolin. Modal logics with transitive closure: Completeness, decidability, filtration. In Nicola Olivetti, Rineke Verbrugge, Sara Negri, and Gabriel Sandu, editors, *13th Conference on Advances in Modal Logic, AiML 2020, Helsinki, Finland, August 24-28, 2020*, pages 369–388. College Publications, 2020.

[13] Marcus Kracht and Frank Wolter. Properties of independently axiomatizable bimodal logics. *The Journal of Symbolic Logic*, 56(4):1469–1485, 1991.

[14] Michael C. Nagle and S. K. Thomason. The extensions of the modal logic K5. *Journal of Symbolic Logic*, 50(1):102–109, 1985.

[15] Rohit Parikh. The completeness of propositional dynamic logic. In J. Winkowski, editor, *Mathematical Foundations of Computer Science 1978*, pages 403–415, Berlin, Heidelberg, 1978. Springer Berlin Heidelberg.

[16] Daniel Rogozin and Ilya Shapirovsky. On decidable extensions of propositional dynamic logic. In *Advances in Modal Logic, Short Papers*, 2022.

[17] Krister Segerberg. Decidability of four modal logics. *Theoria*, 34:21–25, 1968.

[18] Valentin Shehtman. Filtration via bisimulation. In Renate Schmidt, Ian Pratt-Hartmann, Mark Reynolds, and Heinrich Wansing, editors, *Advances in Modal Logic*, number 5, pages 289–308. King's College Publications, 2004.

[19] Valentin Shehtman. Canonical filtrations and local tabularity. In Rajeev Goré, Barteld P. Kooi, and Agi Kurucz, editors, *Advances in Modal Logic*, number 10, pages 498–512. College Publications, 2014.

[20] S. K. Thomason. Independent propositional modal logics. *Studia Logica*, 39(2):143–144, 1980.

[21] Frank Wolter. Fusions of modal logics revisited. In Marcus Kracht, Maarten de Rijke, Heinrich Wansing, and Michael Zakharyaschev, editors, *Advances in Modal Logic 1, papers from the first workshop on "Advances in Modal logic," held in Berlin, Germany, 8-10 October 1996*, pages 361–379. CSLI Publications, 1996.

Received 30 December 2022