Institute of Philosophy
Russian Academy of Sciences

LOGICAL INVESTIGATIONS

Volume 26. Number 1

Special issue dedicated to the memory of Prof. Alexander Karpenko

Moscow
2020

Федеральное государственное бюджетное учреждение науки
Институт философии Российской академии наук

ЛОГИЧЕСКИЕ ИССЛЕДОВАНИЯ

Том 26. Номер 1

Специальный выпуск посвящен памяти профессора А.С. Карпенко

Москва
2020

ISSN 2074-1472 (Print)
ISSN 2413-2713 (Online)

Logical Investigations

Scientific-Theoretical Journal

2020. Volume 26. Number 1

Publisher: Institute of Philosophy, Russian Academy of Sciences

Frequency: 2 times per year

First issue: 1993; the journal is a redesigned continuation of the annual *Logical Investigations* that has been published since 1993 till 2015

The journal is registered with the Federal Service for Supervision of Communications, Information Technology, and Mass Media (Roskomnadzor). The Mass Media Registration Certificate No. FS77-61228 on April 3, 2015

Abstracting and indexing: *Zentralblatt MATH*, *Mathematical Reviews*, *Ulrich's Periodicals Directory*, *EBSCOhost (Philosopher's Index with Full Text)*

The journal is included in the list of peer-reviewed scientific editions acknowledged by the Higher Attestation Commission of the Ministry of Education and Science of the Russian Federation

Subscription index in the United Catalogue *The Russian Press* is 42046

All materials published in *Logical Investigations* undergo peer review process

Editorial address: 12/1 Goncharnaya St., Moscow 109240, Russian Federation
Tel.: +7 (495) 697-96-65; **e-mail:** logicalinvestigations@gmail.com
Website: https://logicalinvestigations.ru

ISSN 2074-1472 (Print)
ISSN 2413-2713 (Online)

Логические исследования

Научно-теоретический журнал

2020. Том 26. Номер 1

Учредитель и издатель: Федеральное государственное бюджетное учреждение науки Институт философии Российской академии наук

Периодичность: 2 раза в год

Выходит с 1993 г.; журнал является прямым продолжением ежегодника «Логические исследования», издававшегося с 1993 по 2015 г.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации СМИ: ПИ № ФС77-61228 от 03 апреля 2015 г.

Журнал реферируется и индексируется: *Mathematical Reviews*, *Zentralblatt MATH*, *Ulrich's Periodicals Directory*, *РИНЦ*, *EBSCOhost (Philosopher's Index with Full Text)*

Журнал включен в Перечень российских рецензируемых научных журналов, рекомендованных ВАК, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук (группа научных специальностей «09.00.00 – философские науки»)

Подписной индекс в Объединенном каталоге «Пресса России» — 42046

Публикуемые материалы прошли процедуру рецензирования и экспертного отбора

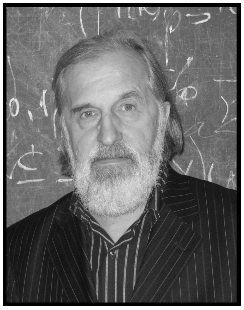
Адрес редакции: Российская Федерация, 109240, г. Москва, ул. Гончарная, д. 12, стр. 1, оф. 426

Тел.: +7 (495) 697-96-65; **e-mail:** logicalinvestigations@gmail.com

Сайт: https://logicalinvestigations.ru

TABLE OF CONTENTS

В НОМЕРЕ



Александр Степанович КАРПЕНКО

Alexander Stepanovich KARPENKO

(07.04.1946 – 07.02.2017)

От Редколлегии

Прошло вот уже более трех лет с тех пор, как 7 февраля 2017 года умер Александр Степанович Карпенко, наш коллега и товарищ. Он совмещал в себе высокий профессионализм в области логики с заразительным жизнелюбием, что привлекало к нему многих логиков как у нас в стране, так и за рубежом. С 1993 начал выходить ежегодник «Логические исследования». В 2015, благодаря организаторским усилиям А. Карпенко, ежегодник стал журналом и теперь выходит два раза в год. В нем публикуют свои работы отечественные и зарубежные авторы, что позволяет поддерживать нашу логическую жизнь и коммуникацию. Мы решили посвятить настоящий номер памяти А. Карпенко, пригласив к публикации своих работ известных логиков, которые не только знали его, но и могут быть названы его друзьями. Наш призыв получил живой отклик, чему мы не были удивлены. Предлагаем вашему вниманию работы на актуальные темы столь любимой нами и А. Карпенко науки.

Editor's note

More than three years have passed since the death of Alexander Stepanovich Karpenko, our colleague and friend, on February 7, 2017. He combined high professionalism in the field of logic with an infectious love of life, which attracted many logicians both in our country and abroad. Since 1993, volumes of the yearbook "Logical Investigations" began to come out. In 2015, thanks to the organizational efforts of A. Karpenko, the yearbook became a journal which now comes out twice a year. Russian and foreign authors publish their works in it, which allows us to maintain our logical life and communication. We decided to devote the present issue to the memory of A. Karpenko by inviting famous logicians who not only knew him, but who could be called his friends, to publish their works. Unsurprisingly, our call received a lively response. We bring to your attention works on current topics of the science so beloved by us and A. Karpenko.

Логические исследования
2020. Т. 26. № 1. С. 9–35
УДК 16

Logical Investigations
2020, Vol. 26, No. 1, pp. 9–35
DOI: 10.21146/2074-1472-2020-26-1-9-35

Философия и логика
Philosophy and logic

Diderik Batens

Devising the set of abnormalities for a given defeasible rule

Diderik Batens
Centre for Logic and Philosophy of Science, Universiteit Gent,
Blandijnberg 2, 9000 Gent, Belgium.
E-mail: `Diderik.Batens@UGent.be`

Abstract: Devising adaptive logics usually starts with a set of abnormalities and a deductive logic. Where the adaptive logic is ampliative, the deductive logic is the lower limit logic, the rules of which are unconditionally valid. Where the adaptive logic is corrective, the deductive logic is the upper limit logic, the rules of which are valid in case the premises do not require any abnormalities to be true. In some cases, the idea for devising an adaptive logic does not relate to a set of abnormalities, but to one or more defeasible rules, and perhaps also to one of the deductive logics. Defeasible rules are not universally valid, but are valid in 'normal situations' or for unproblematic parts of premise set. Where the idea is such, the set of abnormalities has to be delineated in view of the rules. The way in which this task may be tackled is by no means obvious and is the main topic studied in the present paper. The outcome is an extremely simple and transparent recipe. It is shown that, except for very special cases, the recipe leads to an adequate result.

Keywords: adaptive logics, defeasible reasoning, defeasible rules, conditional derivation, dynamic proofs, abnormalities, falsehood, content guidance

For citation: Batens D. "Devising the set of abnormalities for a given defeasible rule", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 9–35. DOI: 10.21146/2074-1472-2020-26-1-9-35

Envoi

This paper is dedicated to the memory of Alexander Karpenko. We got to know each other better when Alexander was responsible for three Moscow institutes in a European project ran by me on behalf of my home university, Ghent, and Salzburg, and Brussels (VUB). I still treasure a booklet with poems by Bielo Cardinal — the White Cardinal, an allusion on Alexander's home country Belarus. I cannot read the poems, let alone understand them. Yet, at a dinner party in his home, the poet read some of them to me, and I associated them with the poems of one of my favourite writers in my home tongue, Willem Elsschot, who, apart from a fat volume of novels, left us some twenty impressive poems.

While Sasha declaimed his poems, he became for me the symbol of man reaching for what cannot be attained. That we reach anyway may, more than anything else, makes our lives meaningful. It incites and motivates us to stand by our fellow humans, to build a better world and to create beauty.

1. Aim Of This Paper

When awake, humans are in a conscious or semi-conscious state. In that state, their brain activity leads to results of many sorts: perceptions, observations, goals, plans, decisions, etc. Philosophers try to explicate most of that brain activity in terms of reasoning. The bulk of this reasoning is defeasible, not deductive.

Allow me to list[1] some reasoning forms that are unavoidably defeasible. One first thinks of all kinds of inductive reasoning [Batens, 2004; Batens, 2005; Batens, 2011; Batens, Haesaert, 2003; Meheus, 2004], including inductive generalization as well as all predictions derived from the obtained generalizations. There is also abductive reasoning, with its ties to explanations of sorts [Batens, 2017; Beirlaen, Aliseda, 2014; Lycke, 2012; Meheus, 2007; Meheus, 2011; Meheus, Batens, 2006; Meheus et al., 2002; Gauderis, Van De Putte, 2012]; but just as well approaches to explanation that do not rely on abduction [Batens, 2005; Batens, Meheus, 2001; Weber, De Clercq, 2002; Weber, Van Dyck, 2001]. A very different topic is compatibility, including inconsistent compatibility. Even finding out whether, in general, a predicative set of statements is inconsistent or not, or whether two predicative sets are incompatible with each other or not requires defeasible reasoning [Batens, Meheus, 2000; Meheus, 2003; Meheus, Provijn, 2004]. Further examples concern the logic of questions [De Clercq, Verhoeven, 2004; Meheus, 2001], handling deontic conflicts [Beirlaen, Straßer, 2013a; Beirlaen, Straßer, 2013b; Goble, 2014; Meheus et al., 2010a; Meheus et al., 2010b; Straßer, 2010; Straßer et al., 2012; Van De Putte et al., in press; Van De Putte, Straßer, 2012] and many more. A whole different family are corrective adaptive logics, like the one for handling inconsistency, started in the 1980s [Batens, 1985; Batens, 1986; Batens, 1989] and having resulted in too many papers to refer to in the present context, and those handling ambiguity [Batens, 2002; Vanackere, 1999a; Vanackere, 1999b; Vanackere, 2000; Vanackere, 2001].

[1]The interspersed references are incomplete, even with respect to adaptive logics proposed for handling the topics.

The adaptive logics programme is one of the attempts to unify all sensible and useful defeasible reasoning. It is rather easy to devise a manifold of model-theoretic, procedural, and other systems that define defeasible reasoning forms that no one could possibly unify. All those systems may prove to be interesting and even useful mathematical structures in some more or less distant future. They may also turn out idle tea table talk. So I propose to spend a reasonable part of our efforts to defeasible reasoning forms that are known to be sensible and useful.

Adaptive logics in standard format — see Section 2. — form a unifying structure that is simple and formal. This requires some comments. The relation between the premises and the conclusion of defeasible reasoning is known to be complex. If the explication in terms of adaptive logics is right, as present insights suggest, the complexity of the consequence relation if up to Π^1_1-complex [Batens et al., 2009; Horsten, Welch, 2007; Odintsov, Speranski, 2012; Odintsov, Speranski, 2013; Verdée, 2009]. Yet the ideas behind the semantics are transparent and unsophisticated. Moreover, there are dynamic proofs. In some cases, the proofs only stabilize at an infinite point — an unavoidable effect of the complexity of the consequence relation. Yet the finite proof stages offer arguably a *sensible estimate*, in view of the information revealed by the stage, of the result obtained when the proof stabilizes — this is called *final derivability*. And indeed, proof stages are constructed by simple means. All rules are *finitary* — unlike for, for example, second order logic. And which lines are IN or OUT in the any given stage of the dynamic proof is decidable. So this basically reflects the human condition: drawing conclusions from the available information is rather unproblematic, but we know this information to be partial and presumably misguided.

I stated that adaptive logics form a *formal* unifying structure. This means what it always meant: that inferences are correct in view of their *form*. This does *not* entail, as some simpletons actually expect, that Uniform Substitution (US) holds. US does not even work for full Classical Logic, **CL**.[2] But a different formal criterion strictly obtains; my preferred name for it is *bijective uniformity*. Technicalities aside, two inference statements $\Gamma \vdash A$ and $\Delta \vdash B$ have the same *characteristic form* iff each of them can be obtained from the other by systematically replacing a referring term by another referring term of the same sort — an individual constant by an individual constant, a predicate of rank r by a predicate of rank r, etc. The result is that, for example, even the propositional inference statements $\neg p \wedge q, p \vee r \vdash r$ and $\neg p \wedge p, p \vee r \vdash r$ do *not*

[2]The closest that comes to it is, to the best of my knowledge, still reported by Witold Pogorzelski and Tadeusz Prucnal [Pogorzelski, Prucnal, 1975]; enjoy.

have the same characteristic form because the former cannot be obtained from the latter by any such *systematic* replacement.

Given the importance of defeasible reasoning, and hence of adaptive logics in standard format as candidates for the unification, it is essential to delineate ways to *devise* adaptive logics. A general feature about defeasible reasoning is that it capitalizes on the fact that a certain feature or situation is *normal* in the sense of frequently occurring, whereas abnormal features or situations are exceptional. This leads to the idea to consider certain conclusions are justified in view of the presumed absence of abnormality. Most studied adaptive logics were obtained by first delineating the set of *abnormalities*, which is characterized by a certain logical form. Thus, even if it turns out that a theory (or data set) requires $\exists x(Px \wedge Qx) \wedge \exists x(Px \wedge \neg Qx)$ to be true, one may still presume that $\exists x(Px \wedge Rx) \wedge \exists x(Px \wedge \neg Rx)$ is false.

Next, one studies which inferences are defeasibly correct, that is correct in view of the presumed falsehood of certain abnormalities. Clearly, $\exists x(Px \wedge Rx) \vdash_{\mathbf{CL}} \forall x(Px \supset Rx) \vee (\exists x(Px \wedge Rx) \wedge \exists x(Px \wedge \neg Rx))$. So if one may, reasoning systematically, consider $\exists x(Px \wedge \neg Rx)$ as false, and one knows that $\exists x(Px \wedge Rx)$ is true, one may conditionally derive $\forall x(Px \supset Rx)$. The justification will go as follows. From the true $\exists x(Px \wedge Rx)$ follows $\forall x(Px \supset Rx) \vee (\exists x(Px \wedge Rx) \wedge \exists x(Px \wedge \neg Rx))$. The second disjunct of the conclusion is an abnormality, which we presume to be false and this presumption can be upheld. So, in the light of present insights, $\forall x(Px \supset Rx)$ is true. Needless to say, this is merely an intuitive description. The matter will be phrased precisely in Section 2. and references to proofs will be given there.

So the traditional approach was to start from a set of *abnormalities* and next to study which defeasible inferences are correct if certain abnormalities may be presumed to be false. As becomes clear in the next section, once we know what the abnormalities are, the relevant adaptive logic is easily defined. Adaptive logics consider abnormalities as *false* in 'normal' situations; as false until and unless proven otherwise.

Often, however, in devising an adaptive logic, one does not know from the beginning which are the abnormalities. Rather, one knows that the reasoning step A/B[3] is correct when 'nothing is *wrong*', when the situation is *normal.* Here "normal" points to a further unknown situation, the situation in which the rule A/B is valid.

Concrete examples follow in subsequent sections, but the problem is to collect general insights on the relation between the abnormalities and such a defeasible rule. Does the rule determine the set of abnormalities? Do several

[3] Rules are phrased in metalinguistic terms. So I use meta-metalinguistic variables for formulas to describe a rule.

sets of abnormalities make the rule valid as a defeasible rule? If so, what are the effects of different choices?

2. Preliminaries

Many introductions to adaptive logics are available — the most recent one is always the best [Batens, 2015]. So I shall be very brief here. Moreover, the reader may skip this section and look up things in it as she or he needs them to understand subsequent sections.

An adaptive logic, **AL**, in standard format is a triple:

(i) a *lower limit logic* **LLL**: a logic that has static proofs and has a nice semantics;[4]

(ii) a *set of abnormalities* Ω: a decidable set of formulas characterized by a (possibly restricted) logical form F; or a union of such sets;[5]

(iii) an *adaptive strategy*: Reliability or Minimal Abnormality.[6]

The *upper limit logic* **ULL** is obtained by extending the lower limit logic **LLL** with an axiom stating that all abnormalities cause triviality. Where a premise set Γ does not require any abnormalities to be true, the **AL**-consequences of Γ provably coincide with its **ULL**-consequences. One of the effects is that the inconsistency-adaptive consequences of a consistent premise set coincide with the set's **CL**-consequences.

In a 'Dab-formula' $\mathrm{Dab}(\Delta)$, Δ is a finite subset of Ω and $\mathrm{Dab}(\Delta)$ denotes the *classical* disjunction of the members of Δ. So classical disjunction needs to be present in the language or has to be added.[7]

$\mathrm{Dab}(\Delta)$ is a *minimal* Dab-*consequence of* Γ iff $\Gamma \vdash_{\mathbf{LLL}} \mathrm{Dab}(\Delta)$ whereas $\Gamma \nvdash_{\mathbf{LLL}} \mathrm{Dab}(\Delta')$ for any $\Delta' \subset \Delta$. A *choice set* of $\Sigma = \{\Delta_1, \Delta_2, \ldots\}$ is a set that contains an element out of each member of Σ. A *minimal choice set* of Σ is a choice set of Σ of which no proper subset is a choice set of Σ. Where $\mathrm{Dab}(\Delta_1), \mathrm{Dab}(\Delta_2), \ldots$ are the minimal Dab-consequences of Γ, $U(\Gamma) = \Delta_1 \cup \Delta_2 \cup \ldots$ and $\Phi(\Gamma)$ is the set of minimal choice sets of $\Sigma = \{\Delta_1, \Delta_2, \ldots\}$.

[4]Read this as a compact Tarski logic with a characteristic semantics. The idea of a nice semantics [Verdée, Batens, 2016] is more sophisticated than that of a characteristic semantics and is fascinating in view of its implications for embedding. Unfortunately, explaining it here would require too long a digression.

[5]Where $\mathcal{F}^a$ is the set of atomic formulas (those containing no logical symbols other than $=$), $\{A \wedge \neg A \mid A \in \mathcal{F}^a\}$ is an example of a restricted logical form.

[6]There are the most important strategies.

[7]As Sergei Odintsov and Stanislav Speranski first pointed out [Odintsov, Speranski, 2013], an alternative is to phrase adaptive logics in multiple conclusion terms.

Definition 1. A **LLL**-model M of Γ is *reliable* iff $\mathrm{Ab}(M) \subseteq U(\Gamma)$.

Definition 2. $\Gamma \vDash_{\mathbf{AL}^r} A$ iff A is verified by all reliable models of Γ.

Definition 3. A **LLL**-model M of Γ is *minimally abnormal* iff there is no **LLL**-model M' of Γ such that $\mathrm{Ab}(M') \subset \mathrm{Ab}(M)$.

Definition 4. $\Gamma \vDash_{\mathbf{AL}^m} A$ iff A is verified by all minimally abnormal models of Γ.

It can be shown that a **LLL**-model M of Γ is *minimally abnormal* iff $\mathrm{Ab}(M) \in \Phi(\Gamma)$.

Although I started with their semantics, adaptive logics were discovered by reflecting on dynamic proofs — the theorizing on dynamic proof theories came much later [Batens, 2009]. An annotated **AL**-proof consists of lines that have four elements: a line number, a formula, a justification (at most referring to preceding lines) and a *condition*. Where

$$A \qquad \Delta$$

abbreviates that A occurs in the proof as the formula of a line that has Δ as its condition, the (generic) inference rules are — $\check{\vee}$ is a classical disjunction:

PREM If $A \in \Gamma$:
$$\begin{array}{ll} \ldots & \ldots \\ \hline A & \emptyset \end{array}$$

RU If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B$:
$$\begin{array}{ll} A_1 & \Delta_1 \\ \ldots & \ldots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \end{array}$$

RC If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B \check{\vee} \mathrm{Dab}(\Theta)$
$$\begin{array}{ll} A_1 & \Delta_1 \\ \ldots & \ldots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \cup \Theta \end{array}$$

A proof *stage* is a list of lines obtained by applications of the generic rules PREM, RU and RC. Let the empty list be stage 0. Where s is a stage, s' is an *extension* of s iff all lines that occur in s occur in the same order in s'. A (dynamic) proof is a chain of stages. That A is derivable on the condition Δ may be interpreted as: it follows from the premise set that A or one of the members of Δ is true. Because the members of Δ, which are abnormalities, are presumed to be false, A is considered as derived, unless and until it shows that the presumption cannot be upheld. The precise meaning of "cannot be upheld" depends on the strategy, which determines the marking definition (see below) and hence determines which lines are marked at a stage. If a line is marked at a stage, its formula is considered as not derived at that stage.

$\text{Dab}(\Delta)$ is a *minimal* Dab-*formula* at stage s of an **AL**-proof iff $\text{Dab}(\Delta)$ was derived at s on the condition $\emptyset$ whereas for no $\Delta' \subset \Delta$ was $\text{Dab}(\Delta')$ derived on the condition $\emptyset$. Where $\text{Dab}(\Delta_1), \ldots, \text{Dab}(\Delta_n)$ are the minimal Dab-formulas at stage s of a proof from Γ, $U_s(\Gamma) = \Delta_1 \cup \ldots \cup \Delta_n$ and $\Phi_s(\Gamma)$ is the set of minimal choice sets of $\{\Delta_1, \ldots, \Delta_n\}$.

Definition 5. Marking for Reliability: Line l is marked at stage s iff, where Δ is its condition, $\Delta \cap U_s(\Gamma) \neq \emptyset$.

Definition 6. Marking for Minimal Abnormality: Line l is marked at stage s iff, where A is derived on the condition Δ on line l, (i) there is no $\varphi \in \Phi_s(\Gamma)$ such that $\varphi \cap \Delta = \emptyset$, or (ii) for some $\varphi \in \Phi_s(\Gamma)$, there is no line on which A is derived on a condition Θ for which $\varphi \cap \Theta = \emptyset$.

Let me rephrase this: where A is derived on the condition Δ on line l, line l is *unmarked* at stage s iff (i) there is a $\varphi \in \Phi_s(\Gamma)$ for which $\varphi \cap \Delta = \emptyset$ and (ii) for every $\varphi \in \Phi_s(\Gamma)$, there is a line at which A is derived on a condition Θ for which $\varphi \cap \Theta = \emptyset$.

A formula A is *derived at stage* s from Γ iff it is the formula of a line that is unmarked in s. Marks may come and go as the proof proceeds. Yet there also is a stable notion of derivability, called *final derivability.*

Definition 7. A is *finally derived* from Γ on line l of stage s iff (i) A is the second element of line l, (ii) line l is not marked at stage s, and (iii) every extension of the stage in which line l is marked may be further extended in such a way that line l is unmarked.

Definition 8. $\Gamma \vdash_{\mathbf{AL}} A$ (A is *finally* **AL**-*derivable* from Γ) iff A is finally derived on a line of a proof stage from Γ.

There are three comments in conclusion of the preliminaries. First, adaptive logics are not competitors of deductive logics, but means to arrive at formal characterizations of methods. Next, one typically needs adaptive logics (and, more generally, defeasible reasoning) when a positive test is absent. Consider any of the examples mentioned before. At the predicative level, the consequence set of the adaptive logics is not semi-recursive. The final comment is that adaptive logics have an impressive metatheory which required the development of novel proof methods. The metatheory includes Soundness and Completeness proofs, but also the proofs of many features that are entirely foreign to deductive logics. I refer to [Batens, 2007] for the metatheory and to [Batens, 2015] for a revised formulation of the theorems, often leaving the straightforward reformulation of the proofs to the reader.

3. The Problem

Consider a defeasible rule $\mathsf{A}_1, \ldots, \mathsf{A}_n/\mathsf{B}$ that we consider as valid in *normal* situations. In rather exceptional cases we have no precise idea of the lower limit logic **LLL**, but let us neglect that problem and suppose that the strategy as well as **LLL** are given.[8] The task is to find the form of a formula C that may serve as the abnormality for the rule, viz. such that $\mathsf{A}_1, \ldots, \mathsf{A}_n \vdash_{\mathbf{LLL}} \mathsf{B}\check{\vee}\mathsf{C}$. Once this C is found, the corresponding conditional rule CR will be: "from $\mathsf{A}_1, \ldots, \mathsf{A}_n$ on the condition Δ to derive B on the condition $\Delta \cup \{\mathsf{C}\}$".

The reader may wonder whether a single abnormality C is introduced rather than a Dab-formula, as was suggested by the way the generic rule RC was phrased in the previous section. This is an interesting point. Let us leave open whether C will be the form of the abnormalities or whether C may indeed be itself a disjunction of abnormalities. Let us also leave open whether the problem is to find a unique C or several — the latter case refers to the second alternative in the description of the set of abnormalities Ω: "or a union of such sets". The first sentence of the present paragraph moreover reminds us of an important matter. We want to delineate Ω in function of the defeasible rule $\mathsf{A}_1, \ldots, \mathsf{A}_n/\mathsf{B}$. Yet, we are after an adaptive logic in standard format. In other words, the generic rule RC will by no means be restricted to the defeasible rule $\mathsf{A}_1, \ldots, \mathsf{A}_n/\mathsf{B}$. The generic rule RC will solely depend on **LLL** and Ω as is obvious from Section 2. We shall see that this consideration will play an important role in subsequent pages.

Consider some examples of defeasible rules in the domain of inductive generalization. There are many adaptive logics in that domain. Each of them characterizes a way to defeasibly infer generalizations. A generalization is a formula $\forall x A(x)$ in which $A(x)$ is a truth function of literals in which no individual constants occur.[9] One of the logics allows one to introduce generalizations as Popperian hypotheses, the defeasible rule then becoming $-/\forall x A(x)$ — given whatever, one may conditionally introduce a generalization. Other such logics require an instance and hence need a defeasible rule $A(\alpha)/\forall x A(x)$, in which α may be any individual constant. Still, other logics require a 'positive instance'

[8]When we are after an ampliative adaptive logic, **LLL** will be the deductive logic we consider suitable in the given context. For many this will be **CL** when the context concerns empirical or classical mathematical theories.

[9]The precise formulation was published elsewhere [Batens, 2011], but is not terribly important in the present context.

as in $B(\alpha) \wedge C(\alpha)/\forall x(B(x) \supset C(x))$. In all of these A and C are disjunctions of one or more literals and B is a conjunction of literals — conjunctions of two generalizations are derived by RU from generalizations derived by RC. So, for each such defeasible rule, the task is to pinpoint an abnormality, which then will determine the set of abnormalities Ω for that logic.

The sets of abnormalities for those inductive generalization rules were delineated a long time ago by tinkering. This was not difficult and they agree nicely with the recipe that will be presented in the present paper. This is a good reason to consider a different type of adaptive logics.

It is desirable to refer to a case where the matter becomes slightly more difficult as well as slightly more interesting. While working on adaptive set theories [Batens, 2019], I came about a case that I never met before. That we are dealing with a corrective adaptive logic is a difference with the logics from the previous paragraphs. Yet, something is more important. The lower limit logic of the set theories is the paraconsistent **CLuNs**, which is specified below, and the strategy is Minimal Abnormality. The well-studied inconsistency-adaptive logic $\mathbf{CLuNs}^m$ is obtained by specifying the set of abnormalities as $\{\mathsf{Q}(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$, in which $\mathcal{F}^a$ is the set of (open and closed) atomic formulas and $\mathsf{Q}(A)$ is (A) preceded by a quantifier over every formula free in A. I give this set a specific name for future reference. It turns out that certain premise sets require a different adaptive logic, one that has a more embracing set of abnormalities and hence assigns a richer consequence set to the premise sets.

While adaptive logics were originally devised as ways to formally characterize methods, it turned out that they may also be profitably invoked to characterize complex theories — viz. theories that are not semi-recursive. Partly relying on work by others, I made attempts to devise adaptive theories for Peano Arithmetic and for Frege's notion of a set. It is the latter that led to the case I now shall outline. I'll just mention some ideas, as the paper will soon be available in print. However, there are some details I have to report explicitly in order to clarify the problem. Readers who are in a hurry may skip to the beginning of Section 4. and return here later if they get interested in the significant example.

As Frege's notion of a set makes inconsistent sets unavoidable, the lower limit logic of the adaptive logic needs to be paraconsistent. For reasons not discussed here, I choose the (very popular) paraconsistent logic (which I prefer to call) **CLuNs** [Batens, Clercq, 2004]. Apart from negation, all logical symbols are exactly as in **CL** and RoI (Replacement of Identicals) holds unrestrictedly.[10] The negation $\neg$ is strictly paraconsistent[11] and reduces complex negations to simpler ones in the usual way: $\neg\neg A \equiv A$, $\neg(A \wedge B) \equiv (\neg A \vee \neg B)$, ... and $\neg\exists x A \equiv \forall x \neg A$.

The set theory obtained by **CLuNs** from (a version of) the Fregean axiom schema Abs and axiom Ext will be called **PFS** (paraconsistent Fregean set theory).[12] Obviously, one wants to move from the paraconsistent theory to an adaptive one, call it **AFS**. While one has unavoidably to allow for some inconsistent sets — sets of which some members are also non-members — one wants that sets are only inconsistent when this is unavoidable, and one wants even inconsistent sets to behave as consistently as possible. For example, one wants $\emptyset$ to be consistent and, while the Russell set R is unavoidably inconsistent, one wants $\emptyset \notin R$ in view of $\emptyset \notin \emptyset$ and one does not want $\emptyset \in R$.

Just like the language of most mathematical theories, the language of set theory is extremely simple. Apart from the logical symbols and the variables of the standard predicative language, it has one binary predicate $\in$ and often *abstracting terms* of the form $\{\alpha \mid A(\alpha)\}$. Where the underlying logic is **CLuNs**, some formulas of this language express triviality,[13] for example $\forall x \forall y(x = y \wedge x \neq y \wedge x \in y \wedge x \notin y)$, which I shall abbreviate as $\perp$.[14] Literally every formula *of the set theoretical language* is **CLuNs**-derivable from this (as well as from some other formulas).[15] Given that material implication is present with all its **CL**-properties, classical negation can be defined: $\check{\neg} A =_{df} A \supset \perp$.

The presence of classical negation has the unexpected consequence that the Abs axiom requires the existence of $R^* =_{df} \{x \mid \check{\neg}\, x \in x\}$. While inconsistency results, $R^* \in R^* \wedge R^* \notin R^*$, it is provable that $R^* \in R^* \wedge \check{\neg} R^* \in R^*$ is not derivable and that the inconsistency-adaptive theory is non-trivial, just like the paraconsistent theory. Yet, the fact that $R^* \in R^*$ is a theorem of the

[10]RoI: $a = b \supset (A \equiv A_{a/b})$ in which $A_{a/b}$ is the result of replacing in A an occurrence of a by b or vice versa. In some paraconsistent logics, RoI does not hold within the scope of a negation.

[11]A negation $\neg$ is paraconsistent iff $A, \neg A \vdash B$ does not hold for all A and B; it is strictly paraconsistent iff there is no A such that $A, \neg A \vdash B$ holds for all B.

[12]Within **CLuNs** there are three different implications that coincide in **CL**: $A \supset B$ is detachable but not contraposable, $A \sqsupset B =_{df} \neg A \vee B$ is not detachable but contraposable, $A \rightarrow B =_{df} (A \supset B) \wedge (\neg B \supset \neg A)$ is both detachable and contraposable; similarly, there are 16 different equivalences that coincide in **CL**. So choices have to be made as one moves from Frege's trivial theory to the provably non-trivial **CLuNs**-theory **PFS**.

[13]Several other paraconsistent logics have the same property.

[14]The abbreviations $t_1 \neq t_2 =_{df} \neg t_1 = t_2$ and $t_1 \notin t_2 =_{df} \neg t_1 \in t_2$ occur for readability.

[15]The formula does not express triviality in some extensions of the language of set theory. So it is a remarkable case of expressing *local* triviality, a feature that also occurs in other mathematical theories — I shall soon publish a brief study of the remarkable phenomenon and its epistemic potential.

paraconsistent theory, and hence also of the adaptive one, reveals a perhaps unpleasant but interesting phenomenon: R^* has members that do not fulfil the *touchstone* of R^*. Indeed, $R^* \in R^*$ is a theorem of the paraconsistent theory, but R^* does not fulfil the touchstone, which is $\check{\neg} R^* \in R^*$; and it *cannot* fulfil the touchstone — the theory is non-trivial. I shall say that R^* is *clean* iff $\forall y(y \in \{x \mid A(x)\} \leftrightarrow A(y))$, in which $A \leftrightarrow B =_{df} (A \equiv B) \wedge (\neg B \equiv \neg A)$. It turns out that both $\forall y(A(y) \supset y \in \{x \mid A(x)\})$ and $\forall y(\neg A(y) \supset y \notin \{x \mid A(x)\})$ can be required to hold, but not their converses, precisely because some sets, for example R^* are unavoidably unclean. The converses have to read $\forall y(y \in \{x \mid A(x)\} \sqsupset A(y))$ and $\forall y(y \notin \{x \mid A(x)\} \sqsupset \neg A(y))$ — remember that $\sqsupset$ is not detachable.[16]

This situation reveals a problem that requires a solution. Indeed, $R =_{df} \{x \mid x \notin x\}$ is inconsistent but there is no need for it to be unclean. Let $\Omega_1 =_{df} \{\mathsf{Q}(A \wedge \neg A) \mid A \in \mathcal{F}^a\}$, in which $\mathcal{F}^a$ is the set of (open and closed) atomic formulas and $\mathsf{Q}(A)$ is (A) preceded by a quantifier over every formula free in A. Consider a **PFS**-model M that is minimally abnormal with respect to Ω_1. Obviously, the domain D of M is uncountable[17] whence some elements of D have no name — are not named by an abstracting term. It turns out that some sets are clean in some **PFS**-models that are minimally abnormal with respect to Ω_1, but are unclean in other **PFS**-models that are also minimally abnormal with respect to Ω_1. A typical example is precisely R. Consider a minimally abnormal **PFS**-model M_1 in which R is clean and consider an element $o \in D$ that stands in the $\in$-relation to R and not also in the $\notin$-relation to R — technically this will be expressed for example by $\langle o, v(R)\rangle \in v^T(\in)$.[18] Next consider a model M_2 that is exactly like M_1 except in that o is not only a member but also a non-member of R. So, in M_2, the set R is unclean as well as inconsistent. Yet, given that no individual constant of the language refers to o, the inconsistency can only be stated as $\exists x(x \in R \wedge x \notin R)$. But this formula is also verified by M_1, because all **PFS**-models verify $R \in R \wedge R \notin R$. So R is clean in M_1 and is unclean in M_2, but both are minimally abnormal and actually $\mathrm{Ab}(M_1) = \mathrm{Ab}(M_2)$. This is not as we want it. The axioms do not require R to be unclean. They do not even require that R is a member

[16] It is not really essential to this paper, but Abs comes to the conjunction of the four implications mentioned in the text, two detachable ones and two non-detachable ones.

[17] Many uncountable **ZF**-hierarchies can be defined in exactly the same way in **PFS** and if their members were countable in **PFS**, then they would be inconsistent. It can be argued that they are consistent in minimally abnormal models of **PFS** if they are consistent in **ZF**.

[18] In this semantic style, the extension of a predicate π of rank n is a triple $\langle \Sigma_1, \Sigma_2, \Sigma_3\rangle$ with $\Sigma_1, \Sigma_2, \Sigma_3 \supseteq D^n$ and $\Sigma_1 \cup \Sigma_2 \cup \Sigma_3 = D^n$. Next, for convenience, the assignment function $v(\pi)$ is seen as three functions: $v^T(\pi) = \Sigma_1$, $v^B(\pi) = \Sigma_2$ and $v^F(\pi) = \Sigma_3$.

of a minimal set of sets one of which is bound to be unclean. So R should be clean.[19]

What does this come to? We want the non-detachable $\forall y(y \in \{x \mid A(x)\} \sqsupset A(y))$ to have the strength of the detachable $\forall y(y \in \{x \mid A(x)\} \supset A(y))$ as much as possible; and similarly for upgrading $\forall y(y \notin \{x \mid A(x)\} \sqsupset \neg A(y))$ to $\forall y(y \notin \{x \mid A(x)\} \supset \neg A(y))$. Obviously, we do not want to interpret all expressions $A \sqsupset B$ as much as possible as $A \supset B$. We only want instances of schema Abs to be as much as possible detachable in all directions. Abs was intended originally as detachable in all directions. This cannot be realized completely because Frege's notion of the extension of a predicate turned out inconsistent. Nevertheless, the original intention can be realized as much as possible. But this cannot be done by minimizing inconsistencies: remember that M_1 and M_2 verify the same members of Ω_1. We have a case where, for ordered pairs $\langle A, B\rangle$ *of certain forms*, we want to derive $A \supset B$ from $A \sqsupset B$ on a certain condition. The task is to find the condition.

4. Solving the Problem

Let us concentrate first on rules with one local premise. The task then is, starting from the defeasible rule A/B, to find a condition or conditions C such that the three following hold:

$$\mathsf{A} \vdash_{\mathbf{LLL}} \mathsf{B}\check{\vee}\mathsf{C} \quad \text{and} \quad \mathsf{A} \nvdash_{\mathbf{LLL}} \mathsf{B} \quad \text{and} \quad \mathsf{A} \nvdash_{\mathbf{LLL}} \mathsf{C}\,. \tag{1}$$

If $\mathsf{A} \vdash_{\mathbf{LLL}} \mathsf{B}$, then A/B is **LLL**-valid and not a defeasible rule. If $\mathsf{A} \vdash_{\mathbf{LLL}} \mathsf{C}$, then the condition C causes the rule A/B to have no sensible applications. Remember indeed that, as stated in Section 3., the idea is to obtain the following particular conditional rule: "from A on the condition Δ to derive B on the condition $\Delta \cup \{\mathsf{C}\}$". If the defeasible rule A/B is to have sensible applications, there must be a premise set Γ and a (empty or non-empty) $\Delta \subset \Omega$ such that (i) A is finally derivable from Γ on the condition Δ and (ii) B is finally derivable from Γ on the condition $\Delta \cup \{\mathsf{C}\}$. In view of (i), $\Gamma \vdash_{\mathbf{LLL}} \mathsf{A}\check{\vee}\mathrm{Dab}(\Delta)$ and there is a $\varphi \in \Phi(\Gamma)$ such that $\Delta \cap \varphi = \emptyset$. In view of (ii), $\Gamma \vdash_{\mathbf{LLL}} \mathsf{B}\check{\vee}\mathrm{Dab}(\Delta \cup \{\mathsf{C}\})$ and there is a $\varphi \in \Phi(\Gamma)$ such that $(\Delta \cup \{\mathsf{C}\}) \cap \varphi = \emptyset$. The latter is impossible because, if $\mathsf{A} \vdash_{\mathbf{LLL}} \mathsf{C}$, then $\Gamma \vdash_{\mathbf{LLL}} \mathsf{A}\check{\vee}\mathrm{Dab}(\Delta)$ is a sufficient condition for $\Gamma \vdash_{\mathbf{LLL}} \mathsf{C}\check{\vee}\mathrm{Dab}(\Delta)$, which is $\Gamma \vdash_{\mathbf{LLL}} \mathrm{Dab}(\Delta \cup \{\mathsf{C}\})$. Whether $\mathrm{Dab}(\Delta \cup \{\mathsf{C}\})$ is a minimal Dab-consequence of Γ or not, every $\varphi \in \Phi(\Gamma)$ contains at least one

[19] The difference between M_1 and M_2 cannot be expressed in the language by a formula stating a contradiction, whether plain or quantified. Does this mean that the difference between M_1 and M_2 cannot be expressed? By no means. We just need different abnormalities.

member of $\text{Dab}(\Delta \cup \{\mathsf{C}\})$. But then the line at which B is derived by A/B is always marked. This ends the justification of the requirements in (1).[20]

Let us move to a concrete case, viz. the defeasible rule

$$A \sqsupset B / A \supset B \,. \tag{2}$$

We are looking for one or more conditions such that the three derivability statements in (1) are fulfilled. I shall first consider the problem in the context of the lower limit logic **CLuNs**, with the classical negation $\check{\neg}$ present or added, and the Minimal Abnormality strategy, but neglecting for the moment that the problem arose in connection with the set theory **AFS**.

In the previous section, (2) was considered in a situation in which the abnormalities were contradictions, as is usual for inconsistency-adaptive logics. Some people will keep repeating that abnormalities of the form $A \wedge \neg A$ justify the defeasible rule (2). Indeed,

$$A \sqsupset B \vdash_{\mathbf{CLuNs}} (A \supset B) \check{\vee} (A \wedge \neg A)$$

holds. Or, even more explicitly in view of **CLuNs**-equivalences,

$$\neg A \vee B \vdash_{\mathbf{CLuNs}} (\check{\neg} A \vee B) \check{\vee} (A \wedge \neg A) \,.$$

However, and as already explained in Section 3., this is not the point.[21] The point is that, for specific ordered pairs $\langle A, B \rangle$, we want (2) to be applied *even if* $A \wedge \neg A$ *is true*.[22]

We are in search of conditions C that fulfill (1) and, by their forms, determine a set Ω that, together with the lower limit logic **CLuNs** and the Minimal Abnormality strategy, defines an adequate adaptive logic. We need a C such that $\mathsf{A} \vdash_{\mathbf{CLuNs}} \mathsf{B} \check{\vee} \mathsf{C}$. Given that **CLuNs** has a nice semantics and given Soundness and Completeness, $\mathsf{A} \vdash_{\mathbf{CLuNs}} \mathsf{B} \check{\vee} \mathsf{C}$ is equivalent to $\mathsf{A} \check{\wedge} \check{\neg} \mathsf{B} \vdash_{\mathbf{CLuNs}} \mathsf{C}$. The strongest such C is obviously $\mathsf{A} \check{\wedge} \check{\neg} \mathsf{B}$ itself and every such C is a **CLuNs**-consequence of $\mathsf{A} \check{\wedge} \check{\neg} \mathsf{B}$.

Let us apply this at once to the defeasible rule (2). The strongest condition C is $(A \sqsupset B) \wedge \check{\neg}(A \supset B)$, which is **CLuNs**-equivalent to $(A \wedge \neg A \wedge \check{\neg} B)$ — as the conjunction is classical in **CLuNs**, there is no need to write $\check{\wedge}$. So the defeasible rule phrased with its strongest condition reads: "from A on the condition Δ to derive B on the condition $\Delta \cup \{A \wedge \neg A \wedge \check{\neg} B\}$". Actually, the

[20]The justification considers only Minimal Abnormality. Where Reliability is the strategy, the justification is much simpler and left to the reader.

[21]Moreover and concerning **AFS**, every unclean set is unavoidably inconsistent: if $t \in \{x \mid A(x)\}$ but $\neg A(t)$, then $t \notin \{x \mid A(x)\}$. However, this too is not the point.

[22]In **AFS** we want all sets to be as clean as possible, even inconsistent sets.

logic **CLuNs** requires[23] that the set of abnormalities contains all formulas of the form $A \wedge \neg A \wedge \check{\neg} B$ in which A is an atomic formula and B is a literal.[24]

Next, consider weaker conditions C for the defeasible rule (2). Let us have a systematic look at 'parts' of $A \wedge \neg A \wedge \check{\neg} B$. The idea is not to find a C that in itself gives us all we want, but to find conditions C that are acceptable, possibly in the presence of other conditions. I neglect the fact that the local premise may come on a condition itself; by now, the reader will have understood the resulting complication. To the left is the **CLuNs**-inference, to the right the effect on a dynamic proof.

$$\frac{A \sqsupset B}{(A \supset B) \vee \mathsf{C}} \qquad\qquad \frac{A \sqsupset B}{A \supset B \quad \{\mathsf{C}\}}$$

Let us consider the possibilities systematically.

(i) We know already that the strongest C is $A \wedge \neg A \wedge \check{\neg} B$.

(ii) That C is $A \wedge \neg A$ is all right *provided* one also wants all conditional inferences that then are correct in view of the standard format, specifically RC. An example is the effect of $A \supset B, \neg B \vdash_{\mathbf{CLuNs}} \neg A \vee (B \wedge \neg B)$: from $A \supset B$ and $\neg B$ to derive $\neg A$ on the condition $\{B \wedge \neg B\}$. So what this comes to is that the choice $A \wedge \neg A$ is all right in case one agrees that $\neg A$ has actually the force of $\check{\neg} A$ whenever $A \wedge \neg A$ can be taken to be false.

(iii) That C is $A \wedge \check{\neg} B$ is not acceptable. Indeed, this condition is simply the classical negation of the conclusion of the defeasible rule. Once $\check{\neg}$ is added to **CLuNs**, $(A \supset B) \vee (A \wedge \check{\neg} B)$ is a **CLuNs**-theorem. So if $A \wedge \check{\neg} B$ is an abnormality, possibly with A restricted to atomic formulas and B to literals, then $A \supset B$ is derivable on the condition $\{A \wedge \check{\neg} B\}$ from any premise set. Unlike what the reader might expect, this would not cause premise sets to have trivial consequence sets; most conditional lines would be marked. Yet, there is no sensible idea behind this choice of an abnormality and the choice does not seem to lead to anything sensible. Nevertheless, I shall return to this choice below.

---

[23] The requirement is related to the avoidance of so-called flip-flop adaptive logics, which are only desirable for specific applications [Batens, 2007]. The point need not further concern us here.

[24] The set of literals is the set of non-equivalent formulas in which occurs an atomic formula preceded by at most unary connectives. Where two negations, $\neg$ and $\check{\neg}$, are present in the language of **CLuNs**, the notion of a literal is a trifle more sophisticated than in **CL**. While this set is $\{A, \neg A\}$ (A a sentential letter) in **CL**, it is $\{A, \neg A, \check{\neg} A, \check{\neg}\neg A, \check{\neg}\check{\neg} A, \check{\neg}\check{\neg}\neg A\}$ (A a sentential letter) in **CLuNs**.

(iv) Somewhat unexpectedly, it seems all right at first sight to choose $\neg A \wedge \check{\neg} B$ as C. Indeed, when one concentrates on the defeasible rule we are considering here, the choice seems unobjectionable, both in case $\neg A \wedge \check{\neg} B$ is true and in case it is false. In the latter case, for example, we obtain: if $\neg A$ is false, the premise warrants that B is true; if $\check{\neg} B$ is false, then B is also true. As B is true, so is $A \supset B$. And yet this choice has consequences we do not want. As $\neg A \vdash_{\mathbf{CLuNs}} B \vee (\neg A \wedge \check{\neg} B)$, the choice would justify that one would derive an arbitrary B from $\neg A$ on the sole condition that $\neg A \wedge \check{\neg} B$ can be taken to be false. As in the previous case, triviality would not result[25] but there is no idea behind this way of proceeding and nothing sensible is expected to result.

(v) To choose A or $\neg A$ or B as C is obviously unacceptable. That literals would be abnormalities, would result in all kinds of turmoil, but in nothing sensible.

(vi) I promised to return to (iii). Choosing $(A \sqsupset B) \wedge (A \wedge \check{\neg} B)$ as C prevents one to introduce detachable implications from the blue. Moreover, these abnormalities nicely express that the premise is true and the desired conclusion false. However, nothing new is arrived at along this road. The chosen abnormality is **CLuNs**-equivalent to $A \wedge \neg A \wedge \check{\neg} B$, which is the abnormality considered in (i).

No other choices are worth commenting upon. Yet it is still interesting to consider combinations, viz. that formulas of different forms are counted as abnormalities, for example $A \wedge \neg A$ and $A \wedge \neg A \wedge \check{\neg} B$. Neglecting some complications, a line is unmarked and its formula is not considered as derived iff its condition can be considered to be false. Suppose that $A \wedge \neg A$ cannot be considered as false. Then it is nevertheless possible that $A \wedge \neg A \wedge \check{\neg} B$ can be considered as false: if A and $\neg A$ are both true, but $\check{\neg} B$ is false, then the conjunction of the three formulas is false.[26] So allowing for abnormalities of both forms has the following effect — I keep restricting attention to crucial insights. On the one hand, including formulas $A \wedge \neg A$ in the set of abnormalities has the effect that a lot of further conditional inferences become valid, as was explained in (ii). On the other hand, even if $A \supset B$ cannot be seen as derived on the condition $A \wedge \neg A$ because this condition cannot be considered as false,

[25] Adaptive logics in standard format have the Strong Reassurance property (also called Stopperedness or Smoothness): if a premise set has **LLL**-models, then it has minimally abnormal models. Proofs were published long ago [Batens, 2000; Batens, 2007].

[26] Spelling the matter out in a precise way for Reliability and (especially) for Minimal Abnormality is much more complicated, but the crucial insight is the one stated in the text.

it is possible that $A \supset B$ can be seen as derived on the condition $A \wedge \neg A \wedge \check{\neg} B$ because $\check{\neg} B$ can still be considered as false.

The matter seems clarified, but there are still two little problems. I comment on them in order to illustrate the complications involved in the systematisation of defeasible reasoning. The easier problem is this: $A \wedge \neg A \wedge \check{\neg} A$ obviously has the form of $A \wedge \neg A \wedge \check{\neg} B$ and is *always* false.[27] That looks frightening. As the line will never be marked, such an adaptive logic seems to extend the lower limit logic with non-defeasible steps. However, this is a pseudo-problem. Whenever $\mathsf{A} \vdash_{\mathbf{CLuNs}} \mathsf{B} \vee \mathsf{C}$ and $\mathsf{C} \in \Omega$ is logically false, then $\mathsf{A} \vdash_{\mathbf{CLuNs}} \mathsf{B}$. So logically impossible abnormalities are harmless; they are obviously also useless.

The second problem is more interesting: if formulas of the form $A \wedge \neg A \wedge \check{\neg} B$ are abnormalities, does it then even make a difference whether formulas of the form $A \wedge \neg A$ are also abnormalities? While $\neg A \nvdash_{\mathbf{CLuNs}} \check{\neg} A \vee (A \wedge \neg A \wedge \check{\neg} B)$, it holds that $\neg A, \check{\neg} B \nvdash_{\mathbf{CLuNs}} \check{\neg} A \vee (A \wedge \neg A \wedge \check{\neg} B)$. So, if any formula of the form $\check{\neg} B$ is derivable, even if only conditionally, an unexpected effect seems to follow. Let me show this by presenting a little proof.

$\vdots$	$\vdots$	$\vdots$	$\vdots$
51	$\check{\neg} s$	$\ldots$	Δ
52	$\neg p$	PREM	$\emptyset$
53	$\check{\neg} p$	51, 52; RC	$\Delta \cup \{p \wedge \neg p \wedge \check{\neg} s\}$

Supposing that line 51 is unmarked, line 53 will be unmarked just in case $p \wedge \neg p$ can be held to be false. If a formula of the form $\check{\neg} B$ is derivable, even conditionally, from the premises, then abnormalities of the form $A \wedge \neg A$ are redundant.

The matter becomes less surprising if one realizes that conditional transition from $\neg p$ to $\check{\neg} p$ may be realized in a way that seems unobjectionable. Recall that $A \sqsupset B$ abbreviates $\neg A \vee B$.

$\vdots$	$\vdots$	$\vdots$	$\vdots$
51	$\check{\neg} s$	$\ldots$	Δ
52	$\neg p$	PREM	$\emptyset$
53	$p \sqsupset s$	52; RU	$\emptyset$
54	$p \supset s$	53; RC	$\{p \wedge \neg p \wedge \check{\neg} s\}$
55	$\check{\neg} p$	51, 54; RU	$\Delta \cup \{p \wedge \neg p \wedge \check{\neg} s\}$

What happens here is that we apply the defeasible rule (2) at line 54 and next apply Modus Tollens — this is correct as $\supset$ is detachable and $\check{\neg}$ is classical negation.

[27]Many will not care about the detail, but it is more correct to say that $A \wedge \neg A \wedge \check{\neg} A$ has no non-trivial models — in some semantic styles no models, in others only the trivial model.

If a bottom constant $\perp$ is present in the applied version of **CLuNs**, the matter is even easier.

1	$\neg p$	PREM	$\emptyset$
2	$\check{\neg} p$	1; RC	$\{p \wedge \neg p \wedge \check{\neg}\perp\}$

Here $\vdash_{\mathbf{CLuNs}} \check{\neg}\perp$ whence $p \wedge \neg p \wedge \check{\neg}\perp$ is **CLuNs**-equivalent to $p \wedge \neg p$.

After discussing (2) independently of the set theoretical context, let us now return to the problem in **AFS**. Two insights are important. (i) If, reasoning systematically, it is possible to consider $A \wedge \neg A$ as false, then it is possible to consider $A \wedge \neg A \wedge \check{\neg} B$ as false, but not conversely. So, given a premise set Γ, the more complex condition allows for final consequences of Γ that are not final consequences if all abnormalities have the form $A \wedge \neg A$. (ii) As noted in Section 3., the aim is not to upgrade expressions $A \sqsupset B$ as much as possible to $A \supset B$, but to do so only for specific formulas that are implicative parts of Abs, viz. $\forall y(y \in \{x \mid A(x)\} \sqsupset A(y))$ and $\forall y(y \notin \{x \mid A(x)\} \sqsupset \neg A(y))$. Consider the formulas

$$\mathsf{Q}(t \in \{x \mid A(x)\} \wedge \check{\neg} A(t)) \quad \text{and} \quad \mathsf{Q}(t \notin \{x \mid A(x)\} \wedge \check{\neg}\neg A(t)) \tag{3}$$

in which t is a set theoretical term and, if it is a variable, Q is a quantifier over that variable. In view of Abs, the left formula **PFS**-entails $\mathsf{Q}(t \notin \{x \mid A(x)\} \wedge t \in \{x \mid A(x)\} \wedge \check{\neg} A(t))$ and the right formula **PFS**-entails $\mathsf{Q}(t \in \{x \mid A(x)\} \wedge t \notin \{x \mid A(x)\} \wedge \check{\neg}\neg A(t))$. So a bit of calculation shows that formulas of the forms in (3) may be safely taken as abnormalities. Proceeding thus, we obtain an adaptive logic that minimizes inconsistencies in view of abnormalities of the form $A \wedge \neg A$ and minimizes uncleanness in view of the abnormalities from (3).

The conclusion then is as follows. First, unless there are very special reasons to refrain from applying the defeasible rule in its full generality, the recipe leads to the following schema, considering multiple local premises (but still restricting to the lower limit **CLuNs**).

$$\begin{array}{ll} \mathsf{A}_1 & \Delta_1 \\ \vdots & \vdots \\ \mathsf{A}_n & \Delta_n \\ \hline \mathsf{B} & \Delta_1 \cup \ldots \cup \Delta_n \cup \{\mathsf{A}_1 \wedge \ldots \wedge \mathsf{A}_n \wedge *\mathsf{B}\} \end{array} \tag{4}$$

Where $*\check{\neg} B = B$ and $*B = \check{\neg} B$ in case the first symbol in B is not $\check{\neg}$. Call $\mathsf{A}_1 \wedge \ldots \wedge \mathsf{A}_n \wedge *\mathsf{B}$ the *typical abnormality* for the defeasible rule $\mathsf{A}_1, \ldots, \mathsf{A}_n/\mathsf{B}$. There is no reason to prefer abnormalities obtained by dropping one or more conjuncts of $\mathsf{A}_1 \wedge \ldots \wedge \mathsf{A}_n \wedge *\mathsf{B}$ because if the shorter formula can be considered as false, then so can the longer one. When one wants to introduce several defeasible rules, each of them may be given its typical abnormality. Two points of attention then are: (i) one should check which rules need to hold in their formal generality and which should be restricted to specific formulas — the **AFS** case illustrates this candidly — and (ii) one should study the effect of the typical abnormality of a rule on other rules in view of the logical and conceptual context — see the example discussed above.

The occurrence of the classical negation $\check{\neg}$ in the typical abnormality may look like causing trouble in paraconsistent contexts, especially for dialetheists. The matter will be discussed in Section 5.

When the lower limit logic is not **CLuNs**, the typical abnormality is easily adjusted. Two adjustments may be required. (i) Sometimes restrictions on the subformulas of abnormalities need to be modified or removed. Thus the restriction $A \in \mathcal{F}^a$ needs sometimes to be replaced, for example by $A \in \mathcal{F}$, in which $\mathcal{F}$ is the set of (open and closed) formulas of standard predicative languages. (ii) Sometimes logical symbols are classical within the considered lower limit logic, whence we do not need $\check{\neg}$ and $*$ — I actually applied this already within the present paper. I do not enter this any further as the matter is mainly technical.

Let us consider some typical abnormalities for other rules. I mentioned three defeasible rules of inductive generalization. In each of them A, B and C are disjunctions of literals and there are some further restrictions. The typical abnormalities for each of the rules can be read off below in somewhat simplified form — as the lower limit is **CL**, the standard negation is classical:

$$\begin{array}{ll} - & \\ \hline \forall x A(x) & \{\exists x \neg A(x)\} \end{array} \qquad \begin{array}{ll} \exists x A(x) & \Delta \\ \hline \forall x A(x) & \Delta \cup \{\exists x A(x) \wedge \exists x \neg A(x)\} \end{array}$$

$$\begin{array}{ll} \exists x (B(x) \wedge C(x)) & \Delta \\ \hline \forall x (B(x) \supset C(x)) & \Delta \cup \{\exists x (B(x) \wedge C(x)) \wedge \exists x \neg (B(x) \wedge \neg C(x))\} \end{array}$$

An inconsistency can be seen as a negation glut: that $v_M(A) = 1$ justifies $v_M(\neg A) = 0$ on the **CL**-semantics, but actually $v_M(\neg A) = 1$. A negation gap is where $v_M(A) = 0$, which justifies $v_M(\neg A) = 1$ on the **CL**-semantics, and nevertheless $v_M(\neg A) = 0$. Along this line, one may consider gluts and gaps for every logical symbol of the standard predicative language. Adaptive logics minimizing gaps and gluts were studied [Batens, 2016]. Suppose that $\dashrightarrow$ is a glutty implication, whereas $\supset$ is the classical implication. A glutty implication is obviously not detachable, as there are models in which $v_M(A \dashrightarrow B) = 1 = v_M(A)$ and $v_M(B) = 0$. This brings us to something very close to (2). In order to minimize implication gluts, we want the following rule and abnormality:

$$\begin{array}{ll} A \dashrightarrow B & \Delta \\ \hline A \supset B & \Delta \cup \{(A \dashrightarrow B) \wedge A \wedge \neg B\} \end{array}$$

If glutty implications are combined with glutty negations (and possibly more oddities), the negation needs to be replaced by $\check{\neg}$. All glutty and gappy logical symbols may be given defeasible rules to minimize them and the insight gained in this section will provide the rules with typical abnormalities.

5. A Puzzle In Inconsistency-Adaptive Logics

Classical negation occurs in the typical abnormalities from the previous section. Some will see this as problematic in paraconsistent contexts. Of course, if classical negation is not definable in a paraconsistent logic, one may add it, possibly forbidding its occurrence in premises and conclusion. Yet especially dialetheists will have objections to such move as they consider classical negation as a tonk-like operator. This conclusion is related to the dialetheist view that all true knowledge should form a single body, phrased within a single language and organized by The True Logic and that this body is necessarily inconsistent in view of the Liar paradox, paradoxes of set theory, etc. I shall not discuss the dialetheist position here, but rather argue that, for two reasons, the typical abnormalities do not lead to a situation that is at odds with dialetheism.

The first reason is that, due to the structural properties and functioning of negation, the typical abnormalities do not require that classical negation ever occurs either in them or elsewhere in a proof. First of all, look at two basic defeasible rules for negation:[28]

$$\begin{array}{ll} \neg A & \Delta \\ \hline \check{\neg} A & \Delta \cup \{A \wedge \neg A\} \end{array} \qquad\qquad \begin{array}{ll} A & \Delta \\ \hline \check{\neg}\neg A & \Delta \cup \{A \wedge \neg A\} \end{array} \tag{5}$$

However, once the adaptive logic is characterized in terms of the Standard Format, these rules need not be mentioned. Applications of the generic conditional rule RC may be phrased completely in the standard language, without ever writing a classical negation. Here are two examples, an application of Disjunctive Syllogism and one of Modus Tollens.

$$\begin{array}{ll} A \vee B & \Delta \\ \neg A & \Theta \\ \hline B & \Delta \cup \Theta \cup \{A \wedge \neg A\} \end{array} \qquad\qquad \begin{array}{ll} A \supset B & \Delta \\ \neg B & \Theta \\ \hline \neg A & \Delta \cup \Theta \cup \{B \wedge \neg B\} \end{array}$$

The classical negation in (5) signifies that, provided the abnormality introduced by the condition can be held to be false, A, respectively $\neg A$, can be considered as consistently false; spelled out, $\check{\neg} A$ signifies that A is consistently

[28] Sometimes the A in the abnormality is restricted, for example to atomic formulas, as is required when **CLuNs** is the lower limit logic. Sometimes several defeasible rules are required as in **AFS**.

false and $\check{\neg}\neg A$ that A is consistently true. In the application of Disjunctive Syllogism, if the local premises are true and A is consistent, then B is bound to be true. That A is consistent is nowhere stated. The fact that the conclusion line is unmarked indicates that (the members of $\Delta \cup \Theta$ as well as) $A \wedge \neg A$ can be held to be false, which means that A is consistent and in that case B is bound to be true if the local premises are true.[29] The reasoning is similar for the application of Modus Tollens, except that here the consistency of B matters. The situation is analogous for all applications of the generic conditional rule RC in $\mathbf{CLuNs}^m$ and similar inconsistency-adaptive logics.

There is a second reason why the dialetheist should not eschew adaptive logics. The typical abnormality as defined in Section 4. works not only with classical negation, but works equally well with a paraconsistent negation. So where the symbol $\check{\neg}$ (defining $*$) in (4) is a negation that is paraconsistent and not also paracomplete, (4) still works fine: if all members of the condition can be held to be false, the conclusion follows from the premises.

This comment does not concern (5), the basic rule for negation. This rule, or rather both of them, are still unacceptable for the dialetheist because in it the symbol $\check{\neg}$ in the conclusion needs to be classical. As explained, however, there is no need for $\check{\neg}$ to occur anywhere in inconsistency-adaptive logics. Yet there still is a catch. Suppose that the dialetheist position gets generally recognized, that the methodology of the sciences is spelled out in terms of, say, the **LP**-negation [Priest, 1987], and that scientists would actually apply **LP** rather than requiring, presupposing and sometimes pretending that their theories be consistent, then the dialetheist might phrase the whole scientific methodology in terms of adaptive logics based on **LP**. If the condition is false, the local conclusion will follow from the local premises. A hindrance for dialetheists will be that, in the preceding sentence, "false" needs to have the meaning with which I use it: consistently false, not false as meant by Graham Priest [Priest, 1987]. The latter meaning is that A is false iff $\neg A$ is true; this entails that A and $\neg A$ are both false in case they are both true, as may happen in paraconsistent contexts. The situation seems rather crucial. *All* instances of $A \wedge \neg A$ and *all* instances of $A \wedge \neg A \wedge \check{\neg} B$ are false in the sense of Priest. There is no point in asking whether they can be held to be false in view of the premises. They are false in Priest's sense, now, yesterday, tomorrow and *always* because their negation is true, even logically true: $\neg(A \wedge \neg A)$ is an **LP**-theorem; it is **LP**-equivalent to $\neg A \vee A$. From here on I return to my use of false.

Just for the record, a comment on two related negation-like entities. A paracomplete negation, according to which A and its negation may be jointly false,

[29] Obviously, from $\neg A \vee B$ and A follows B on the condition $A \wedge \neg A$.

is insufficient for adaptive logics to work decently. If the $\check{\neg}$ in (4) is paracomplete, the falsehood of the condition is insufficient for the local conclusion to follow from the local premises. The second negation-like entity is the arrow-bottom construction, $A \to \bot$, in which $\to$ is a detachable implication and $\bot$ a bottom operator.[30] This is obviously a kind of negation of A. Dialetheists have argued that $A \to \bot$, for $\to$ a relevant implication,[31] allows them, just as much as the classical logician or intuitionist, to commit themselves to the falsehood of a certain statement A in that $A \to \bot$ connects the truth of A to triviality.

Some paragraphs ago, I argued that there is a problem for dialetheists to apply adaptive logics. Quite unexpectedly, however, there seems to be a way out. I am not a dialetheist, recently I even got doubts on the viability of dialetheism. Yet, these doubts are not related to what follows. Consider the following rules and their typical abnormalities:

$$\frac{\neg A \qquad \Delta}{A \to \bot \quad \Delta \cup \{\neg A \wedge \neg(A \to \bot)\}} \qquad \frac{A \qquad \Delta}{\neg A \to \bot \quad \Delta \cup \{A \wedge \neg(\neg A \to \bot)\}} \tag{6}$$

Dialetheists claim that true inconsistencies are exceptional. So, in non-exceptional situations, that $\neg A$ is given justifies one to defeasibly connect A to triviality and that A is given justifies one to defeasibly connect $\neg A$ to triviality.

The typical abnormality may look problematic, but it is not. For most relevant implications, $\neg(A \to B)$ is derivable from $A \wedge \neg B$. Where this is the case, $A \wedge \neg A$ is sufficient to derive both $\neg A \wedge \neg(A \to \bot)$ and $A \wedge \neg(\neg A \to \bot)$ because $\neg\bot$ is a theorem of **LP**. To prevent readers from getting overoptimistic, let me point out that the 'negation' $\ominus$, defined by $\ominus A =_{df} A \to \bot$, is a paracomplete negation. Clearly, $A \vee \ominus A$ is not a theorem unless the relevant $\to$ is downgraded to a detachable material implication.

It seems to me that the rules and abnormalities in (6) look extremely interesting from a dialetheist point of view. They allow dialetheists to *express* their commitment to the falsehood of a statement in the sense that the falsehood of A connects the truth of A to triviality. Moreover, they may do so without ever using classical negation — dialetheist may continue to catalogue that as a tonk-like operator. So (6) seems to provide a means for dialetheists to apply an inconsistency-adaptive logic without committing themselves to classical negation. Exploring the consequences of this insight obviously deserves a careful study, but that goes beyond the present paper. Moreover, adaptive

[30] A bottom operator is characterized by the rule "from $\bot$ to derive A".

[31] The intended relevant implications are not those from the well-known and very rich systems devised by Ackermann [Ackermann, 1956], Church [Church, 1951] and especially Anderson and Belnap [Anderson, Belnap 1975; Anderson et al., 1992] but of much weaker systems surveyed by Routley [Routley, 1982] and Brady [Brady, 2006].

consequences derived on a condition Δ, remain to be justified in terms of the joint falsehood, in the dialetheist sense, of the members of Δ.

6. In Conclusion

The problem I set out to solve concerned cases where one has an idea for devising an adaptive logic in terms of a defeasible rule. The easier case was where the set of abnormalities was given together with the lower limit logic for ampliative adaptive logics or together with the upper limit logic for corrective adaptive logics. If the idea for the adaptive logic comes in terms of a rule, the set of abnormalities has to be delineated. I presented an extremely simple and transparent recipe for doing so and argued that, except for very special cases, the recipe leads to an adequate result.

A further important point deserves to be mentioned. I have shown that there is a number of formerly unsolved difficulties for dialetheists who try to invoke inconsistency-adaptive logics. For me logics are instruments. Instruments may be independent of the philosophical and ideological views of those who use them. So it seems an important feature that adaptive logics as well as the proposed recipe work fine for dialetheists. Disagreements with dialetheists is not an excuse for hiding that, unlike what one might expect, adaptive logics turn out sensible and useful instruments for them.

A very different conclusion is not about generality but about specificity. There is a huge number of different adaptive logics of inductive generalization. This is not only required because of the many disagreements between philosophers of science on inductive methods. It is also necessary in view of the very different domains of application. To give just one example, the non-logical terms of one language may be well entrenched technical terms and those of another language may be taken straight from natural language. Further differences will depend on the underlying conceptual framework, on the presence of articulated observational criteria, and so on. Such differences may have an effect on the suitability of a specific inductive method. The situation for other ampliative adaptive logics is analogous.

Similar comments apply to corrective adaptive logics. The upper limit logic is known beforehand, but there are many ways to approach it: different strategies, different lower limit logics, and for each combination of a strategy and lower limit logic, different sets of abnormalities. Of course, not every specific circumstance determines a single adaptive logic. Nevertheless, the choice of a suitable adaptive logic will be largely determined by properties of the theory or domain to which it is applied. Mathematical theories have generally conceptual structures that are much simpler that most empirical theories. So they usually require a stronger lower limit logic, validating full Replacement of Identicals and reducing all statements to truth-functions of literals — truth-functions in the broad sense including quantifiers. But apart from such rough classifications, both mathematical and empirical theories will require careful analysis in order to select the specific non-logical axioms in view of the lower limit logic. Adaptive mathematical theories [Batens, 2014; Batens, 2019] are a case in point.

Part of the importance of the present paper and of the recipe is related to insights that have grown over the years. In the early days, adaptive logics seemed to present an attractive approach to handle certain problems. Examples were (i) inconsistencies coming up unexpectedly in a theory that was intended as consistent or (ii) devising a precise formulation of a given method. By and large, the impression was that adaptive logics were very general tools that could be efficiently applied in nearly all circumstances. Only over the years did it become clear that especially the choice of corrective adaptive logics depends heavily on the context. When a problem is located, adaptive logics do not provide one with a tool that in itself warrants success. One has to carefully choose a language in which to formulate the problem. One has to carefully select the way in which the theory or the data, in which the problem arises, are phrased. Recently, especially with the application to Fregean set theories (sic), it turned out that sometimes one even has to tailor the adaptive logic in view of its application. On the one hand, this shows to what extend Dudley Shapere was right in propagating content guidance and learning how to learn [Shapere, 2004]. On the other hand, it made necessary the search for the present recipe: content guidance provokes more frequently the need for adaptive logics devised in view of defeasible rules.

Acknowledgements. I want to thank the referee and the editor for their careful work.

References

Ackermann, 1956 – Ackermann, W. "Begründung einer strengen Implikation", *Journal of Symbolic Logic*, 1956, Vol. 21, pp. 113–128.

Anderson, Belnap 1975 – Anderson, A.R., Belnap, N.D. *Entailment. The Logic of Relevance and Necessity*, Vol. 1. Princeton University Press, 1975.

Anderson et al., 1992 – Anderson, A.R., Belnap, N.D., Dunn, J.M. *Entailment. The Logic of Relevance and Necessity*, Vol. 2. Princeton University Press, 1992.

Batens, 2019 – Batens, D. "Adaptive Fregean set theory", *Studia Logica*, 2019. `https://doi.org/10.1007/s11225-019-09882-1`

Batens, 1985 – Batens, D. "Dynamic dialectical logics as a tool to deal with and partly eliminate unexpected inconsistencies", in: *The Logic of Discovery and the Logic of Discourse*, J. Hintikka, F. Vandamme (eds.), Plenum Press, New York, 1985, pp. 263–271.

Batens, 1986 – Batens, D. "Dialectical dynamics within formal logics", *Logique et Analyse*, 1986, Vol. 114, pp. 161–173.

Batens, 1989 – Batens, D. "Dynamic dialectical logics", in: *Paraconsistent Logic. Essays on the Inconsistent*, G. Priest, R. Routley, J. Norman (eds.), Philosophia Verlag, München, 1989, pp. 187–217.

Batens, 2000 – Batens, D. "Minimally abnormal models in some adaptive logics", *Synthese*, 2000, Vol. 125, pp. 5–18.

Batens, 2002 – Batens, D. "On some remarkable relations between paraconsistent logics, modal logics, and ambiguity logics", in: *Paraconsistency. The Logical Way to the Inconsistent*, W.A. Carnielli, M.E. Coniglio, and I.M.L. D'Ottaviano (eds.), Marcel Dekker, New York, 2002, pp. 275–293.

Batens, 2004 – Batens, D. "The basic inductive schema, inductive truisms, and the research-guiding capacities of the logic of inductive generalization", *Logique et Analyse*, 2004, Vol. 185–188, pp. 53–84. (Appeared 2005)

Batens, 2005 – Batens, D. "On a logic of induction", in: *Confirmation, Empirical Progress, and Truth Approximation. Essays in Debate with Theo Kuipers. Vol. 1*, volume 83 of *Poznan Studies in the Philosophy of the Sciences and the Humanities*, R. Festa, A. Aliseda, J. Peijnenburg (eds.). Rodopi, Amsterdam/New York, 2005, pp. 221–242. (Contains uncorrected proofs; see [Batens, 2006] for correct version.)

Batens, 2005 – Batens, D. "The theory of the process of explanation generalized to include the inconsistent case", *Synthese*, 2005, Vol. 143, pp. 63–88. doi:10.1007/s11229-005-3114-3.

Batens, 2006 – Batens, D. "On a logic of induction", *L&PS – Logic & Philosophy of Science*, 2006, Vol. IV, No. 1, pp. 3–32. (Corrected version of [Batens, 2005].)

Batens, 2007 – Batens, D. "A universal logic approach to adaptive logics", *Logica Universalis*, 2007, Vol. 1, pp. 221–242.

Batens, 2009 – Batens, D. "Towards a dialogic interpretation of dynamic proofs", in: *Dialogues, Logics and Other Strange Things. Essays in Honour of Shahid Rahman*, C. Dégremont, L. Keiff, H. Rückert (eds.). College Publications, London, 2009, pp. 27–51.

Batens, 2011 – Batens, D. "Logics for qualitative inductive generalization", *Studia Logica*, 2011, Vol. 97, No. 1, pp. 61–80. doi:10.1007/s11225-010-9297-8.

Batens, 2014 – Batens, D. "The consistency of Peano Arithmetic. A defeasible perspective", in: *Modestly Radical or Radically Modest. Festschrift for Jean Paul van Bendegem on the Occasion of His 60th Birthday*, P. Allo, B. Van Kerkhove (eds.), Vol. 24 of *Tributes*, College Publications, London, 2014, pp. 11–59.

Batens, 2015 – Batens, D. "Tutorial on inconsistency-adaptive logics", in: *New Directions in Paraconsistent Logic*, J.-Y. Béziau, M. Chakraborty, S. Dutta (eds.), Vol. 152 of *Springer Proceedings in Mathematics & Statistics*, Springer, 2015, pp. 3–38.

Batens, 2016 – Baten, D. "Spoiled for choice?", *Journal of Logic and Computation*, 2016, Vol. 26(1) pp. 65–95. E-published 2013: doi:10.1093/logcom/ext019.

Batens, 2017 – Batens, D. "Abduction logics illustrating pitfalls of defeasible methods", in: *Applications of formal philosophy: the road less travelled*, R. Urbaniak, G. Payette (eds.), Vol. 14 of *Logic, argumentation & reasoning*, Springer, Berlin, 2017, pp. 169–193.

Batens, Clercq, 2004 – Batens, D., De Clercq, K. "A rich paraconsistent extension of full positive logic", *Logique et Analyse*, 2004, Vol. 185–188, pp. 227–257. (Appeared 2005)

Batens et al., 2009 – Batens, D., De Clercq, K., Verdée, P., Meheus, J. "Yes fellows, most human reasoning is complex", *Synthese*, 2009, Vol. 166, pp. 113–131.

Batens, Haesaert, 2003 – Batens, D., Haesaert, L. "On classical adaptive logics of induction, *Logique et Analyse*, 2001, Vol. 173–175, pp. 255–290. (Appeared 2003)

Batens, Meheus, 2000 – Batens, D., Meheus, J. "The adaptive logic of compatibility", *Studia Logica*, 2000, Vol. 66, pp. 327–348.

Batens, Meheus, 2001 – Batens, D., Meheus, J. "On the logic and pragmatics of the process of explanation", in: *Explanatory Connections. Electronic Essays Dedicated to Matti Sintonen*, M. Kiikeri, P. Ylikoski (eds.), 2001, pp. 22.

Beirlaen, Aliseda, 2014 – Beirlaen, M., Aliseda, A. "A conditional logic for abduction", *Synthese*, 2014, Vol. 191, pp. 3733–3758.

Beirlaen, Straßer, 2013a – Beirlaen, M., Straßer, C. "Non-monotonic reasoning with normative conflicts in multi-agent deontic logic", *Journal of Logic and Computation*, 2013, Vol. 24, No. 6, pp. 1179–1207.

Beirlaen, Straßer, 2013b – Beirlaen, M., Straßer, C. "Two adaptive logics of norm-propositions", *Journal of Applied Logic*, 2013, Vol. 11, No. 2, pp. 147–168.

Brady, 2006 – Brady, R. *Universal Logic*. CSLI Publications, Stanford, Cal., 2006.

Church, 1951 – Church, A. "The weak theory of implication", in: *Kontrolliertes Denken. Untersuchungen zum Logikkalkül und zur Logik der Einzelwissenschaften*, A. Menne, A. Wilhemy, H. Angsil (eds.), Vol. 16, Karl Alber, München, 1951, pp. 22–37.

De Clercq, Verhoeven, 2004 – De Clercq, K., Verhoeven, L. "Sieving out relevant and efficient questions", *Logique et Analyse*, 2004, Vol. 185–188, pp. 53–84. (Appeared 2005)

Gauderis, Van De Putte, 2012 – Gauderis, T., Van De Putte, F. "Abduction of generalizations", *Theoria*, 2012, pp. 345–363.

Goble, 2014 – Goble, L. "Deontic logic (adapted) for normative conflicts", *Logic Journal of the IGPL*, 2014, Vol. 22, No. 2, pp. 206–235.

Horsten, Welch, 2007 – Horsten, L., Welch, P. "The undecidability of propositional adaptive logic", *Synthese*, 2007, Vol. 158, pp. 41–60.

Lycke, 2012 – Lycke, H. "A formal explication of the search for explanations: the adaptive logics approach to abductive reasoning", *Logic Journal of the IGPL*, 2012, Vol. 20, pp. 497–516.

Meheus, 2001 – Meheus, J. "Adaptive logics for question", *Logique et Analyse*, 2001, Vol. 173–175, pp. 135–164. (Appeared 2003)

Meheus, 2003 – Meheus, J. "Paraconsistent compatibility", *Logique et Analyse*, 2003, Vol. 183–184, pp. 251–287. (Appeared 2005)

Meheus, 2004 – Meheus, J. "Adaptive logics and the integration of induction and deduction", in: *Induction and Deduction in the Sciences*, ed. by F. Stadler, Kluwer, Dordrecht, 2004, pp. 93–120.

Meheus, 2007 – Meheus, J. "Adaptive logics for abduction and the explication of explanation-seeking processes", in: *Abduction and the Process of Scientific Discovery*, O. Pombo, A. Gerner (eds.), Centro de Filosofia das Ciências da Universidade de Lisboa, Lisboa, 2007, pp. 97–119.

Meheus, 2011 – Meheus, J. "A formal logic for the abduction of singular hypotheses", in: *Explanation, Prediction, and Confirmation*, D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, M. Weber (eds.), Springer, Dordrecht, 2011, pp. 93–108.

Meheus, Batens, 2006 – Meheus, J., Batens, D. "A formal logic for abductive reasoning", *Logic Journal of the IGPL*, 2006, Vol. 14, pp. 221–236.

Meheus et al., 2010a – Meheus, J., Beirlaen, M., Putte, F.V.D. "Avoiding deontic explosion by contextually restricting aggregation", in: *Deontic Logic in Computer Science*, G. Governatori, G. Sartor (eds.), Springer, Dordrecht, 2010, pp. 148–165.

Meheus et al., 2010b – Meheus, J., Beirlaen, M., Putte, F.V.D., Straßer, C. "Nonadjunctive deontic logics that validate aggregation as much as possible", 2010, [https://www.clps.ugent.be/sites/default/files/publications/nadl.pdf,accessedon09.04.2020].

Meheus, Provijn, 2004 – Meheus, J., Provijn, D. "Direct dynamic proofs for classical compatibility", *Logique et Analyse*, 2004, Vol. 185–188, pp. 305–317. (Appeared 2005)

Meheus et al., 2002 – Meheus, J., Verhoeven, L., Van Dyck, M., Provijn, D. "Ampliative adaptive logics and the foundation of logic-based approaches to abduction", in: *Logical and Computational Aspects of Model-Based Reasoning*, L. Magnani, N.J. Nersessian, C. Pizzi (eds.), Kluwer, Dordrecht, 2002, pp. 39–71.

Odintsov, Speranski, 2012 – Odintsov, S.P., Speranski, S.O. "On algorithmic properties of propositional inconsistency-adaptive logics", *Logic and Logical Philosophy*, 2012, Vol. 21, pp. 209–228.

Odintsov, Speranski, 2013 – Odintsov, S.P., Speranski, S.O. "Computability issues for adaptive logics in multi-consequence standard format", *Studia Logica*, 2013, Vol. 101, No. 6, pp. 1237–1262. doi:10.1007/s11225-013-9531-2.

Pogorzelski, Prucnal, 1975 – Pogorzelski, W.A., Prucnal, T. "The substitution rule for predicate letters in the first-order predicate calculus", *Reports on Mathematical Logic*, 1975, Vol. 5, pp. 77–90.

Priest, 1987 – Priest, G. *In Contradiction. A Study of the Transconsistent.* Oxford University Press, Oxford, 2006. Second expanded edition (first edition 1987).

Routley, 1982 – Routley, R. *Relevant Logics and their Rivals*, Vol. 1, Ridgeview, Atascadero, Ca., 1982.

Shapere, 2004 – Shapere, D. "Logic and the philosophical interpretation of science", in: *Alternative Logics. Do sciences need them?*, ed. by P. Weingartner, Springer, Berlin, Heidelberg, 2004, pp. 41–54.

Straßer, 2010 – Straßer, C. "An adaptive logic framework for conditional obligations and deontic dilemmas", *Logic and Logical Philosophy*, 2010, Vol. 19, No. 1–2, pp. 95–128.

Straßer et al., 2012 – Straßer, C., Meheus, J., Beirlaen, M. "Tolerating deontic conflicts by adaptively restricting inheritance", *Logique et Analyse*, 2012, Vol. 219, pp. 477–506.

Van De Putte et al., in press – Van De Putte, F., Beirlaen, M., Meheus, J. "Adaptive deontic logics", in: *Handbook of Deontic Logic and Normative Systems*, Vol. 2, D. Gabbay, J. Horty, X. Parent, R. van der Meyden, L. van der Torre (eds.), College Publications, in press.

Van De Putte, Straßer, 2012 – Van De Putte, F., Straßer, C. "A logic for prioritized normative reasoning", *Journal of Logic and Computation*, 2012, Vol. 23, pp. 563–583. doi:10.1093/logcom/exs008.

Vanackere, 1999a – Vanackere, G. "Ambiguity-adaptive logic", *Logique et Analyse*, 1997, Vol. 159, pp. 261–280. (Appeared 1999)

Vanackere, 1999b – Vanackere, G. "Minimizing ambiguity and paraconsistency", *Logique et Analyse*, 1999, Vol. 165–166, pp. 139–160. (Appeared 2002)

Vanackere, 2000 – Vanackere, G. "**HL2**. An inconsistency-adaptive and inconsistency-resolving logic for general statements that might have exceptions", *Journal of Applied Non-Classical Logics*, 2000, Vol. 10, pp. 317–338.

Vanackere, 2001 – Vanackere, G. "The role of ambiguities in the construction of collective theories", *Logique et Analyse*, 2001, Vol. 173–175, pp. 189–214. (Appeared 2003)

Verdée, 2009 – Verdée, P. "Adaptive logics using the minimal abnormality strategy are Π^1_1-complex", *Synthese*, 2009, Vol. 167, pp. 93–104.

Verdée, Batens, 2016 – Verdée, P., Batens, D. "Nice embedding in classical logic", *Studia Logica*, 2016, Vol. 104, pp. 47–78. E-published since 2015: doi:10.1007/s11225-015-9622-3.

Weber, De Clercq, 2002 – Weber, E., De Clercq, K. "Why the logic of explanation is inconsistency-adaptive", in: *Inconsistency in Science*, ed. by J. Meheus, Kluwer, Dordrecht, 2002, pp. 165–184.

Weber, Van Dyck, 2001 – Weber, E., Van Dyck, M. "Adaptive logic and covering law explanations", *Logique et Analyse*, 2001, Vol. 173–175, pp. 237–254. (Appeared 2003)

Логические исследования
2020. Т. 26. № 1. С. 36–47
УДК 16

Logical Investigations
2020, Vol. 26, No. 1, pp. 36–47
DOI: 10.21146/2074-1472-2020-26-1-36-47

Ilkka Niiniluoto

Perception, memory, and imagination as propositional attitudes

Ilkka Niiniluoto
University of Helsinki,
P.O. Box 24 (Unioninkatu 40 A), 00014 Helsinki, Finland.
E-mail: ilkka.niiniluoto@helsinki.fi

Abstract: Jaakko Hintikka started in 1969 the study of the logic of perception as a special case of his more general approach to propositional attitudes by means of the possible worlds semantics. His students and co-workers extended this study to the logic of memory and imagination. The key elements of this approach are the distinction between physical and perspectival cross-identification and the related two kinds of quantifiers, which allow a formulation of the syntax and semantics of various types of statements about perceiving, remembering and imagining. This paper surveys the main results of these logical investigations.

Keywords: Hintikka, cross-identification, epistemic logic, imagination, memory, perception, possible worlds semantics, propositional attitudes

For citation: Niiniluoto I. "Perception, memory, and imagination as propositional attitudes", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 36–47. DOI: 10.21146/2074-1472-2020-26-1-36-47

1. Introduction

After discovering model sets in 1955 and (simultaneously with Stig Kanger) the possible worlds semantics in 1957, Jaakko Hintikka published his pioneering work *Knowledge and Belief* in 1962. This study formulated, by using the framework of model sets (as partial descriptions of possible worlds), the fundamental ideas of epistemic and doxastic logic. In *Models for Modalities* (1969) Hintikka then generalized his approach from knowledge and belief to a general theory of propositional attitudes (see also [Hintikka, 1980]). This book includes an article "On the Logic of Perception" [Hintikka, 1969], where Hintikka proposes to analyze perceptual statements (with seeing, hearing, and feeling) within modal logic in a similar way as knowing and believing. This paper used as its tool the distinction between two ways of cross-identifying individuals in alternative possible worlds. In a subsequent article "Information, Causality, and the Logic of Perception" [Hintikka, 1975a] Hintikka incorporated causal aspects to his logic of perception.

The logic of perception is an important part of Hintikka's legacy within intensional logic. It became an actively studied field in the 1970s and 1980s, with contributions (among others) by Robert Howell [Howell, 1972], Richmond Thomason [Thomason, 1973], John Bacon [Bacon, 1979], Jon Barwise [Barwise, 1981] and James Higginbotham [Higginbotham, 1983] — and from Finland Ilkka Niiniluoto [Niiniluoto, 1979; Niiniluoto, 1982] and Esa Saarinen [Saarinen, 1983]. But it is fair to say that, while epistemic and doxastic logics have become more and more popular within philosophical logic and artificial intelligence (see e.g. Hintikka [Hintikka, 2013] and the important collection edited by van Ditmarsch and Sandu [van Ditmarsch, Sandu, 2018]), the logic of perception has received relatively little attention (see, however, Rantala [Rantala, 2007] and Bourget [Bourget, 2017]). Apart from some scattered examples by Hintikka, the article by Aho and Niiniluoto [Aho, Niiniluoto, 1990] has remained the only systematic investigation of the logic of memory. On the other hand, the logic of imagination, introduced by Niiniluoto [Niiniluoto, 1983; Niiniluoto, 1985a; Niiniluoto, 1985b] along Hintikka's lines (see also [Aho, 1994]), has experienced a recent renaissance with several new contributions (see [Costa-Leite, 2010; Wansing, 2017; Berto, 2017]).

2. Hintikka on Propositional Attitudes

Let a be a person or agent (a proper name in language) and p a proposition (a factual statement in language). Then examples of *propositional attitudes*, which are relations between a and p, include

$$\begin{aligned}
K_a p &= a \text{ knows that } p\\
B_a p &= a \text{ believes that } p\\
S_a p &= a \text{ sees that } p\\
R_a p &= a \text{ remembers that } p\\
I_a p &= a \text{ imagines that } p.
\end{aligned}$$

According to Hintikka, a general truth condition for an attitude $\varnothing$ can be formulated as follows:

> Sentence 'a $\varnothing$s that p' is true in world w if and only if p is true in all possible worlds which are compatible with what a $\varnothing$s in world w.

Similarly,

> Sentence 'a $\varnothing$s that p' is false in world w if and only p is false in some possible world which is compatible with that a $\varnothing$s in world w.

Here the condition

w' is compatible with what a $\varnothing$s in w

defines an alternativeness relation for $\varnothing$ in the sense of possible worlds semantics. Thus, 'a $\varnothing$s that p' is true in w if and only if p is true in all $\varnothing$-alternatives of w.

Immediate consequences of the truth condition for any attitude $\varnothing$ include

($\varnothing$1) $\varnothing_a(A \rightarrow B) \rightarrow (\varnothing_a A \rightarrow \varnothing_a B)$

($\varnothing$2) $\varnothing_a(A\&B) \equiv (\varnothing_a A \,\&\, \varnothing_a B)$

($\varnothing$3) $\varnothing_a T$, if T is a tautology

($\varnothing$4) $\varnothing_a A \rightarrow \varnothing_a(A \vee B)$.

When $\varnothing$ is replaced by K, B, S, R, or I, we obtain basic principles for these specific propositional attitudes. Besides these principles[Hintikka, 1962] argued that knowledge K (unlike belief B) satisfies the success condition

(K5) $K_a A \rightarrow A$

and the KK-principle

(K6) $K_a K_a A \equiv K_a A$.

Hintikka's truth definition for propositional attitudes leads to a problem which is called *logical omniscience* in epistemic logic: an agent knows all tautologies and all logical consequences of her knowledge. This is unrealistic, if knowledge is understood as an actual mental state of a person. Similar problems arise for "logical omniperception" (in watching an ice hockey match, do I see that Lionel Messi is playing or Lionel Messi is not playing?) or "logical omnimemory" (do I remember all logical and mathematical truths as Plato's slave boy in *Meno*?). One solution is to accept that we in fact know and see tautologies: when $S_a p$ means that according to the perceptions of a it is the case that p, then trivially a tautology T is true in the actual world and all of its S-alternatives. But there are also many other more technical solutions to logical omniscience. Hintikka himself proposed in 1975 the use of "impossible worlds", which were developed as "urn models" by Veikko Rantala [Rantala, 1982]. If one allows non-normal possible worlds, where ordinary laws of logic are not satisfied, then propositional attitudes do not satisfy closure conditions for logical consequence. This proposal has been recently applied in the logic of imagination as "hyperintensionality" (see [Berto, 2017]). Hintikka also argued that one can use "small worlds", which need not include all possible individuals (like Lionel Messi), and the same restriction can be obtained by Barwise's "situations" [Barwise, 1981]. Fagin and Halpern proposed an "awareness logic" [Fagin, Halpern, 1985], where explicit knowledge concerns only such propositions about which the agent is aware, but this is a very strong restriction, since actual awareness need not satisfy even the closure condition for conjunctions (cf. (∅2)).

3. The Logic of Perception

Hintikka's proposal to treat *perception* as a propositional attitude ([Hintikka, 1969]) was inspired by Elizabeth Anscombe's thesis about the intensionality of perceptual ascriptions. It is also related to Edmund Husserl's phenomenological approach to intentionality as directness. At the same time this choice reflects Hintikka's "neo-Kantian" conviction that perception is thoroughly conceptual, always mediated by conceptual schemes. He even blames Husserl for assuming that in our sensuous experience there exists a non-conceptual ingredient or *hyle*, which is changed into an experience about an object by the act of *noesis* (see [Hintikka, 1975b, p. 198]). By the same argument, Hintikka would reject the idea of non-conceptual content in experience (see e.g. [Crane, 1992]). Perception differs from imagination by the fact that it involves causal interaction with external objects. With reference to the psychologist James Gibson's view of senses as information systems, Hintikka characterizes perception as a method of reaching information about the world.

The logic of perception can be understood as an attempt to develop an explicit semantics for sentences containing perceptual terms [Niiniluoto, 1982]. But the truth conditions of perceptual sentences provide also a formal syntax which exhibits the systematic interconnections between different grammatical constructions with perceptual terms. Just like epistemic logic shows how expressions like 'know who', 'know where', 'know when' etc. can be reduced to propositional 'know that'(see [Hintikka, 1962]), the logic of perception shows that 'seeing that' is the basic form of perceptual statements. In particular, the propositionality of perception is reflected in the result that all direct object *de re* constructions (about things or events) are reduced to sentences with seeing that. And, by the intensionality of perception, the truth conditions for statements of the form $S_a p$ have to refer to several alternative possible worlds of states of affairs at the same time.

Perception is usually understood as a species of knowledge, even though errors of observation are common (illusions, hallucinations). Evolutionary arguments suggest that human perception is relatively reliable in ordinary circumstances. Some early attempts to develop logics of perception imitated epistemic logic. For example, Richmond Thomason [Thomason, 1973] assumed that seeing satisfies the success condition

$(S5)\ \ S_a A \to A$

and John Bacon [Bacon, 1979] suggested an SS-principle

$(S6)\ \ S_a S_a A \equiv S_a A$

But Hintikka realized that it is better to start from a weaker interpretation, where $S_a p$ means something like 'it appears to a that p', 'it looks to a that p' or 'a seems to see that p'. In this sense, the S-operator does not satisfy the success condition $S5$, so that it belongs to the same group of propositional attitudes as belief. A stronger notion of veridical seeing $^*S$, which satisfies the success principle $^*S_a p \longrightarrow p$, can be obtained from the weaker S by adding conditions which are sufficient to guarantee the truth of the perceived p. It is also interesting to investigate the interplay of the operators K and S [Hintikka, 1975a; Niiniluoto, 1979].

Similar remarks apply the notion of *memory* [Aho, Niiniluoto, 1990]. As a propositional attitude, memory is more complex than perception, since 'a remembers that p' allows for many temporal alternatives, where p may be an eternal, past tense, present tense, or future tense sentence. For example, 'I remember that $5 + 6 = 11$', 'I remember that Jaakko was lecturing on information in 1967', 'I remember that today is my daughter's birthday', and 'I remember that tomorrow is my wife's birthday'. Again memory is relatively reliable, but mistakes are common. So in the logic of memory one should start from a weak interpretation of R, which does not satisfy the success principle

$(R5)\ \ R_a A \to A,$

but a strong notion of remembering $^*R$ can be obtained by adding conditions so that $^*R_a A \to A$ is satisfied. At least for the strong notion we have the principle that $^*R_a p$ at t implies $(Et' < t)S_a p$ at t', i.e. reliable memories are based on earlier perceptions. Instead of the RR-thesis $(R6)$ it is plausible to assume that $K_a R_a A \equiv R_a A$.

For *imagination*, which a mental faculty of creating fictional worlds, it is even more straightforward to observe that the principle $I_a A \to A$ is not valid [Niiniluoto, 1983; Niiniluoto, 1985a]. Still, it would be too strong to assume an anti-success principle $I_a A \to \neg A$, since our imagination may be accidentally true. It can be debated whether it is possible to imagine physically impossible or logically contradictory states of affairs (see [Niiniluoto, 1985b; Costa-Leite, 2010; Berto, 2017]). Berto, whose dialetheism accepts the conceivability of real contradictions, gives an affirmative answer to this question. In order to emphasize imagination as an activity, Wansing analyses imagination by combining a neighbourhood semantics with a modal logic of agency [Wansing, 2017].

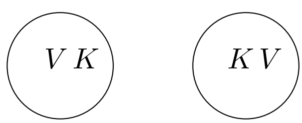
4. Quantifiers and Propositional Attitudes

The expressive force of Hintikka's treatment of propositional attitudes is seen only when we move from propositional logic to a framework with existential and universal quantifiers. This requires a solution to the problem of quantifying into an intensional context, i.e. a method of identifying the same individual in different possible worlds. In Hintikka's approach, identified individuals constitute *world lines*, which as intensional entities serve as interpretations of quantified variables (cf. [Tulenheimo, 2017]). The *cross-identification* of individuals can be achieved by two different method: *physical* (descriptive) world lines rely on physical properties of individuals, such as their permanent public attributes and spatio-temporal continuity, while *perspectival* world lines depend on the role of individuals in the agent's perspective. In the case of perception, the perspectival method identifies those individuals who play the same role in the visual field of the percipient (cf. [Rantala, 2007]). These two methods of cross-identification are correlated with two different quantifiers: the physical existence quantifier is denoted by (Ex) and the perspectival by $(\exists x)$. Then the truth conditions for quantified sentences with the S-operator can be formulated as follows:

1. $(Ex)S_aA(x)$ is true at world w if and only if there is a physical world line f which picks out an individual in each S-alternative w' of w such that $f(w')$ satisfies $A(x)$ at w';

2. $(\exists x)S_aA(x)$ is true at world w if and only if there is a perspectival world line f which picks out an individual in each S-alternative w' of w such that $f(w')$ satisfies $A(x)$ at w'.

For example, assume that I meet on the road two familiar brothers, Ville and Kalle, but I am not able to recognize who is who of them. The worlds compatible with by perception are two:



Then the perspectival world line picks out the brother on the left side, i.e. V and K, while the physical word line identifies Kalle (resp. Ville) in the two alternative worlds.

According to the causal theory of perception, sense experience is normally caused by external objects and events in the real world. Hintikka (see [Hintikka, 1975a]) complemented his logic of perception by requiring that perspectival world lines are extended to the actual world by means of a *causal* connection. For example, the line connecting the brother on the left is continued to the individual who in the actual world has caused the observation. Memory involves typically two causal processes: first learning that p by perception and then maintaining this memory content in the mind over time. Due to their temporal dimension, the world lines for memory are more complex, since they may pick out temporally extended individuals from possible world histories (see [Aho, Niiniluoto, 1990]).

With this machinery, we can formalize a variety of different epistemic and perceptual statements (see [Niiniluoto, 1982]). Examples of sentences with a direct reference to the object of perception include the following:

$(\exists x)K_a(x = b)$	a knows b
$(\exists x)S_a(x = b)$	a sees b
$(\exists x)(x = b \,\&\, S_a(\exists x)(y = x))$	a looks at b

The sentence 'a sees b' is intensional in the sense that the object b may be misidentified or a mere illusion. But instead 'a looks at b' implies that $(\exists x)(x = b)$, i.e. b exists. The construction of seeing as, which was important to Ludwig Wittgenstein, has a natural formalization (see [Howell, 1972]):

$(\exists x)(x = b \,\&\, S_a(x = c))$	a sees b as c
$(\exists x)(x = b \,\&\, S_aFx)$	a sees b as an F

Additional examples with a physical quantifier include

$(Ex)K_a(x = b)$	knows who b is
$(Ex)S_a(x = b)$	a sees who b is

Besides perceiving things and states of affair, one may speak about perceiving events, when we allow quantifiers to range over events (or world-line connecting events in alternative possible worlds). For example, we may distinguish between

$S_a(\text{Esa runs})$	a sees that Esa runs
$(\exists e)(e = \text{Esa's running} \,\&\, S_a(\exists x)(x = e))$	a sees Esa run

(see [Niiniluoto, 1982]). The former sentence is intensional, so that I can be mistaken by the observed person or his activity. The latter sentence 'I see Esa run' is known in English as "naked infinitive". Jon Barwise (see [Barwise, 1981]) proposed in his situation semantics that the sentence 'I see Esa run' is true, if there is a situation which I see and which supports the truth of the sentence 'Esa runs'. Seeing a situation is a purely extensional relation for Barwise. Thus, such extensional perceptual statements are associated with a success condition: if I see Esa run, then Esa runs. This holds also of the Hintikka style formalization, which implies that $(\exists e)(e = \text{Esa's run})$. In the same way, the statement 'I see the birch tree blowing in the wind' can be formalized by the formula

$$(\exists e)(e = \text{the tree is blowing in the wind } \& S_a(\exists x)(x = e)).$$

Here Barwise's extensional success condition is satisfied, but the problem of his situation semantics is its inability to treat the intensionality of perception (cf. [Saarinen, 1983; Higginbotham, 1983; Niiniluoto, 1985a; Niiniluoto, 1985b]).

In Hintikka's formalism, one may distinguish the epistemologically important cases ([Niiniluoto, 1979]):

$(\exists x)(x = b \,\&\, Fx \,\&\, S_a Fx)$	veridical perception
$(\exists x)(\sim Fx \,\&\, S_a Fx)$	visual illusion
$(\exists x)(S_a Fx) \,\&\, \sim (\exists x)((Ey)(y = x) \,\&\, S_a Fx)$	visual hallucination

As an example of hallucination, in the morning after a heavy party I may see a pink elephant on the wall (F), but the associated perspectival world line cannot be extended to the actual world. The sentence

$$(\exists x)(S_a Fx \,\&\, K_a \sim Fx)$$

expresses a conscious illusion: it seems to me that the oar is bent in the water, even though I know that this is not really the case. Hence, illusions need not always be mistaken beliefs, as many theories of perception claim.

By combining perceptual and epistemic operators further interesting cases are obtained (see [Niiniluoto, 1979]):

$(\exists x)(x = b \,\&\, S_a(\exists y)(y = x) \,\&\, B_a(x = c))$	a visually holds b as c
$(\exists x)(x = b \,\&\, S_a(\exists y)(y = x) \,\&\, K_a(\exists y)(y = x))$	a notices b
$(\exists x)(S_a(\exists y)(y = x) \,\&\, K_a(x = b))$	a recognizes b

For similar reasons perception may fail in many ways: don't look at, don't see, don't notice, don't recognize.

Corresponding formulations for memory (e.g. 'I remember you', 'I remember Jaakko lecturing', 'I am reminiscing about her', 'I remember who this girl is') and imagination (e.g. 'I am imagining about my friend', 'I imagine her as Anna Karenina', 'I imagine that Esa is running') can be given by the two kinds of quantifiers combined with the operators R_a and I_a (see [Aho, Niiniluoto, 1990; Niiniluoto, 1985b]). It is also easy to formulate sentences for remembering when, where, what, and who. But a complete formalization of memory statements should be combined with temporal logic: the statement 'I remember Esa as a young student' is directed to a person living now, but 'Jaakko remembers Gödel' should not entail that Gödel exists now.

An interesting special feature of memory and imagination is self-identification. Memories of past event are personal in the sense that the agent has to be able to place himself or herself in the remembered scene. If I remember that Jaakko was lecturing in 1967, I have to identify myself as a person in the audience. David Lewis (see [Lewis, 1979]) has called such epistemic abilities *de se* attitudes. More generally, contexts involving *de se* attitudes may involve interplay of physical and perspectival identification.

5. Concluding Remarks

The logic of perception is mainly interesting for epistemology and philosophy of language, but it may have potential applications with the psychology of perception and cognitive neuroscience. Hintikka himself was excited by the fact that his philosophical distinction between the physical and perspectival methods of cross-identification has a counterpart within neuroscience: the what- and where-systems of visual perception [Vaina, 1990] and the semantic and episodic memory [Tulving, 1972] (see [Hintikka, 1990; Hintikka, Symons, 2003]). But while the neuroscientists have postulated two different kinds of visual perception or memory, Hintikka's system is more economical, as it assumes only one perceptual operator (seeing that) or memory operator (remembering that).

Given the strong emphasis on the concept-laden nature of perception and memory, one may ask whether the Hintikka-type of approach is applicable to animals and children before they have learnt a symbolic language. One possibility is that the logic of perception is a third-person analysis of perceptual processes independently whether the agent has linguistic abilities. But Hintikka's own discussion seems to assume that the framework describes perceptual experiences of actual subjects. Then one might surmise that the physical cross-identification is not yet successful for a creature on the pre-linguistic level, as this presupposes mastery of temporal and spatial concepts and the objective distinction between "you" and "me". Perspectival cross-identification is simpler, as it allows a dog to "know" its master or a child to "know" her mother. The formula 'a looks at b' presupposes only that a is able to see b as an existing object separate from its environment, which is possible already in the pre-conceptual level of consciousness. But here it is somewhat perplexing that Tulving argues that animals possess the semantic memory but lack the episodic memory (see [Tulving, 1972]). Perhaps such animal abilities should be formalized by statements involving remember-how in analogy with know-how.

Similar question arise, if the logic of perception and the logic of memory are applied to theories and practices of artificial intelligence, such as pattern recognition and machine learning. There the human agent is replaced by a robot or a self-regulating computer program, which does not have intentional mentality or *de se* attitudes. Still, such machines can be taught to be in causal interaction with their environment, to store perceptual data and to use them in recognition and inference.

References

Aho, 1994 – Aho, T. *On the Philosophy of Attitude Logic*, Acta Philosophica Fennica 57, Societas Philosophica Fennica, Helsinki, 1994.

Aho, Niiniluoto, 1990 – Aho, T., Niiniluoto, I. "On the Logic of Memory", in: *Language, Knowledge, and Intentionality: Perspectives on the Philosophy of Jaakko Hintikka*, L. Haaparanta, M. Kusch, and I. Niiniluoto (eds.), Acta Philosophica Fennica, Societas Philosophica Fennica, Helsinki, 1990, Vol. 49, pp. 408–429.

Bacon, 1979 – Bacon, J. "The Logical Form of Perception Sentences", *Synthese*, 1979, Vol. 41, pp. 271–308.

Barwise, 1981 – Barwise, J. "Scenes and Other Situations", *The Journal of Philosophy*, 1981, Vol. 78, pp. 369–397.

Berto, 2017 – Berto, F. "Impossible Worlds and the Logic of Imagination", *Erkenntnis*, 2017, Vol. 82, pp. 1277–1297.

Bourget, 2017 – Bourget, D. "Intensional Perceptual Ascriptions", *Erkenntnis*, 2017, Vol. 82, pp. 513–530.

Costa-Leite, 2010 – Costa-Leite, A. "Logical Properties of Imagination", *Abstracta*, 2010, Vol. 6, pp. 103–116.

Crane, 1992 – Crane, T. "The Nonconceptual Content of Experience", in: *The Contents of Experience: Essays on Perception*, ed by T. Crane, Cambridge University Press, 1992, pp. 136–157.

Fagin, Halpern, 1985 – Fagin, R., Halpern, J.Y. "Belief, Awareness and Limited Reasoning", *Artificial Intelligence*, 1985, Vol. 34, pp. 39–76.

Higginbotham, 1983 – Higginbotham, J. "The Logic of Perceptual Reports: An Extensional Alternative to Situation Semantics", *The Journal of Philosophy*, 1983, Vol. 80, pp. 100–127.

Hintikka, 1962 – Hintikka, J. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press, Ithaca, NY, 1962.

Hintikka, 1969 – Hintikka, J. "On the Logic of Perception", in: *Models for Modalities*, ed by D. Reidel, Dordrecht, 1969, pp. 151–183.

Hintikka, 1975a – Hintikka, J. "Information, Causality, and the Logic of Perception", *The Intentions of Intentionality*, D. Reidel, Dordrecht, 1975, pp. 59–75.

Hintikka, 1975b – Hintikka, J. "The Intentions of Intentionality", *The Intentions of Intentionality*, D. Reidel, Dordrecht, 1975, pp. 192–222.

Hintikka, 1980 – Hintikka, J. *Logical-Epistemological Studies*, ed. by V.N. Sadovski and V.A. Smirnov, Progress, Moscow, 1980. (In Russian)

Hintikka, 1990 – Hintikka, J. "The Cartesian Cogito, Epistemic Logic and Neuroscience: Some Surprising Connections", *Synthese*, 1990, Vol. 83, pp. 133–157.

Hintikka, 2013 – *Open Problems in Epistemology – Problèmes Ouverts en Épistemologie*, ed by J. Hintikka, Acta Philosophica Fennica, Societas Philosophica Fennica, Vol. 90, Helsinki, 2013.

Hintikka, Symons, 2003 – Hintikka, J., Symons, J. "Systems of Visual Identification in Neuroscience: Lessons from Epistemic Logic", *Philosophy of Science*, 2003, Vol. 70, pp. 89–104.

Howell, 1972 – Howell, R. "Seeing As", *Synthese*, 1972, Vol. 23, pp. 400–422.

Lewis, 1979 – Lewis, D. "Attitudes de dicto and de se", *The Philosophical Review*, 1979, Vol. 88, pp. 513–543.

Niiniluoto, 1979 – Niiniluoto, I. "Knowing that One Sees", in: *Essays in Honour of Jaakko Hintikka*, E. Saarinen, R. Hilpinen, I. Niiniluoto, and M.B. Provence Hintikka (eds.), D. Reidel, Dordrecht, 1979, pp. 249–282.

Niiniluoto, 1982 – Niiniluoto, I. "Remarks on the Logic of Perception", in: *Intensional Logic: Theory and Applications*, I. Niiniluoto and E. Saarinen (eds.), Acta Philosophica Fennica 35, Societas Philosophica Fennica, Helsinki, 1982, pp. 116–129.

Niiniluoto, 1983 – Niiniluoto, I. "On the Logic of Imagination", in: *Vexing Questions*, I. Patoluoto et al. (eds.), Reports from the Department of Philosophy 3, University of Helsinki, 1983, pp. 23–28.

Niiniluoto, 1985a – Niiniluoto, I. "Remarks on the Logic Imagination", in: *Action, Logic, and Social Theory*, G. Holmström and A.J.I. Jones (eds.), Acta Philosophica Fennica 38, Societas Philosophica Fennica, Helsinki, 1985, pp. 183–202.

Niiniluoto, 1985b – Niiniluoto, I. "Imagination and Fiction", *Journal of Semantics*, 1985, Vol. 4, pp. 22–29.

Rantala, 1982 – Rantala, V. "Quantified Modal Logic: Non-Normal Worlds and Propositional Attitudes", *Studia Logica*, 1982, Vol. 41, pp. 41–65.

Rantala, 2007 – Rantala, V. "Perception in Context and Its Logic", in: *Approaching Truth: Essays in Honour of Ilkka Niiniluoto*, S. Pihlström, P. Raatikainen, and M. Sintonen (eds.), College Publications, London, 2007, pp. 113–127.

Saarinen, 1983 – Saarinen, E. "On the Logic of Perception Sentences", *Synthese*, 1983, Vol. 54, pp. 115–128.

Thomason, 1973 – Thomason, R. "Perception and Individuation", in: *Logic and Ontology*, ed by M. K. Munitz, New York University Press, New York, 1973, pp. 261–285.

Tulenheimo, 2017 – Tulenheimo, T. *Objects and Modalities: A Study in the Semantics of Modal Logic*, Springer, Cham, 2017.

Tulving, 1972 – Tulving, E. “Episodic and Semantic Memory”, in: *Organization of Memory*, E. Tulving, W. Donaldson (eds.), Academic Press, New York, 1972, pp. 381–402.

Vaina, 1990 – Vaina, L. “ ‘What’ and ‘Where’ in Human Visual System: Two Hierarchies of Visual Modules”, *Synthese*, 1990, Vol. 83, pp. 49–52.

van Ditmarsch, Sandu, 2018 – *Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, van Ditmarsch H., Sandu G. (eds.), Outstanding Contributions to Logic 12, Springer, Cham, 2018.

Wansing, 2017 – Wansing, H. “Remarks on the Logic of Imagination: A Step Toward Understanding Doxastic Control Through Imagination”, *Synthese*, 2017, Vol. 194, pp. 2843–2861.

Логические исследования
2020. Т. 26. № 1. С. 48–59
УДК 510.67+510.22

Logical Investigations
2020, Vol. 26, No. 1, pp. 48–59
DOI: 10.21146/2074-1472-2020-26-1-48-59

Graham Priest

Metatheory and dialetheism

Graham Priest
Departments of Philosophy, the CUNY Graduate Center,
365 5th Ave, New York, NY 10016, USA.
and the University of Melbourne,
Grattan Str., Parkville, Victoria, 3010, Australia.
E-mail: priest.graham@gmail.com

Abstract: Given a formal language, a metalanguage is a language which can express — amongst other things — statements about it and its properties. And a metatheory is a theory couched in that language concerning how some of those notions behave. Two such notions that have been of particular interest to modern logicians — for obvious reasons — are truth and validity. These notions are, however, notoriously deeply entangled in paradox. A standard move is to take the metalanguage to be distinct from the language in question, and so avoid the paradoxes. One of the attractions of a dialetheic approach to the paradoxes of self-reference is that this move may be avoided. One may have a language with the expressive power to talk about — among other things — itself, and a theory in that language about how notions such as truth and validity for that language behave. The contradictions delivered by these notions are forthcoming, but they are quarantined by the use of a paraconsistent logic. The point of this paper is to discuss this project, the extent to which it has been successful, and the places where issues still remain.

Keywords: truth, validity, dialetheism, paraconsistency, paradox, self-reference, model theory, material detachment, set theory

For citation: Priest G. "Metatheory and dialetheism", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 48–59. DOI: 10.21146/2074-1472-2020-26-1-48-59

Dedication: Sasha Karpenko was an outstanding logician, and a key player in logic in Russia. I met him on a couple of trips to Moscow. He was an engaging philosophical interlocutor, and a kind and warm host. I dedicate this paper fondly to his memory.

1. Introduction

Logic is about arguments. Arguments are expressed in languages; and for modern logicians, these are formal languages. For such a language, a metalanguage is a language which can express — amongst other things — statements about that language and its properties. And a metatheory is a theory couched in that language concerning how some of those notions behave. Two such notions that have been of particular interest to modern logicians — for obvious reasons — are truth and validity. These notions are notoriously, however, deeply entangled in paradox. A standard move, since Tarski's *Warheitsbegri* [Tarski, 1935] is to take the metalanguage to be distinct from language in question, and so avoid the paradoxes.

I think that most logicians would now agree with Tarski that this move is just an artifice — one with little justification other than to avoid contradiction. Natural languages, in which, of course, such paradoxes find their home, most certainly do not seem to be structured in this hierarchical fashion.

One of the attractions of a dialetheic approach to the paradoxes of self-reference is that this move may be avoided. One may have a language with the expressive power to talk about — among other things — itself, and a theory in that language about how notions such as truth and validity for that language behave. The contradictions delivered by these notions are forthcoming, but they are quarantined by the use of a paraconsistent logic.

The point of this paper is to discuss this project, the extent to which it has been successful, and the places where issues still remain. In the first part of the paper I will discuss truth; in the second and much longer part, I will discuss validity.

2. Truth

As far as truth is concerned, what we need is a language which contains a way of referring to its own sentences, and a truth predicate that applies to these. As is now standard, a simple way of talking about sentences is to suppose that our theory contains arithmetic, and use a gödel coding. I will assume this in what follows. In particular, given any sentence, A, of the language, I will write $\langle A \rangle$ for the numeral of the gödel code of A. This is its name.

The theory must also tell us how truth behaves. Given that avoiding paradox is no longer necessary, the natural and obvious thought is that it should deliver all instances of the T-schema:

- $T\langle A \rangle$ iff A

There is a question about how to understand the 'iff' here, and there are various possibilities. One is as the biconditional of some relevant logic, $\leftrightarrow$; another is as bi-deducibility, $\dashv\vdash$; another is as a material biconditional, $\equiv$ (where $A \equiv B$ is $(\neg A \vee B) \wedge (\neg B \vee A)$). Given some simple assumptions, the first of these options is the strongest. In particular, any results about what cannot be proved using this notion of conditionality carry over to the weaker notions. So let me discuss this option.

Take the logic of our theory to be an appropriate relevant logic;[1] and take a theory in this logic which contains enough arithmetic, plus the axiom schema:

- $T\langle A\rangle \leftrightarrow A$

for any closed sentence, A. It is now well known that such a theory is inconsistent but non-trivial. Thus, one can show that some sentences are both true and not true — for example, the sentence L of the form $\neg T\langle L\rangle$. However, one cannot prove everything. In particular, any sentence in the $\rightarrow$-free fragment in the language which is grounded in Kripke's sense [Kripke, 1975] behaves consistently. The proof of this and references may be found in [Priest, 2008, § 8].

Of course, one may wish for more from a theory of truth than this. In particular, Tarski showed how to give a theory of truth in which truth conditions are given recursively, and the T-Scheme is then proved. One can do this too. What is essentially the Tarski construction can also be carried out in a paraconsistent logic. The details of the construction can be found in [Priest, 1987, ch. 9].

Before we move on to validity, let me make a couple of comments on two notions cognate with truth: satisfaction and denotation. The construction which shows the non-triviality of the T-Schema may be used to establish the non-triviality of the (one-place) Satisfaction-Schema:[2]

- $yS\langle A\rangle \leftrightarrow A_x(y)$

where S is the satisfaction predicate, A is any formula of one free variable, x, and $A_x(y)$ is the result of substituting y for x in A (relabelling bound variables if necessary to avoid a clash).

We may define the denotation predicate, D, in the obvious way: $\langle t\rangle Dy := yS\langle x = t\rangle$, where t is any (closed) term, to deliver the Denotation Schema:

- $\langle t\rangle Dy \leftrightarrow y = t$

The same results then follow for denotation.

There is an extra complication in this case, however. I have tacitly assumed that the language we are dealing with till now does not contain descriptive terms. If we add such terms to the language, complications arise in the case of denotation. Non-triviality results are obtainable, at least when the

[1]Exactly what this is, we do not need to go into here. But I assume that the logic contains the Principle of Excluded Middle, $\models A \vee \neg A$, but not Absorption, $A \rightarrow (A \rightarrow B) \models A \rightarrow B$. A suitable such logic is BX. See [Priest, 2008, 10.4a.12].

[2]And its generalisation to an arbitrary number of free variables.

biconditional of the Scheme is bi-deducibility.[3] However, matters are less than straightforward, since they are entangled with assumptions one makes about denotation.

3. Validity: Preliminary Considerations

Let us now move from truth to the somewhat move vexed notion of validity; and let us start by getting some relatively straightforward matters out of the way.

In modern logic, validity can be defined syntactically, in terms of some proof system, or semantically, in terms of interpretations of the language. Any axiomatic theory that contains arithmetic can define its own syntactic validity relation, at least for finite premise-sets. Thus, given an axiom system for the theory, whose axioms are the members of some decidable set X, then if Y is finite set of sentences, $Y \vdash A$ iff there is a finite sequence of formulas, $A_1, ..., A_n$ such that A is A_n, and for any $i < n$, A_i is either in X, or in Y, or follows from some sentences earlier in the sequence by some rule of inference. All this can be expressed in arithmetic in a familiar fashion. Nothing about paraconsistency changes this matter.

Of course, if the arithmetic theory is axiomatic and consistent, it cannot prove its own consistency. On the other hand, there are complete and inconsistent arithmetics based on a paraconsistent logic which can prove their own non-triviality.[4]

So let us switch our attention to a semantic definition of validity. According to such a definition, $X \models A$ iff every interpretation (appropriate for the logic in question) which makes all the members of X true makes A true. First, note that the notion of truth here is truth-in-an-interpretation, not truth *simpliciter*. Of course, one might hope that there is some interpretation such that truth in that interpretation is extensionally equivalent to truth *simpliciter* (a *standard* interpretation); but such is not required for a definition of validity. Next, note the the notion of an interpretation, as it is standardly understood, is a set-theoretic one. (An interpretation comprises a domain of quantification, a denotation function, etc.) Hence, to give such a definition requires one to deploy the language of set theory.

How to do so is familiar to anyone with a knowledge of the elements of model-theory. Thus, suppose our language is that of first-order set-theory, and our theory is ZFC. Then, given any language, L, and a notion of interpretation for that language, we can define the relation $X \models A$ in a straightforward fashion. In particular, L can be the language of first-order set theory, and the

[3] See [Priest, 1998; Priest, 1999] and [Priest, 2005, ch. 8].

[4] See [Priest, 1987, 17.4].

notion of interpretation can be that of classical logic. The notion of validity defined can then be the one deployed in ZFC. In that sense, ZFC can define its own validity relation.

What it cannot do, at least if ZFC is consistent, is to prove that there is a standard model. That is, it cannot establish the existence of an interpretation, I, such for any sentence, A, in the language of set theory, $\langle A\rangle$ is true in I iff A. This fact deprives us of a rationale as to why one may legitimately deploy this notion of validity when reasoning in ZFC itself — where we are, presumably, interested in deploying a notion of validity that preserves truth *simpliciter*. It is not at all obvious how to address this issue. Probably the best known approach is to apply the notion of informal rigour, as suggested by [Kreisel, 1967].

4. Paraconsisent Validity

Having got these matters straight, let us now turn to the issue of a model-theoretic definition of validity appropriate for a dialetheic solution to the paradoxes of self-reference using a paraconsistent logic.

Of course, if one holds that ZFC is the correct set theory, matters are exactly the same as in the case of classical logic. One simply replaces the notion of a classical interpretation with that of the notion of interpretation appropriate for the paraconsistent logic at hand. This approach is hardly available to someone who endorses a dialetheic solution to the paradoxes of set-theory, however. For in such an approach one endorses the naive comprehension schema:

- $\exists x\forall y(y \in x \text{ iff } A)$

where A is arbitrary.[5] The set-theoretic paradoxes are then forthcoming, but they are quarantined by the use of the paraconsistent logic. Naturally, this commits one to a set-theory quite different from ZFC.

And here we meet our first real problem. What is that set theory? The matter turns again on how one is to understand the conditional 'iff' in the schema. There are presently two approaches to the problem. The first is to take the underlying logic of the theory to be an appropriate relevant logic, and take the biconditional to be that of this logic. This approach has been developed at greatest length by Weber [Weber, 2010; Weber, 2012], who has shown how to prove most of the standard results concerning ordinal and cardinal arithmetic (and many other interesting things) in this theory. The theory is also known to be non-trivial, due to a proof of Brady.[6] It is also clear that the

[5]Even if one insists that x does not occur free in A, the more general case follows. A proof of this for set theory based on a substructural logic can be found in [Cantini, 2003, Theorem 3.20]. The same proof works in relevant logic.

[6]For discussion and references, see, again, [Priest, 2002, § 8].

standard model-theoretic definition of validity can be given for the logic used. One simply defines it and the notions it requires in the obvious way. One has at one's disposal, after all, the naive comprehension schema. Unfortunately, it is not known how much of standard model theory can be established in this way, since the usual proofs deploy inferences not available in the theory. Naturally, the definitions themselves are not much good unless we can show that the notions defined have at least some minimal properties, such as that the proof system of the logic is sound with respect to the notion of validity. Until more is known about these matters, it is impossible to say anything much for the issue at hand.[7]

A quite different approach to paraconsistent set theory is to take its underlying logic to be that of the paraconsistent logic LP, and to take the biconditional in the comprehension scheme to be the material biconditional of that logic.[8] If one takes this approach then, since the conditional of this logic does not detach, one cannot *prove* anything much from the set theoretic axioms. A different approach is required. This is itself model-theoretic. It can be shown that there are interpretations of the language of set-theory which are models of both the naive comprehension schema *and* all the theorems of ZFC. The set of sentences true in such models is inconsistent, since one can prove paradoxes such as Russell's, but non-trivial. And now, if one may assume that the universe of sets — or universes of sets if there is more than one — is/are given by such (an) interpretation(s), then one may simply help oneself to anything that can be engineered in ZFC, including the definition of validity for the logic LP, and all those results about it that may be proved.[9]

One thing that this approach can do, which can not be done classically, to deliver us a standard model. For it can be shown that there are interpretations of the kind just indicated in which all instances of the following are true:

- $\exists x(x \text{ is an } LP \text{ interpretation} \wedge (A \equiv (x \Vdash^{+} \langle A \rangle)))$

[7]In [Weber, 2016], Weber shows one way in which semantic validity may be defined for propositional logic, and proves soundness and completeness. His approach comes with a steep downside, however: every inference is invalid (though some are valid too). Non-triviality proofs also become trivially easy, and so somewhat vacuous. It seems to me that many of the problems arise because Weber endorses the exclusivity of truth and falsity in an interpretation (p. 539). This assumption, it seems to me, could be jettisoned. However, this is not the place to go into that matter.

[8]This account of set-theory is proposed in the second edition of [Priest, 1987, ch. 18], and explored at greater length in [Priest, 2017, §§ 10–12].

[9]Naturally, one may ask why such an assumption is justified. Perhaps there is no better answer than that it seems to validate the things we take to be true of sets — though of course it does not justify these beliefs.

(Here, $\Vdash^+$ denotes truth in a *LP* interpretation.) Assuming, again, that the universe(s) of sets is (are) like this then justifies the application of this notion of validity in reasoning about sets. Of course, one might reject the claim that the "intended" models of set theory do contain a standard model. In that case, we would be no worse off than in the classical case, and we would have to deploy some strategy such as Kreisel's in an attempt to justify using *LP* to reason about sets.

A second positive fact about this notion of validity, is that it solves a version of the validity-Curry paradox, first proposed by [Beall, Murzi, 2013]. Let us write $V(\langle A\rangle, \langle B\rangle)$ to express the fact that the inference from A to B is valid. By standard techniques of self-reference, we can construct a sentence, D, of the form $V(\langle D\rangle, \langle\bot\rangle)$, where $\bot$ is a logical constant such that, in *LP*, $\bot \models A$, for any A. Then there is a natural argument for $\bot$. This argument fails in the models we are dealing with, since they are closed under *LP* consequence, and they are not trivial. Where the argument breaks down may depend on exactly how it is formulated; but essentially, it fails due to the invalidity of material detachment.[10]

5. Validity and Detachment

In the last couple of sections, I will take up two issues arising from this account of validity. In what follows, I will stick to the one-premise case of validity for simplicity. The considerations clearly carry over the general case.

The first issue concerns the fact that if we define validity in the way described, the connection between the premises of a valid inference and its conclusion is only a material one.[11] The definition of validity has the following form:[12]

- $\forall I(I \Vdash^+ A \supset I \Vdash^+ B)$

Even though an inference is valid, then, the move from $I \Vdash^+ A$ to $I \Vdash^+ B$ is not valid in *LP*. And if I is the standard model, the same goes for truth *simpliciter*. This does not mean that no inference from the first to the second is possible. Failure of detachment occurs in *LP* only when the antecedent of the conditional is both true and false. Hence the move is legitimate provided that this is not the case. This observation can be built into a formal non-monotonic logic, *LPm*, in which the inference from C and $C \supset D$ to D is a valid default

[10] For further discussion, see [Priest, 2017, § 4.2].

[11] The following comes from [Priest, 2017, § 14].

[12] $\Vdash^+$ is a relationship between an interpretation and a sentence, so it would be more correct to write the relation as $I \Vdash^+ \langle A\rangle$. However, here and in what follows I omit the quotation marks, as logicians usually do.

inference.[13] Using LPm, we can then move from $I \Vdash^+ A$ to $I \Vdash^+ B$ provided that we do not have $I \nVdash^+ A$ as well (which is quite different from $I \Vdash^+ \neg A$).

However, it remains the case that this is a default inference. What is the significance of this? This depends on how one understands the model-theory. A straightforward way to understand the model-theoretic definition of validity is as specifying the *meaning* of 'valid'. In this case, even valid deductive inferences are, in the last instance, default inferences.

It might be thought odd to have the validity of a deductive inference grounded in a defeasible inference. But a little thought may assuage this worry. The difference between a material $I \Vdash^+ A \ \supset \ I \Vdash^+ B$ and a detachable $I \Vdash^+ A \rightarrow I \Vdash^+ B$ is not as great as might be thought. Both are simply true (or false) statements. Inference, by contrast, is an action. Given the premises of an argument, an inference is a jump to a new state. No number of truths is the same thing as a jump. (This is the moral of Lewis Carroll's celebrated dialogue between Achilles and the Tortoise [Carroll, 1895].) None the less, truths of a certain kind may ground the jump, in the sense of making it a reasonable action. There is no reason why a sentence of the form $C \supset D$ may not do this, just as much as one of the form $C \rightarrow D$. It is just that one of the latter kind always does so, while one of the former kind does so only sometimes (normally).

If it is not clear how a defeasible warrant for an action can work, merely consider sentences of the form:

(*) You promised to do x.

The truth of (*) is normally a ground for doing x, in the sense of making it reasonable to do so. But, to use a celebrated example, suppose that (*) is true, where the x in question is the returning of a weapon to a certain person. And suppose that the person comes requesting the weapon, but you know that they intend to use it to commit murder. Then the truth of (*) does not, in this context, ground the action. So with validity and the material conditional: the truth of a sentence of a certain kind may ground an appropriate action in normal circumstances, but fail to do so in unusual circumstances.

Another way to take the model-theoretic account of validity is as providing, not the meaning of 'valid', but merely an extensional characterisation of what is valid. The meaning of 'valid' itself can be characterised in a different way, say proof-theoretically (or simply taken as an indefinable primitive). The model-theoretic account merely gives us a characterisation of what inferences are or are not deductively valid — nothing more. Valid inferences can then simply license detachment of their conclusions, though this aspect of things may not be captured by the characterisation. In a similar way, an inferentialist who

[13] See the second edition of [Priest, 1987, ch. 16].

takes validity to be defined in terms of the meanings of the logical constants, spelled out in terms of introduction and elimination rules, may yet hold that a model-theoretic definition of validity delivers an extensionally equivalent characterisation (if sound and complete), though this may miss aspects of validity itself.

6. Dialetheic Validity

The second issue I will take up concerns the extent to which validity is itself a dialetheic notion. Let us suppose that we are working within one of the models of the kind we saw to exist in §4; and let us suppose that the model does verify the existence of a standard model.

The following argument is due, in effect, to Young [Young, 2005]. Let M be the standard model. Then given the resources of self-reference, we can find a sentence, D, of the form $M \nVdash^+ D$. The facts about the standard model then deliver:

- $M \Vdash^+ D \ \equiv\ M \nVdash^+ D$

It follows in LP that $M \Vdash^+ D\ \wedge\ M \nVdash^+ D$. Since $A, \neg B \models \neg(A \supset B)$, it also follows that

- $\neg(M \Vdash^+ D\ \supset\ M \Vdash^+ D)$

So $\exists x \neg(x \Vdash^+ D\ \supset\ x \Vdash^+ D)$, i.e., $\neg\forall x(x \Vdash^+ D\ \supset\ x \Vdash^+ D)$. That is, the inference from D to D is invalid. It follows that $p \not\models p$, since p has an invalid substitution instance — even though $p \models p$ as well. Perhaps this is not surprising. Truth is intimately connected with validity — at least when we have a standard model around. So one might expect self-referential constructions to deliver inconsistencies concerning validity; and the inference from p to p is not a terribly useful one!

It might be thought that Young's argument can be extended to establish that other valid inferences are also invalid. Thus, consider, for example, the inference from $p \wedge q$ to p. If we could show that:

- [1] $\neg(M \Vdash^+ D \wedge D\ \supset\ M \Vdash^+ D)$

we would have a similar counter-example. Now, the truth conditions for conjunction give us:

- $I \Vdash^+ D \wedge D\ \equiv\ I \Vdash^+ D\ \wedge\ I \Vdash^+ D$

or equivalently:

- [2] $I \Vdash^+ D \wedge D\ \equiv\ I \Vdash^+ D$

In LP material equivalents are inter-substitutable, in the sense that $A \equiv B \models C(A) \equiv C(B)$.[14] So from [2], when I is M, we have:

- $\neg(M \Vdash^+ D \supset M \Vdash^+ D) \equiv \neg(M \Vdash^+ D \wedge D \supset M \Vdash^+ D)$

But we cannot infer [1] because of the failure of detachment for the material biconditional. The inference is not even a valid default inference, since we know that the left hand side is contradictory. The same sort of problem is going to beset similar extensions of Young's argument to other inferences.

At present, it is not known how inconsistent validity is on the approach under consideration. The natural generalisation of Young's argument does not go through, but that does not mean that there are no others; and at present, there are no arguments which establish that the domain of inconsistency concerning validity is bounded, of the kind that show this for truth. This is, hence, an area where more work is required.

But let us suppose a worst case scenario: *every* inference is invalid. How damaging a conclusion would this be? Less than one might have thought. (Certainly, much less than a conclusion to the effect that every sentence is not true.) One should remember that every inference is an instance of *some* formally invalid inference (e.g., $p \vdash q$). An inference is acceptable if it is a substitution instance of *some* formally valid inference. Thus, it is perfectly acceptable to use an inference that is formally valid — it's a substitution instance of itself — even if it is invalid too![15]

[14] See [Priest, 2017, 2.3].

[15] I note that there is a way to avoid conclusions to the effect that some valid inferences are invalid, whether they are arguments of Weber's kind or of Young's kind. The definition of semantic validity given in [Priest, 1987, 5.2], is slightly different from the one considered above. It uses a truth predicate, and amounts to this:

- $A \models B$ is $\forall I\, T \langle I \Vdash^+ A \Rightarrow I \Vdash^+ B \rangle$

(where $\Rightarrow$ is either $\rightarrow$ or $\supset$, depending on how one thinks of the underlying set theory). Of course, given the T-Schema, this makes no difference to what is valid. However, it may well make a difference to what is invalid. One can establish that $\neg(M \Vdash^+ D \Rightarrow M \Vdash^+ D)$, and so $T \langle \neg(M \Vdash^+ D \Rightarrow M \Vdash^+ D) \rangle$; but one cannot establish that $\neg T \langle M \Vdash^+ D \Rightarrow M \Vdash^+ D \rangle$ unless negation commutes with truth; that is, unless the T-schema contraposes. There are reasons to suppose that it does not. See [Priest, 1987, 4.6]. A reader of the 1987 text would probably not even have noticed the use of truth in the definition there, or might have supposed its use to be merely stylistic. It would not be unreasonable to do so. Indeed, I myself have ignored this subtlety ever since. However, the use of the truth predicate was not an innocent one: I phrased the definition like this precisely because I thought that arguments of the kind in question might be possible.

7. Conclusion

We have now looked at many of the most important aspects of a dialetheic account of metatheory. While it can hardly be claimed that all of these are resolved, the project seems in a more than satisfactory state.

Acknowledgements. Many thanks go to Zach Weber for comments on an earlier draft of this paper. A version was read at the UNILOG conference, Vichy, June 2018.

References

Beall, Murzi, 2013 – Beall, J., Murzi, J. "Two Flavors of Curry's Paradox", *The Journal of Philosophy*, 2013, Vol. 110, pp. 143–165.

Cantini, 2003 – Cantini, A. "The Undecidability of Grišin's Set Theory", *Studia Logica*, 2003, Vol. 74, pp. 345–368.

Carroll, 1895 – Carroll, L. "What the Tortoise Said to Achilles", *Mind*, 1895, Vol. 4, pp. 278–280.

Kreisel, 1967 – Kreisel, G. "Informal Rigour and Completeness Proofs", in: *Problems in the Philosophy of Mathematics*, ed. by I. Lakatos, Amsterdam: North Holland, 1967, pp. 138–157.

Kripke, 1975 – Kripke, S. "Outline of a Theory of Truth", *Journal of Philosophy*, 1975, Vol. 72, pp. 690–716.

Priest, 1987 – Priest, G. *In Contradiction*, Dordrecht: Martinus Nijhoff; 2nd ed., Oxford: Oxford University Press, 2006.

Priest, 1997 – Priest, G. "On a Paradox of Hilbert and Bernays", *Journal of Philosophical Logic*, 1997, Vol. 26, pp. 45–56.

Priest, 1998 – Priest, G. "The Trivial Object, and the Non-Triviality of a Semantically Closed Theory with Descriptions", *Journal of Applied and Non-Classical Logics*, 1998, Vol. 8, pp. 171–183.

Priest, 1999 – Priest, G. "Semantic Closure, Descriptions, and Non-Triviality", *Journal of Philosophical Logic*, 1999, Vol. 28, pp. 549–558.

Priest, 2002 – Priest, G. "Paraconsistent Logic", in: *Handbook of Philosophical Logic*, Vol. 6, D. Gabbay, F. Guenthner (eds.), Dordrecht: Kluwer Academic Publishers, 2002, pp. 287–393.

Priest, 2005 – Priest, G. *Towards Non-Being*, Oxford: Oxford University Press; 2nd ed., 2016.

Priest, 2008 – Priest, G. *Introduction to Non-Classical Logic: From If to Is*, Cambridge: Cambridge University Press, 2008.

Priest, 2017 – Priest, G. "What *If*: the Exploration of an Idea", *Australasian Journal of Logic*, 2017, Vol 14:1, article 1, `https://ojs.victoria.ac.nz/ajl/article/view/4028/3574`.

Tarski, 1935 – Tarski, A. "Der Wahrheitsbegriff in den Formalisierten Sprachen", *Studia Philosophica*, 1935, Bd. 1, S. 261–405. Translated as "The Concept of

Truth in Formalized Languages", in: *Logic, Semantics, Metamathematics*, ed. by J. Corcoran, 2nd ed., Indianapolis: Hackett, 1983, pp. 152–278.

Weber, 2010 – Weber, Z. "Transfinite Numbers in Paraconsistent Set Theory", *Review of Symbolic Logic*, 2010, Vol. 3, pp. 71–92.

Weber, 2012 – Weber, Z. "Transfinite Cardinals in Paraconsistent Set Theory", *Review of Symbolic Logic*, 2012, Vol. 5, pp. 269–293.

Weber, 2016 – Weber, Z. "What is an Inconsistent Truth Table?", *Australasian Journal of Philosophy*, 2016, Vol. 94, pp. 533–548.

Young, 2005 – Young, G. *Revenge: Dialetheism and its Expressive Limitations*, PhD Thesis, University of Glasgow, 2005.

Логические исследования
2020. Т. 26. № 1. С. 60–77
УДК 16

Logical Investigations
2020, Vol. 26, No. 1, pp. 60–77
DOI: 10.21146/2074-1472-2020-26-1-60-77

Gabriel Sandu

Gödelian sentences and semantic arguments

Gabriel Sandu
University of Helsinki,
P.O. Box 24 (Unioninkatu 40 A), 00014 Helsinki, Finland.
E-mail: gabriel.sandu@helsinki.fi

Abstract: This paper contains some philosophical reflections on Gödelian (undecidable) sentences and the recognition of their truth using semantic arguments. These reflections are not new, similar matters have been extensively addressed in the philosophical literature. The matter is rather one of emphasis.

Keywords: Gödelian sentences, Gödel's incompleteness theorem, semantical argument, truth theory, arithmetic, proof, provability

For citation: Sandu G. "Gödelian sentences and semantic arguments", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 60–77. DOI: 10.21146/2074-1472-2020-26-1-60-77

To the memory of Alexandr Karpenko, such a great friend

1. Gödel incompleteness theorem

Let L be the *language of arithmetic*, consisting of

- variables, x_0,$x_1, x, y, ...$
- logical constants: $\neg, \vee, \exists x, =$
- nonlogical constants: $\mathbf{0}, \mathbf{S}, +, \times$.

(Here '**0**' is an individual constant, '**S**', is a one-place function symbol and '+', and '×' are two place function symbols.)

From these items, the terms and formulas of the language of L are formed in the standard way.

As Tarski observed, the object language of a formalized science, comes together with a theory, usually given by listing its axioms and rules of inference. In our case the starting point is the theory Q (minimal arithmetic) which is the set of logical consequences of the following axioms:

1. $\forall x \forall y(\boldsymbol{S}x = \boldsymbol{S}y \rightarrow x = y)$

2. $\forall x(\boldsymbol{S}x \neq \mathbf{0})$

3. $\forall x(x \neq \mathbf{0} \rightarrow \exists y(s = \boldsymbol{S}y))$

4. $\forall x(x + \mathbf{0} = x)$

5. $\forall x \forall y(x + \boldsymbol{S}y = \boldsymbol{S}(x + y))$

6. $\forall x(x \times \mathbf{0} = \mathbf{0})$

7. $\forall x \forall y(x \times \boldsymbol{S}y = (x \times y) + x)$.

Notice that this theory is finitely axiomatizable. The language of Q is interpreted in a metalanguage in which '$\mathbf{0}$' is assigned the the natural number zero, '**S**' is assigned the successor function, '$+$' is assigned the operation of addition '$\times$' is assigned multiplication. It is known that Q is a rather strong theory which is able to represent all recursive functions (in a technical sense of the notion of 'representation', which is assumed to be known. It is also known that Q defines (in a technical sense assumed to be known) its own syntax and many semantical notions. This happens, as shown by Gödel, via the notion of gödel numbering. As a result, each term t in the language L gets associated with a gödel number $\ulcorner t \urcorner$; and each formula A receives its gödel number $\ulcorner A \urcorner$. Recalling that every natural number m has a name $\underline{m}$ in L, where $\underline{m}$ is an abbreviation for (the numeral) $\underbrace{\boldsymbol{SS}...}\mathbf{0}$ (m times), we see that every term t and every formula A have names in the arithmetical language, $\underline{\ulcorner t \urcorner}$ and $\underline{\ulcorner A \urcorner}$, respectively. This fact, together with the ones mentioned earlier, makes possible to introduce, for any formula A in the language of arithmetic, the diagonalization of A, which is the expression

$$\exists x(x = \underline{\ulcorner A \urcorner} \wedge A).$$

When A is a formula with one free variable, then we see that asserting the diagonalization of A amounts to predicating A of its own gödel number.

From Gödel's results, it follows that for any theory T extending Q, the set of gödel numbers of theorems of T is not definable in T, from which it can be further inferred that the set of Gödel numbers of true arithmetical sentences ("true in the standard model") is not definable. This last statement is usually known as "Tarski's theorem"; it is somehow debatable in the literature whether Gödel himself was aware of this result or not, but this matter will not concern us here. The first statement is standardly proved by reductio using the diagonalisation lemma which asserts that for any theory T which extends Q, for any formula $B(y)$ there is a sentence A such that

$$T \vdash A \leftrightarrow \neg B(\underline{\ulcorner A \urcorner}).$$

The second statement follows directly from it, by observing that the set of true arithmetical sentences is an extension of Q.

The variant of the Gödel's incompleteness theorem we are interested in is proved by first showing that for every extension T of Q there is a formula $Pr_T(x)$ in the language of arithmetic which has the form $\exists y Prov_T(x, y)$ and is such that for any sentence A in the language of arithmetic:

- $T \vdash A$ if and only if $\exists y Prov_T(\underline{\ulcorner A \urcorner}, y)$ is true (in the standard model) if and only if for some natural number m, $Prov_T(\underline{\ulcorner A \urcorner}, \underline{m})$ is true if and only if (given the representability of $Prov_T$ in Q), $Q \vdash Prov_T(\underline{\ulcorner A \urcorner}, \underline{m})$ for some m.

Here $Prov_T(x, y)$ is a primitive recursive formula, that is, a formula which contains only bounded quantifiers and is closed under the standard propositional connectives. Thus, from the above we get that if $T \vdash A$ then $Q \vdash Prov_T(\underline{\ulcorner A \urcorner}, \underline{m})$ for some m, and given that T is an extension of Q we also get $T \vdash \exists y Prov_T(\underline{\ulcorner A \urcorner}, y)$, i.e., $T \vdash Pr(\underline{\ulcorner A \urcorner})$. Now applying the Diagonalization lemma to the formula $\exists y Prov_T(\underline{\ulcorner A \urcorner}, y)$ Godel showed that there is a sentence, usually denoted by G such that

$$T \vdash G \leftrightarrow \neg \exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$$

The sentence G is called a *Gödel sentence for T*. It is taken to say: "I am unprovable".

We recall that a theory T is called $\omega-$inconsistent if there is a formula $F(x)$ such that $T \vdash \exists x F(x)$ but $T \vdash \neg F(\underline{0})$, $T \vdash \neg F(\underline{1})$, $T \vdash \neg F(\underline{2})$,...(for every natural number $0, 1, 2, ...$). T is called $\omega-$consistent if it is not $\omega-$inconsistent. Now Gödel proved

Theorem 1. *(Gödel First Incompleteness Theorem). Let T be a consistent, axiomatizable extension of Q and let G be a Gödel sentence for T. Then $T \nvdash G$. If T is $\omega-$consistent, then $T \nvdash \neg G$.*

The proof is well known but we rehearse it here (we follow Boolos, Jeffrey and Burgess), because it serves as a basis for extracting, later on, a semantic argument. Suppose that $T \vdash G$. Hence, by our previous comments, $\exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$ is true (in the standard model) and by a well known result, $Q \vdash \exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$; given that T is an extension of Q we also have

$T \vdash \exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$. From the Diagonalization lemma we also know that $T \vdash \neg\exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$. Thus T is inconsistent, a contradition. Hence $T \nvdash G$. For the second claim, suppose that $T \vdash \neg G$. By the diagonalization lemma, $T \vdash \exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$. But given that T is consistent and $T \vdash \neg G$, we must have $T \nvdash G$. This implies that for no natural number n, n is the code of a proof of G in T, that is, $\neg Prov_T(\underline{\ulcorner G \urcorner}, \underline{0})$, $\neg Prov_T(\underline{\ulcorner G \urcorner}, \underline{1})$, $\neg Prov_T(\underline{\ulcorner G \urcorner}, \underline{2})$..., are all true (in the standard model), where each of these formulas are primitive recursive. Hence $Q \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{0})$, $Q \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{1})$, $Q \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{2})$.... and since T is an extension of Q we also have $T \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{0})$, $T \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{1})$, $T \vdash \neg Prov(\underline{\ulcorner G \urcorner}, \underline{2})$.... Hence T is $\omega-$inconsistent, which contradicts our assumption. We conclude $T \nvdash \neg G$.

After reviewing these results, let us return to the question which is the main concern in this paper, namely Gödel's method to produce undecidable sentences such as G, and especially a claim often made in this connection to the effect that these sentences are true and *recognized to be true*. Here is, for instance, how Dummett describes Gödel's result:

> By Gödel's theorem there exists, for an intuitively correct formal system for elementary arithmetic, a statement [G] expressible in the system but not provable in it, which not only is true but can be recognized by us to be true... [Dummett, 1963].

The puzzling question is: how do we "recognize" that G (or any statement equivalent to it) is true?

The above proof of the theorem does not give an explicit argument about how we come to recognize G as true, neither did Gödel provide one. But it is not very difficult to extract one. From the Diagonalization lemma we know that the statement G is equivalent to a universal statement, viz. $\neg\exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$ (i..e $\forall y \neg Prov_T(\underline{\ulcorner G \urcorner}, y)$). From the second part of the proof we see that every numerical instance is provable (and true) in the system. Since G is the universal quantification over all these numerical instances, then G is true. Of course in this last step we rely on our grasp of the standard model (this is what the $\omega-$consistency is supposed to ensure).

In fact, this is Dummett's argument for the truth of Gödel's sentence:

> The statement [G] is of the form $\forall x A(x)$,where each one of the statements $A(\underline{0})$,$A(\underline{1})$,$A(\underline{2})$, ...is true: since $A(x)$ is recursive, the notion of truth for these statements is unproblematic. Since each of the statements $A(\underline{0})$,$A(\underline{1})$,$A(\underline{2})$, ...is true in every model of the formal system, every model of the system in which G is false must be a non-standard model...whenever, for some predicate $B(x)$, we

> can recognize all of the statements $B(\underline{0})$,$B(\underline{1})$,B$A(\underline{2})$, ...as true in the standard model, then we can recognize that $\forall x A(x)$ is true in that model. This fact ...we know on the strength of our clear intuitive conception of the structure of the model [Dummett, 1963, p. 191].

As we see from this quote, we come to appreciate that the undecidable Gödel sentence G for Q is true not by working inside the system but rather by conducting a so called *semantical argument* which makes an essential use of the concept of truth itself. Dummett is not the only one to have seen the importance of semantical arguments. There is another semantical argument which uses the truth predicate, distinct from Dummett's argument, which goes back to Alfred Tarski [Tarski, 1956]. In order to present it, we need to say somehting about arithmetical induction.

The system Q of minimal arithmetic is knowingly defficient in that it fails to prove many universal statements about numbers which are usually proved by mathematical induction. Typically, if we want to prove that every number has a given property, we prove it by showing that 0 has that property, and then we show, from the assumption that an arbitrary number x has that property, that the successor Sx has that property. To accommodate induction one needs a more adequate set of axioms for number theory. To this effect we add to the 7 axioms of the system Q all sentences of the form

8. $[A(\mathbf{0}) \wedge \forall x(A(x) \to A(S(x))] \to \forall x A(x)$

(8) is usually known as the *Induction axiom scheme.* The theory which is the set of all sentences in the language of arithmetic which are logical consequences of (1)–(8) is known as Peano Arithmetic (PA). It is a simple mathematical fact that definability and representability in Q entail definability and representability in any extension of Q and thus in PA in particular. From now on we shall operate with PA. Tarski's semantical argument which proves the truth of the Gödelian statement G for PA, uses a universal statement which cannot be proved in Q but needs PA.

1.1. The representability of the syntax in arithmetic

Tarski's truth-definition for arithmetic exploits the representability of the syntax of PA in PA.

It is a mathematical fact that there are functions $f_\neg, f_\vee, f_\exists$ defined on the natural numbers such that the following hold:

- $f_\neg(\ulcorner A \urcorner) = \ulcorner \neg A \urcorner$, for every formula A in the object language;
- $f_\vee(\ulcorner A \urcorner, \ulcorner B \urcorner) = \ulcorner A \vee B \urcorner$, for every formulas A, B in the object language;
- $f_\exists(\ulcorner A \urcorner, n) = \ulcorner \exists x_n A \urcorner$, for every formula A and natural number n.

There is also a function f_{sub} (the susbstitution function) which has the property:

$$f_{sub}(\ulcorner A \urcorner, \ulcorner x_i \urcorner, \ulcorner t \urcorner) = \ulcorner A(t) \urcorner$$

for every formula A in the language of arithmetic, variable x_i and term t in the same language.

All these functions are recursive, thus representable in Q and hence in PA which means there are formulas $Neg(x, y), Dis(x, y, z), Ex(x, y, z)$ and $Sub(x, y, z, w)$ in the language of arithmetics so that for all formulas A, B, term t, and natural number n we have

a) $PA \vdash \forall y \, (Neg(\underline{\ulcorner A \urcorner}, y) \leftrightarrow y = \underline{\ulcorner \neg A \urcorner})$

b) $PA \vdash \forall y \, (Dis(\underline{\ulcorner A \urcorner}, \underline{\ulcorner B \urcorner}, y) \leftrightarrow y = \underline{\ulcorner A \vee B \urcorner})$

c) $PA \vdash \forall y \, \big(Ex(\underline{\ulcorner A \urcorner}, \underline{n}, y) \leftrightarrow y = \underline{\ulcorner \exists x_n A \urcorner}\big)$

d) $PA \vdash \forall y \, \Big(Sub(\underline{\ulcorner A \urcorner}, \underline{\ulcorner x_i \urcorner}, \underline{\ulcorner t \urcorner}, y) \leftrightarrow y = \underline{\ulcorner A(t) \urcorner}\Big)$

Similarly, the function $f_=$ on the natural numbers such that

$$f_= (\ulcorner t \urcorner, \ulcorner s \urcorner) = \ulcorner t = s \urcorner$$

for all terms t, s in the language of arithmetic is representable in PA by, say, the expression $Id(x, g, z)$, that is,

$$PA \vdash \forall y \, (Id(\ulcorner t \urcorner, \ulcorner s \urcorner, y) \leftrightarrow y = \ulcorner t = s \urcorner) \, .$$

If in (a) we instantiate y with $\underline{\ulcorner \neg A \urcorner}$ we get

$$PA \vdash Neg(\underline{\ulcorner A \urcorner}, \underline{\ulcorner \neg A \urcorner}) \leftrightarrow \underline{\ulcorner \neg A \urcorner} = \underline{\ulcorner \neg A \urcorner}.$$

The formula on the right side is a theorem of the predicate calculus (with identity), hence PA proves it. Thus $PA \vdash Neg(\underline{\ulcorner A \urcorner}, \underline{\ulcorner \neg A \urcorner})$. We can show that for each formula A of the object language there is exactly one formula B of the object laguage such that $PA \vdash Neg(\underline{\ulcorner A \urcorner}, \underline{\ulcorner B \urcorner})$ and B is $\neg A$. Therefore we can take Neg to be a function and write $Neg(\underline{\ulcorner A \urcorner}) = \underline{\ulcorner \neg A \urcorner}$.

In a similar way we can also take $Dis, Ex, Sub, Id, Less$ to be also functions. Thus we shall have

a*) $PA \vdash Neg(\underline{\ulcorner A \urcorner}) = \underline{\ulcorner \neg A \urcorner}$, for every formula A in the object language.

b*) $PA \vdash Dis(\underline{\ulcorner A \urcorner}, \underline{\ulcorner B \urcorner}) = \underline{\ulcorner A \vee B \urcorner}$, for every formulas A, B in the object language

c*) $PA \vdash Ex(\underline{\ulcorner A \urcorner}, \underline{n}) = \underline{\ulcorner \exists x_n A \urcorner}$, for every formula A in the object language and natural number n.

d*) $PA \vdash Sub(\underline{\ulcorner A \urcorner}, \underline{\ulcorner x_i \urcorner}, \underline{\ulcorner t \urcorner}) = \underline{\ulcorner A(t) \urcorner}$, for every formula A and term t of the object language and every natural number i.

e*) $PA \vdash Id(\ulcorner t \urcorner, \ulcorner s \urcorner) = \ulcorner t = s \urcorner$, for all terms t, s of the object language.

In a similar way it can be shown that PA defines its own syntax: being a closed term, a variable, a formula and a sentence (of the language of arithmetic). That is, there are formulas $ct(x), var(x), form(x)$ and $sen(x)$ in the object language such that the following holds:

f) $PA \vdash ct(\underline{\ulcorner t \urcorner})$, for every closed term t.

g) $PA \vdash var(\underline{\ulcorner x_i \urcorner})$, for every natural number i.

h) $PA \vdash form(\underline{\ulcorner A \urcorner})$, for every formula A.

j) $PA \vdash sen(\underline{\ulcorner A \urcorner})$, for every closed sentence A.

PA also defines some semantical properties. There is a formula $Den(x)$ in the object language (that we can take to be a function) such that

k) $PA \vdash t = s \leftrightarrow Den(\underline{\ulcorner t \urcorner}) = Den(\underline{\ulcorner s \urcorner})$, for all terms t, s in the object language.

2. Tarski's truth theory

In the case of Tarski's truth theory for arithmetic we do not need to go via the notion of satisfaction but use directly the truth-predicate Tr. The reason for this is that each natural number has a name in the object language.

The axioms of the truth-definition are given in the metalanguage containing Tr is a predicate symbol:

Ax1 $\forall x(Tr(x) \rightarrow sen(x))$

(If x is true, then x is of a sentence)

Ax2 $\forall x \forall y(ct(x) \wedge ct(y) \rightarrow (Tr(Id(x, y)) \leftrightarrow Den(x) = Den(y)))$

(The identity between two closed terms x and y is true iff their denotations are the same)

Ax3 $\forall x(Sen(x) \rightarrow (Tr(Neg(x)) \leftrightarrow \neg Tr(x)))$

(The negation of the sentence is true iff the sentence is not true)

Ax4 $\forall x \forall y(sen(x) \wedge sen(y) \rightarrow (Tr(Dis(x,y)) \leftrightarrow Tr(x) \vee Tr(y)))$

(A disjunction is true iff either sentence is true)

Ax5 $\forall x_1 \forall x_2(form(x_1) \wedge var(x_2) \rightarrow (Tr(Ex(x_1,x_2)) \leftrightarrow \exists t(Tr(Sub(x_1,x_2,t))))$

(An existential sentence is true iff there is a closed term t such that the sentence which is the result of the substitution of the free variable x_2 in x_1 by t is true.)

Let $PA(Tr)$ be the set of sentences which are the logical consequences of the 7 axioms of PA, the five axioms (Ax1)–(Ax5), and plus the Induction schema (8) which allows occurrences of the truth-predicate in the formulas $A(x)$. It can be shown that $PA(Tr)$ is materially adequate, that is,

$$PA(Tr) \vdash Tr(\underline{\ulcorner A \urcorner}) \leftrightarrow A,$$

for any sentence A in the language of arithmetic.

It is well known that the Tarskian truth theory proves the following universal statements:

- *The principle of noncontradiction (consistency).* For every sentence y of the object language it is not the case that both y and its negation are true:
$$PA(Tr) \vdash \forall y\,(Sen(y) \rightarrow \neg(Tr(y) \wedge Tr(neg(y))))\,.$$
This property follows directly from Ax3.

- *The principle of excludded middle.* Every sentence of the object language is true ot its negation is true:
$$PA(Tr) \vdash \forall y\,(Sen(y) \rightarrow Tr(y) \vee Tr(neg(y)))\,.$$
This property follows from the other direction of Ax3.

- *The principle of soundness.* All theorems are true:
$$PA(Tr) \vdash \forall x(Pr_{PA}(x) \rightarrow Tr(x)).$$

This principle fully exploits the occurrence of the truth-predicate in the Induction scheme. We omit its proof but it consists, informally, of the following steps:

1. All the axioms of PA are true.

2. The rules of inference of PA preserve truth.

3. Hence every theorem of PA is true (i.e. $PA(Tr) \vdash \forall x(Pr_{PA}(x) \rightarrow Tr(x))$.

2.1. Tarski's semantical argument

In the postscript to the English translation of his seminal article, Tarski adds some interesting parallels between his results and those of Gödel:

> Moreover Gödel has given a method for constructing sentences which- assuming the theory concerned to be consistent- cannot be decided in any direct way in this theory. All sentences constructed according to Gödel's method possess the property it can be established whether they are true or false on the basis of the metatheory of higher order having a correct definition of truth [Tarski, 1956, p. 274].

To establish the truth of such a Gödelian sentence Tarksi uses the principle of soundness listed in the previous sesction. We present Tarski's semantical argument (Tarski, 1936, Theorem 5) for the Gödelian sentence $\neg Pr_{PA}(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner})$ (that we shall abbreviate by Con_{PA}) which is taken to express the consistency of PA. The semantical argument for G is similar. There is nothing original in my presentation, this argument has been rehearsed many times [Ketland, 1999] and [Shapiro, 1998].

Gödel's second incompleteness theorem shows that $PA \nvdash Con_{PA}$ and $PA \nvdash \neg Con_{PA}$. But Tarski shows

$$PA(Tr) \vdash Con_{PA}.$$

The argument is straightforward. From the soundness principle we get

$$(i) \quad PA(Tr) \vdash Pr_{PA}(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner}) \rightarrow Tr(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner}).$$

We also know that the theory of truth proves all the T-instances, i.e.,

$$(ii) \quad PA(Tr) \vdash Tr(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner}) \leftrightarrow \neg \mathbf{0} = \mathbf{0}.$$

But PA proves $\mathbf{0} = \mathbf{0}$, and thus $PA(Tr) \vdash \mathbf{0} = \mathbf{0}$, which together with (ii) entails

$$(iii) \quad PA(Tr) \vdash \neg Tr(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner}).$$

From (i) and (iii) we get

$$(iv) \quad PA(Tr) \vdash \neg Pr_{PA}(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner})$$

that is, $PA(Tr) \vdash Con_{PA}$.

Tarski's *semantical argument* is usually expressed in words, in order to enhance its explanatory power:

- In a first step we establish the principle of soundness as we showed earlier:

1. All the axioms of PA are true.
2. The rules of inference of PA preserve truth.
3. Hence every theorem of PA is true,

$$PA(Tr) \vdash \forall x(Pr_{PA}(x) \to Tr(x)).$$

- A second step established that the sentence '$\neg \mathbf{0} = \mathbf{0}$' is not true:

$$PA(Tr) \vdash \neg Tr(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner})$$

(see (iii))

- In a third step we combined the conclusion of the first and of the second step and concluded that '$\neg \mathbf{0} = \mathbf{0}$' is not a theorem:

$$PA(Tr) \vdash \neg Pr_{PA}(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner})$$

(see (iv))

- Finally we note that $\neg Pr_{PA}(\underline{\ulcorner \neg \mathbf{0} = \mathbf{0} \urcorner})$ is the Consistency statement Con_{PA}.

The crucial role in this argument is the universal generalization which is the Principle of soundness. It confers the semantic argument the form of a nomological argument which shows the *explanatory role of the truth predicate:*

> Let us return to the Gödelian statement G (or Con_{PA}). Let us suppose a logic teacher asserts that Con_{PA} is true, and the puzzled student asks for an explanation. The student believes the teacher's word that Con_{PA} is true, but he wants to be shown why Con_{PA} is true. The student wants something like a convincing proof or an explanatory proof. The natural answer is to remark that all the axioms of PA are true and the rules of inference preserve truth. Thus every theorem of PA is true. It follows that '$\neg 0 = 0$' is not a theorem and thus PA is consistent.... It seems to me that this informal version of the derivability of Con_{PA} is as good an *explanation* as there is. The argument shows why Con_{PA} is true or why Con_{PA} is a consequence- and the move through the notion of truth provides the explanation [Shapiro, 1998, p. 505].

3. Feferman's program

Tennant [Tennant, 2002] argues against Ketland [Ketland, 1999] and Shapiro [Shapiro, 1998] that Tarski's theory of truth is not the only way we can come to recognize the truth of the Gödel sentence. In particular, Tennant claims, the generalization "All theorems are true" is not the only way to express the soundness of an arithmetical system S. There is, instead, another way to express it, viz., using reflection principles of the form

(pa) If $\overline{\varphi}$ is a primitive recursive sentence and $\overline{\varphi}$ is provable in S, then φ.

As we see, this reflection principle does not use the truth-predicate. Tennnat follows here Feferman [Feferman, 1962], who emphasizes that "Reflection principles are axioms schemata ...which express, insofar as is possible without use of the formal notion of truth, that whatever is derivable in S is true".

Let us take stock. We have discussed two semantic arguments invoked in how we come to recognize that Gödelain sentences are true.

One such argument, due to Tarski, and explicitly described in Shapiro's quote in the last section, uses the generalization "All theorems are true" and can be run in an extension $PA(Tr)$ of PA which, in addition to the truth axioms, allows occurrences of the truth predicate in the induction scheme.

The other semantic argument, described earlier in the second quote from Dummett also uses the truth-predicate. However, Tennant [Tennant, 2002] rephrases it, so that the reference to "the structure of the model" is deleted and the truth-predicate lifted out as required by Feferman's reflection principles. Here is Tennant's formulation of his own semantic argument:

> G is a universally quantified sentence (as it happens, one of Goldbach type, that is, a universal quantification of a primitive-recursive predicate). Every numerical instance of that predicate is provable in the system S. (This claim requires a subargument exploiting Gödel numbering and the representability in S of recursive properties.) Proof in S guarantees *truth*. Hence every numerical instance of G is *true*. So, since G is simply the universal quantification over those numerical instances, it too must be *true* [Tennant, 2002, p. 556].

Tennant shows that this argument can be faithfully represented in a "sufficiently strong" arithmetical system S enriched with reflection principles (with no occurrence of the truth-predicate) in Feferman's style.

I will now describe shortly the main lines of Tennant's argument. Before doing that let me mention what it means for a formal system of arithmetic S to be "sufficiently strong": S represents recursive properties and proves the Diagonalization lemma (i.e., there is a proof in S leading from G to $\neg\exists y Prov_T(\underline{\ulcorner G \urcorner}, y)$;

and there is a proof in the other direction too), and S also proves the equivalence between the Gödelian sentence G and the consistency sentence Con_S. It is known that there are several systems which satisfy this requirement, e.g. PA.

Tennant proposes an extension of S with Feferman's *principle of uniform primitive recursive reflection* (which is more general than the principle (pa) mentioned above):

(UR) Add to S all sentences of the form

$$\forall n(Pr_S(\underline{\ulcorner \psi(\underline{n}) \urcorner})) \rightarrow \forall m \psi(m)$$

where ψ is a primitive recursive formula and $\underline{n}$ is, as before the numeral corresponding to the natural number $n$and $Pr_S(\underline{\ulcorner \psi \urcorner})$ is, like before, an abbreviation for $\exists y Prov_S(\underline{\psi}, y)$

He then shows that in this extension a faithful formalization of the semantical argument described above can be run. The proof goes like this [Tennant, 2002, p. 577]. (We let S^* denote the system S plus (UR)).

Suppose m codes a proof of G in S. Hence by representability (a natural number being the code of a proof in S of a formula is a primitive recursive relation), $S \vdash Prov_S(\underline{\ulcorner G \urcorner}, \underline{m})$, where $Prov_S$ is a primitive recursive formula. But S proves also, from the assumption G, the sentence $\forall y \neg Prov_S(\underline{\ulcorner G \urcorner}, y)$ (i.e. the diagonalization lemma), which by universal instantiation implies $\neg Prov_S(\underline{\ulcorner G \urcorner}, \underline{m})$. Given our assumption that S is consistent, we have a contradiction, from which we conclude that m does not code a proof of G in S. Again by representability we get $S \vdash \neg Prov_S(\underline{\ulcorner G \urcorner}, \underline{m})$. But n has been chosen arbitrarily, hence for every n, there is some proof of $\neg Prov_S(\underline{\ulcorner G \urcorner}, \underline{n})$ in S, from which with the help of (UR) we derive (in S^*) that $\forall y \neg Prov_S(\underline{\ulcorner G \urcorner}, y)$. Finally, by the Diagonalization Lemma, we get G (in S^*).

The penultinate steps requires perhaps some additional clarification. If I understood correctly, "for every n, there is some proof of $\neg Prov_S(\underline{\ulcorner G \urcorner}, \underline{n})$ in S" is just the sentence $\forall n Pr_S(\underline{\ulcorner \psi(\underline{n}) \urcorner})$ in the antecedent of (UR), where $\psi(\underline{n})$ is the primitive recursive sentence $\neg Prov_T(\underline{\ulcorner G \urcorner}, \underline{n})$.

We are then told:

> The foregoing proof justifies the assertion of G. The stronger system S^* contains methods for reflecting on the justification resources of the weaker system S. These methods can be seen at work, in the application, in the proof just give, of various rules of inference that are available in S^* but not in S [Tennant, 2002, p. 577].

The thing which I find somehow problematic in the proof are the penultimate steps:

...But n has been chosen arbitrarily, hence for every n, there is some proof of $\neg Prov_S(\underline{\ulcorner G \urcorner}, \underline{n})$ in S, from which, with the help of (UR), we derive (in S^*) that $\forall y \neg Prov_S(\underline{\ulcorner G \urcorner}, y)$.

I take them to correspond to the informal steps of Tennant's own semantic argument listed earlier in this section. It seems to me that we can justify these steps only on the basis of our intuitive understanding of the standard model, as Dummett pointed out. The principle of uniform recursive reflection (UR) just expresses this understanding in a formal way. We may have eliminated the truth-predicate as required by a minimlist conception of truth, but the justification of (UR) is still grounded in such understanding. This matter is orthogonal to the goal of this essay, so I will not dwell on it.

One can still perhaps argue that Tarski's truth-definition is more general, because it can also account for the intuition that all $S-$theorems are true (sound), and not just the primitive recursive ones. Tennant's response to this objection is that we could add as well to S^* the schema (soundness principle)

$$Prov_S(\underline{\ulcorner \varphi \urcorner}) \rightarrow \varphi,$$

where φ is any sentence in the language of arithmetic. It is known from Löb's theorem that this principle cannot be derivable in S without making S inconsistent. But in the present case we add the soundness principle not to S directly but to S extended with the principle of uniform primitive recursive reflection, and this avoids the inconsistency.

To sum up, I agree with Tennant that the difference between the two semantic arguments is that between saying (Tarski) and showing (Feferman). That is, Tarski's truth theory can state the principle of soundness in one single universal statement "All theorems are true". In this case the "recognition" of the truth of the Gödelian sentence takes the form of a nomological explanation which uses that universal statement [Ketland, 1999; Shapiro, 1998]. On the other side, the Feferman-Tennant framework (S^* extended with the soundness axiom scheme) uses an axiom scheme which can be seen as a list of the infinitely many instances of the universal statement $\forall x(Pr_S(x) \rightarrow Tr(x)$:

$$\begin{gathered} Pr_S(\underline{\ulcorner \varphi_1 \urcorner}) \rightarrow Tr(\underline{\ulcorner \varphi_1 \urcorner}) \\ Pr_S(\underline{\ulcorner \varphi_2 \urcorner}) \rightarrow Tr(\underline{\ulcorner \varphi_2 \urcorner}) \\ \vdots \end{gathered}$$

in which the truth-predicate has been eliminated in virtue of the equivalences

$$\begin{gathered} Tr(\underline{\ulcorner\varphi_1\urcorner}) \leftrightarrow \varphi_1 \\ Tr(\underline{\ulcorner\varphi_2\urcorner}) \leftrightarrow \varphi_2 \\ \vdots \end{gathered}$$

In this case the recognition of the truth of G does not take the form of a nomological argument (because there is no collection of all these instances into one universal statement). It consists in the apprehension of the proof of G in the extension of e.g. PA with the soundness principle. Truth does not "transcend" proof, truth is just proof (in the extended system).

4. The justification of the extensions

A question which arises quite naturally at this stage is about the justification of different extensions which settle the Gödelian statements, and about the nature of these statements themselves. Is a given extension more justified than another? This question revives an older discussion which goes back to Gödel concerning intrinsic versus extrinsic extensions of a theory which has been the inspiring source for the Feferman program.

Gödel's reflections took place in the context of set theory (*What is Cantor's continuum problem*? [Gödel, 1947]) but they also apply *mutatis mutandis* to arithmetic. Gödel introduced a distinction between an *intrinsic* and *extrinsic* extension of an axiom system. An intrinsic extension, unlike an extrinsic one, is justified on the basis of one grasping the concepts of the base theory. Gödel gave as an example the *Axiom of Determinacy* in set theory that he regarded as an extrinsic axiom because it is not justified by our understanding of sets, in contrast to *Mahlo's axioms* for big cardinals. In addition, Gödel also mentioned intrinsic extensions with undecidable statements (Gödelian sentences) that one recognizes as true in virtue of their meaning, that is, by reflecting on their undecidability.

Gödel's remarks suggest the idea to treat the truth axioms of Tarski's theory of truth as examples of intrinsic extensions of the base theories, whose justification is grounded in our grasping of the concepts of the base theory, that is, natural numbers and operations on natural numbers. In fact this suggestion, which was not made by Gödel, has been explicitly advocated later on by Koellner in his reflections on Gödel's distinctions:

> Let us consider first our conception of natural numbers which is underlying PA. This conception of natural numbers not only justifies the principle of mathematical induction for the language of PA, but for any other extension of the language of PA which has

> a sense. For instance if we extend the language of PA by adding the tarskian truth-predicate and we extend the axioms of PA by adding the tarskan axioms for truth, then, on the basis of our conception of natural numbers, we are justified in accepting the instances of the induction scheme in which the truth-predicate occurs. In the resulting system one can prove Con_{PA}....By contrast, the Axiom of determinacy AD is not justified by our understanding of natural numbers [Koellner, 2006].

Similar ideas have been expressed by Feferman. Starting with the 60's and inspired by Gödel, he addressed the question of the extensions of *schematic formal systems* (formal systems which contain axiom shemes, like ZFC and PA) with new axioms. He started looking for the possibility to generate systematically extensions of such systems whose acceptance was already implicit in the base theory. One of the mechanisms Feferman proposed is *reflection principles.* We saw an illustration of this mechanism when presenting Tennant's ideas. Little by little Feferman also came to consider extensions which contains explicitly a truth-predicate and developed the notion of *reflexive closure of a schematic theory* [Feferman, 1991], which allows for the Induction scheme to range over the truth-predicate. In this case the extended system can prove statements of the form $\forall x(Pr_{PA}(x) \to Tr(x))$. This has been, as we saw, Tarski's way.

I think there is an important difference between Gödel's notion of intrinsic extension where the new axioms display or unfold the content of the notions of the base theory, and the two extensions of PA introduced in this paper. It seems to me that neither Tarski's extension of PA with his theory of truth, nor Tennant's extension of a sufficiently strong arithmetical system S (e.g. PA) with reflection principles $Prov_S(\underline{\ulcorner \varphi \urcorner}) \to \varphi$, "unfold" the content of the notion of natural number. None of this extensions is, in my opinion, grounded in our knowledge and understanding of natural numbers but rather "reflect" on the properties of certain methods of proof that have been adopted. That is, although these methods of proof operate on arithmetical and logical resources, they also possess certain properties confered to them by certain philosophical positions which are constitutive of their definitions. The extension axioms or schemata are about these properties (e.g. soundness, truth, consistency) and not about the content of the notion of natural number. Gödelian arithmetical statements as well as their analogues in set theory contain explicit references to these methods of proof, as a consequence of which they inherit an additional content which is not purely arithmetical, or set-theoretical, for that matter. One can find a partial recognition of this point in [Horsten, 2011]:

> Gödelian proofs of G_{ZFC} and Con_{ZFC} are certainly *partly* mathematical in nature. The proof cited above, for example, involves an instance of the principle of mathematical induction, which is a mathematical principle if there ever was one. It is just that such Gödelian proofs are not *purely* mathematical proofs. For they essentially contain the notion of *truth*, which is itself not a mathematical but a philosophical notion. This is not to deny that mathematics can be applied to produce interesting theories of truth. It is just that mathematical theories of truth do, on this view, belong not to pure mathematics but at least to *applied* mathematics, or to the more mathematical part of philosophy [Horsten, 2011].

Horsten refers here to the philosophical notion of truth, and to Gödelian proofs using a truth-predicate, but my main point in this paper is slightly different. It concerns the notions of proof and provability. It is a metamathematical notion which reflects a certain finitistic, philosophical standpoint. By making explicit reference to such notions, Gödelian sentences acquire also a higher-order, not purely numerical content, which depends on the properties of these notions and cannot be reduced to the concept of natural number. One possible way to be more explicit about the higher-order content of Gödelian sentences is through some remarks made by Isaacson [Isaacson, 1991; Isaacson, 1996]. He contrasts arithmetical sentences provable in PA with the Gödelian sentences: the former have a pure arithmetical content, and the system PA which proves them arises out of our undertanding of natural numbers. On the other side, the meaning of Gödelian statements involve our reflections on our understanding of natural numbers.

The ideas discussed in this paper have been debated many times in the post Gödelian era. The contribution of the paper is simply one of emphasis. Myhill, for instance expresses similar ideas in an often quoted passage:

> Indeed it seems to me that the use of the word 'proof' in ordinary non-philosophical mathematical discussion is rather clearly neither a syntactical nor a semantical term. It is as self-contradictory to use methods of proof without admitting their correctness, as it is to make statements without admitting their truth. (I am not using 'self-contradictory' in the sense of formal logic, but roughly as a synonym for 'irrational'.) Therefore if a person who has been using certain methods for proving arithmetical theorems succeeds in making these methods explicit, he is ipso facto committed to the perfectly definite proposition that the use of those methods cannot lead to a false arithmetical statement, for example the statement

> that 0 is equal to 1. By Gödel's technique of arithmetization, which translates every statement of formal deducibility into a statement of arithmetic, any such person is compelled to admit a new arithmetical statement, namely the arithmetized version of the statement that his methods cannot lead to a proof of the statement that 0 is equal to 1. By Gödel's theorem, he could not have established this statement by his previous methods. Hence, as soon as a person makes explicit the tools which he has been using in the construction of arithmetical proofs, he is ipso facto in a position to obtain new arithmetical proofs which he could not have obtained by using those tools alone. The whole process is closely related to what the British philosophical logician W.E. Johnson called 'intuitive induction'; we find ourselves making certain inferences and we thereupon realize that the pattern of those inferences is such as to confer validity on arguments in which they occur. This realization is a demonstrative and rational step quite apart from any question of formalization, though of course the *results* of an intuitive induction can be formalized after the induction has taken place [Myhill, 1960, p. 461].

It is difficult to disagree with these remarks. Myhill, like other commentators I discussed (Horsten) is concerned with the distinction between different kind of proofs. My concern in this paper was, however, with the other side of the coin: the meaning of the Gödelian sentences which are settled by these proofs. The minor point I tried to make was that, by making reference to notions like proof (provability), these sentences have a content which transcend the arithmetical content of purely numerical statements. This is the internal, conceptual reason for which, in some cases (not all; there are Gödelian statements like "I am provable" which are provable), their proof has to mobilize higher-order (meta-theoretical) resources, be they in the form of a truth-theory, a la Tarski, or reflection principles, a la Feferman. I think that Gödel was aware of this fact when he made a distinction between intrinsic extensions with Gödelian sentences and intrinsic extensions with other kind of axioms which unfold the content of the basic notions like natural numbers.

References

Dummett, 1963 – Dummett, M. "The Philosophical Significance of Gödel's Theorem", *Ratio*, Vol. 5, pp. 140–155. Reprinted in: Dummett, M. *Truth and Other Enigmas*, London, Duckworth, 1979, pp. 186–201.

Feferman, 1962 – Feferman, S. "Transfinite recursive progressions of axiomatic theories", *The Journal of Symbolic Logic*, 1962, Vol. 27, pp. 259–316.

Feferman, 1991 – Feferman, S. "Reflections on incompleteness", *Journal of Symbolic Logic*, 1991, Vol. 56, pp. 1–49.

Feferman, 2005 – Feferman, S. "Predicativity", in: *The Oxford handbook of philosophy of mathematics and logic*, ed. by S. Shapiro, Oxford: Oxford University Press, 2005, pp. 590–624.

Gödel, 1947 – Gödel, K. "What is Cantor's continuum problem?", *The American Mathematical Monthly*, 1947, Vol. 54, pp. 515–525.

Horsten, 2011 – Horsten, L. *The Tarskian Turn. Deflationism and Axiomatic Truth*, MIT Press, 2011.

Hyttinen, Sandu, 2004 – Hyttinen, T., Sandu, G. "Deflationism and Arithmetical Truth", *Dialectica*, 2004, Vol. 58, pp. 413–426.

Isaacson, 1991 – Isaacson, D. "Some considerations on arithmetical truth and the omega-rule", in: *Proof, logic, and formalization*, ed. by M. Detlefsen, Routledge, 1991, pp. 49–138.

Isaacson, 1996 – Isaacson, D. "Arithmetical truth and hidden higher-order concepts", in: *Logic Colloquium '85*, Amsterdam: North-Holland, 1987, pp. 147–169. Reprinted in: *The philosophy of mathematics*, ed. by W.D. Hart, Oxford University Press, 1996, pp. 203–224.

Ketland, 1999 – Ketland, J. "Tarski's Paradise and Deflationist Truth", *Mind*, 1999, Vol. 108, pp. 69–94.

Koellner, 2006 – Koellner, P. "On the question of absolute undecidability", *Philosophia Mathematica*, 2006, Vol. 14, pp. 153–188. Revised and reprinted in: *Kurt Gödel: Essays for his Centennial*, S. Feferman, C. Parsons, S.G. Simpson (eds.), Lecture Notes in Logic, Vol. 33. Association of Symbolic Logic, 2009.

Myhill, 1960 – Myhill, J. "Some Remarks om the notion of proof", *Journal of Philosophy*, 1960, Vol. 57, pp. 461–471.

Shapiro, 1998 – Shapiro, S. "Truth and Proof: Through Thick and Thin", *Journal of Philosophy*, 1998, Vol. 95, pp. 493–521.

Tarski, 1956 – Tarski, A. "The Concept of Truth in Formalized Languages", in: *Logic, Semantics, Metamathematics*, ed. by A. Tarski, 2d edition, Oxford University Press, 1956, pp. 152–278.

Tennant, 2002 – Tennant, N. "Deflationism and the Gödel Phenomena", *Mind*, 2002, Vol. 111, pp. 551–582.

Логические исследования
2020. Т. 26. № 1. С. 78–90
УДК 510.64

Logical Investigations
2020, Vol. 26, No. 1, pp. 78–90
DOI: 10.21146/2074-1472-2020-26-1-78-90

Неклассическая логика
Non-classical Logic

Jean-Yves Béziau

Many-valuedness from a universal logic perspective

Jean-Yves Béziau
Federal University of Rio de Janeiro,
Largo São Francisco de Paula 1 - Centro, Rio de Janeiro, RJ, 20051-070, Brazil.
Brazilian Academy of Philosophy,
R. Riachuelo, 303 - Centro, Rio de Janeiro - RJ, 20230-011, Brazil.
E-mail: jyb@ufrj.br

Abstract: We start by presenting various ways to define and to talk about many-valued logic(s). We make the distinction between on the one hand the class of many-valued logics and on the other hand what we call "many-valuedness": the meta-theory of many-valued logics and the related meta-theoretical framework that is useful for the study of any logical systems. We point out that *universal logic*, considered as a general theory of logical systems, can be seen as an extension of many-valuedness. After a short story of many-valuedness, stressing that it is present since the beginning of the history of logic in Ancient Greece, we discuss the distinction between dichotomy and polytomy and the possible reduction to bivalence. We then examine the relations between singularity and universality and the connection of many-valuedness with the universe of logical systems. In particular, we have a look at the interrelationship between modal logic, 3-valued logic and paraconsistent logic. We go on by dealing with philosophical aspects and discussing the applications of many-valuedness. We end with some personal recollections regarding Alexander Karpenko, from our first meeting in Ghent, Belgium in 1997, up to our last meeting in Saint Petersburg, Russia in 2016.

Keywords: many-valued logic, many-valuedness, universal logic, modal logic, 3-valued logic, paraconsistent logic

For citation: Béziau J.-Y. "Many-valuedness from a universal logic perspective", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 78–90. DOI: 10.21146/2074-1472-2020-26-1-78-90

En toute sobriété
j'ai eu de mutliples visions de multiples choses
et si je suis arrivé à maintenir ma sérénité
c'est en voyant l'unité au-dela de la diversité
la cohérence au-delà de l'incohérence.
Baron de Chambourcy

1. Many-Valuedness and Universal Logic

The aim of this paper is to develop a better understanding of what many-valuedness is and what universal logic is. Universal logic has emerged as a general theory of logical systems (see [Béziau, 1994] and [Béziau, 2012c]), so it is directly linked to many-valuedness in two different ways:

- Many-valued logics are objects of study of universal logic.
- Many-valuedness, including in particular many-valued logical matrices, is a tool for developing universal logic.

But, as we have pointed out in previous papers (see [Béziau, 2006b] and [Béziau, 2018b]), universal logic is not restricted to a mathematical meta-theory, it encompasses also philosophical and methodological questions. Many-valuedness with its twofold relation with universal logic is a good opportunity to discuss the many virtues of both many-valuedness and universal logic.

2. Many-Valued Logic(s), Many-Valuedness and Universal Logic

As for many concepts, such as e.g. *human being*, *number* or *time*, there is not only one and true definition of *many-valued logic*.

First let us make a distinction between "Many-valued logic" and "Many-valued logics". Here we are putting quotes because we are talking about the linguistic expressions rather than the notions, for which we, as above, use italics.[1] Although it has become trendy, following the fashion of pluralism, to put an "s" at the end of everything, a small snake tailing any idea, let us emphasize that we can still sanely and safely make the distinction between plurality and singularity, not to say unity. No doubts that there are many girls, cars, numbers, but we still can, even without being a Platonist, consider the notions of *girl*, *car* and *number*.

There are many different many-valued logics, but nevertheless we can consider the notion of *many-valued logic* which encompasses all these logics. Although it is rather trivial, it is worth emphasizing that the notion of *many-valued logic* is not itself a logic, in the same way that the notion of *girl* is not

[1] About the notion of notion, see [Béziau, 2018a].

itself a girl, by contrast to the notion of *notion* which can itself be considered as a notion.

To conceptualize what a girl is, we need a general theory of psychology, zoology, archaeology... Of course we can also give a first idea, as stressed by Quine (cf. [Quine, 1960]), just by ostentation, pointing at and/or focusing on a canonical example, such as the Girl from Ipanema. We can indeed do the same with many-valued logic, although less beautifully and musically (see Figure 1).

NOT(A)

A	¬A
F	T
U	U
T	F

AND(A, B)

A ∧ B		B: F	B: U	B: T
A	F	F	F	F
	U	F	U	U
	T	F	U	T

OR(A, B)

A ∨ B		B: F	B: U	B: T
A	F	F	U	T
	U	U	U	T
	T	T	T	T

Fig. 1. Definition by Ostentation of the notion of *many-valued logic*.

That gives a rough idea of what it is. It is necessarily biased, as any tentative to think the general through the particular. But it is fair enough for childish games. If we want to get more scientific, that is another kettle of fish. And if we want to get more philosophical, that is a true cassoulet, not to say feijoada. Let us present three definitions of *A MANY-VALUED LOGIC* on the basis on which we can go a step further than ostentation:

- A logic which does not reduce to truth and falsity.
- A logic that can be characterized only by a logical matrix of more than 2-values (including or not infinite matrices).[2]
- A logic that can be characterized by any semantics with more than 2 truth-values.[3]

A careful look at these definitions shows that they are pairwise different but not pairwise exclusive. In particular the first does not use the notion of *value*, the third one uses it but does not use the notion of *matrix*, by contrast with the second one.

MANY-VALUED LOGIC itself can be considered as

- The class of many-valued logics.
- The meta-theory of many-valued logics.

[2] For details see [Béziau, 1997].

[3] In this case classical proposotional logic can be considered as a many-valued logic, for details see [Béziau, 1997].

- A meta-theoretical tool / framework that is useful for the study of any logical systems.

It is often difficult, not to say artificial, to distinguish between the two first meanings, as in the case of other logics, that is why it is not necessarily useful to introduce two different names. It is also difficult to find a good name for the third meaning. But the expression "MANY-VALUEDNESS" looks pretty good to encompass the 2nd and 3rd above meanings.

The idea of *universal logic* is to promote a general theory inside which / with which, we can turn conceptualization of logical notions and systems easier. In this sense *universal logic* is neither a logic, nor a bunch of logics. It is a kind of extension of many-valuedness as just characterized above: it is the meta-theory of all logical systems, therefore extending the above 2nd meaning and it is a meta-theoretical framework including the above 3rd meaning.

3. A Short Short History of Many-Valuedness

We can say without much exaggeration that many-valuedness exists since the beginning of the world, or better the beginning of the logical world, considering that it is directly connected with Aristotle who is considered himself as the father of logic, as the science of reasoning (Aristotle did not create logic as reasoning, cf. [Béziau, 2010]). This is the famous story of future contingents according to which "Tomorrow will be the end of the world" is neither true, nor false, unless we believe in determinism or apocalypse.

And also without romancising too much we can say that the next step in the story is with Jan Łukasiewicz who, directly influenced by the Stagirite, built a three-valued logic [Łukasiewicz, 1920]. But Łukasiewicz did much more than that, not only he also built a four-valued logic [Łukasiewicz, 1953], but he developed with other Polish logicians, in particular with Alfred Tarski [Łukasiewicz, Tarski, 1930], a whole theory of many-valued logical matrices that is a basic framework and tool for a general theory of logics. At this level we see therefore a strong connection between universal logic and many-valuedness.

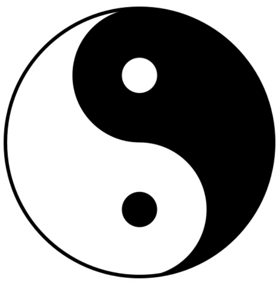
This connection was independently promoted by Paul Bernays and Emil Post at approximately the same time. And it is also worth mentioning that many-valued logic did not escape to Charles Sanders Peirce who had a very general view of logic both from a philosophical and mathematical point of views. In particular he was the first to draw three-valued "truth-tables" (see [Béziau, 2012a]).

This is of course a very short and synoptic story of many-valuedness. We will not go in more details since our objective here is more to look a the present and the future than the past. But the moral of the story is that many-valuedness is present along the whole story of logic and that it is not a crucial difference between traditional logic and modern logic.

4. Dichotomy and Polytomy

Before examining the opposition between many-valuedness and bivalent logic closer and if we can reduce or not logicality to bivalence, let us go beyond logic *stricto sensu* and broaden our horizon to general thinking / conceptualization.

Dichotomy can be found both in the East and the West. In the East we have Taoism, with Yin and her brother Yang, in the West Pythagoras with his table of opposites. Taoism is more radical and interactive: there are only two things from which everything is derived by combination. Pythagoras's table of opposites has at least ten different dichotomic oppositions. But this theory of multiple oppositions was then developed in a very abstract theory of dichotomic oppositions, more abstract than the Chinese one, very logical, connected to the emergence of classical negation, a very powerful tool that can apply to any thing, as we have recently argued in [Béziau, 2019].



One	Many
Right	Left
Male	Female
Square	Oblong
Light	Darkness
Good	Bad
Straight	Crooked
Rest	Motion
Close	Open
Happy	Sad
Odd	Even

Fig. 2. Taosim vs. Pythagoricism.

Polytomy can also be seen both in the East and the West. By its own nature polytomy is multiple: it can be 3, 4, 5, up to infinity. But what is predominating is small size polytomy: trichotomy, quadritomy, or pentagony. At the religious level we have in India the trimurti with Brahma, Vishnu, Shiva and in the West the trinity of Christianity with The Son, The Father and the Holy Spirit. At a more physical level, we have the theory of four elements in the Occident and the theory of five elements in China.

Vivaldi naturally promoted quadritomy with his masterpiece *The Four Seasons*. Schopenhauer was also found of a fourfold approach, that he developed at different levels (see [Béziau, 2020]). On the other hand Peirce was very found of trichotomy as well as Robert Blanché, who duplicated it as a colorful hexatomy (see [Blanché, 1966] and [Béziau, 2012d]). One may wonder in which sense this symphony of polytomies is part of many-valuedness. In the case of Blanché is hexagon is clearly part of it, if we consider that it can be applied to metalogic (see [Béziau, 2013]).

5. The Value of Reduction to Bivalence

If we put aside monotheism, the most impressive reduction of multiplicity is the binary notation. "In the beginning was the word" (John 1:1) can be coded as a sequence of 0s and 1s:

01001001 01101110 00100000 01110100 01101000 01100101 00100000 01100010 01100101 01100111 01101001 01101110 01101110 01101001 01101110 01100111 00100000 01110111 01100001 01110011 00100000 01110100 01101000 01100101 00100000 01110111 01101111 01110010 01100100

It is less poetic and maybe the meaning of the sentence is lost in some way, but it makes sense for a computer. However we are not (yet) computers and what is good for our understanding is something not tooooooooooooo big, but also not too small. For numbers we use decimal notations, an alphabet has an average of 25 signs and the average of phonemes in a language is 31.

Reduction in logic can be considered either from a pragmatic viewpoint or an objective viewpoint. A pseudo-Fregean may claim that there are only two real truth-values: truth and falsity, a pseudo-Peircean may say that three values are quite useful. Peirce proved that all binary connectives can be reduced to only one, but it was not for him a reason to use only one.

A three-valued logic like Łukasiewicz logic L3 cannot be defined by a two-valued truth-functional semantics, however it can be defined by a two-valued non-truth functional semantics, the charateristic functions of relatively maximal theories, like many logics. This result can be considered as a typical result of universal logic (see [Béziau, 2012b]). It is a valuable and interesting result but nevertheless something is lost in the reduction, i.e. truth-functionality.

What we can say, against Suszko's reductionist thesis (see [da Costa et al., 1996]), is that truth-functional semantics helps to give meaning. However we have to be careful with meaning! In standard many-valued matrix semantics the values are divided in two sets: designated and non-designated values. It makes sense to still apply the dichotomy truth/falsity to them. For example in the case of a four-valued matrix semantics with two designated and two non-designated elements, we can use the terminology: strong truth, weak truth, weak falsity, strong falsity, or necessary truth, possible truth, possible falsity, necessary falsity (see [Béziau, 2011]).

Moreover many-valued matrices can be used to refine the notion of consequence relation, as it has been done by G. Malinowski [Malinowski, 1990], Shramko and Wansing [Shramko, Wansing, 2011].

6. Singularity vs. Universality

Let us consider the following truth-table:

$\multimap$	1	2	3	4	5	6	7
1	3	7	2	1	4	4	6
2	1	7	4	2	3	3	1
3	5	2	3	4	5	2	4
4	3	2	5	4	3	6	1
5	2	6	7	3	1	3	1
6	4	1	4	1	5	7	3
7	7	3	2	6	2	6	5

Fig. 3. A Binary Connective in a 7-valued logic.

This is intended to be a truth-table for a binary connective, $\multimap$, in a 7-valued logic. This connective is very singular, peculiar not to say idiosyncratic. What can we say about it? What can we do with it? And why should we waste our time focusing on it?

We can ask the same questions about any particular individual, whether it is a stone, a tree, a number or a human being. Can we say that the number 5987 has an interest by itself? Maybe yes, maybe not. Some particular numbers are more interesting than others, like the number π, to give a classical example of celebrity. And some particular connectives are more interesting than other ones like Sheffer stroke, in bivalent classical logic.

The value of a singular connective in many-valued logic really makes sense only from a general perspective and this is true for any singular object of any field. A singular object is singular only in relation with other objects. A universal approach helps to stress singularity. However a particular case can be a starting point.

It is worth to find some general positive and negative results about all finite valued logics. This is very important. For example if we consider Dugundji theorem stating that S5 cannot be characterized by a finite valued logic (cf. [Dugundji, 1940]), then we will not lose our time looking for a possible 256-valued matrix semantics for it. On the other hand one may explore some particular cases in view of a specific goal or based on an intuitive interpretation. One may develop a beautiful 9-valued logic with wonderful applications.

Another methodology is to connect these general investigations with other mathematical properties and theories. This is what Karpenko did making a connection between prime numbers and many-valued matrix semantics (see [Karpenko, 2006]).

7. Many-Valuedness and the Universe of Logical Systems

If we consider many-valuedness as a general tool, in particular logical matrices, it is related to many logics, including bivalent classical logic, in the sense that it can be applied to them at the meta-level, for example for proof of independence of axioms, as Bernays originally did (see Chapter 2 of [Béziau, 2012c]).

Now if we consider many-valuedness as a tool for constructing logical systems, it is related to many other non-classical logics, in particular modal logic, paraconsistent logic, probability logic, fuzzy logic.

Many-valued logic was developed by Łukasiewciz in view of modality, this line of research was in some sense aborted on the one hand due to the negative result of Dugundji [Dugundji, 1940], on the other hand due to the success of possible world semantics. Nevertheless it still makes sense to use many-valuedness to develop modal logic, either using logical matrices or non truth-functional many-valued semantics. In both cases the problem is with self-extensionality, i.e. the failure of the replacement theorem, but this is not necessarily a problem despite the fact that paradoxically modal logic has been qualified as extensional logic.

Three-valued logic has been used for the developement of paraconsistent logic by Asenjo [Asenjo, 1966], da Costa and D'Ottaviano [D'Ottaviano, da Costa, 1970], Priest [Priest, 1979], Avron [Avron et al., 2018] and Beziau [Béziau, Franceschetto, 2016; Béziau, 2016b].

Asenjo's logic is a logic which is both paraconsistent and 3-valued but not modal, the paraconsistent logic Z [Béziau, 2006a] is both modal and paraconsistent but not finite-valued and we have investigated logics which are at the same time paraconsistent, modal and many-valued but considering 4-values instead of 3-values (see [Béziau, 2011]).

8. Philosophy of Many-Valuedness

In the last 100 years there was a proliferation of logical systems, due in particular to the formalization and mathematization of logic. This is the door open to infinite non-sense. Quine wrote about modern many-valued logic: "Primarily the motivation of these studies has been abstractly mathematical: the pursuit of analogy and generalization. Studied in this spirit, many-valued logic is logic only analogically speaking; it is uninterpreted theory, abstract algebra" [Quine, 1960].

Mathematics is nice and can lead us to the sky of ideas, but it is good to always try to have some meaningful constructions, which can help us to land back down to earth. And it is important to work out the interaction between philosophy and mathematics.

If we say that a proposition is something which is either true or false, formulas of many-valued logic are not propositions. Is this a problem? And then what are they? By contrast we may want to introduce a three-value logic exactly because we believe that we should consider formulas that are neither true nor false as corresponding to propositions.

What happens is that in modern logic people are considering "formulas" in an informal way without asking what they are or/and what they are representing. This is not necessarily a problem, this is the path to abstraction and generalization. But on the one hand it is good to go at a higher level of abstraction, on the other hand to go down to earth to connect to reality.

The idea of universal logic is indeed to consider a structure where a consequence relation acts on objects whose nature is not further specified. These objects can be events, thoughts, information, etc. They can be interpreted in many ways.

It is important to take in account philosophical motivations to develop a mathematical framework. Matrix semantics can look as a non-intuitive, not to say absurd, construct. One may want to build semantics with:

- Formulas having no truth-values.
- Formulas having as value a set of values, e.g. truth and falsity.

These two cases are comically nicknamed respectively gap and glut. It is true that this can be simulated in matrix semantics but simulation is not strictly speaking the same as reality. And although it can make sense to call many-valued the glut case, it is not clear that this makes sense for the gap case.

Let us also stress that a central problem of the philosophy of many-valued logic is how to interpret the additional values. A straightforward interpretation is degrees of truth and degrees of falsity. But in the simplest case, i.e. three-valued logic this does not necessarily make sense in particular due to dissymmetry. The third value is seen most of the time as something at the middle between truth and falsity. It is often called "undetermined" and funny enough it is indeed quite undetermined: it can be considered as designated or non-designated, as neither true nor false, or as both true and false.

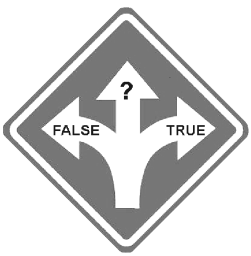


Fig. 4. Indetermination lying at the middle.

9. Applications of Many-Valuedness

We can make a clear distinction between many-valuedness as a meta-theory to study logical systems and many-valuedness as a basis to develop some interesting logical systems that can have useful applications. The first point is quite clear, and we have good examples, the second point is not so obvious, it depends on particular what we really consider as a many-valued logic (cf. our definitions from the first section).

Let us note that there is no many-valued system of logic which really solves the liar paradox (cf. [Béziau, 2016a]). It is also not clear that many-valued logic can be applied in case of physics, in particular to Heisenberg's principle of indeterminacy. Paulette Février [Février, 1937] tried to do so many years ago, but she had very few, not to say no, followers.

Applications to computer science also are not clear, computer scientists in fact use another name *multiple-valued logic* (change of terminology: change of subject?), no to speak about fuzzy logic. Most of the time it is rather something corresponding to many-valuedness or/and algebraic systems rather than to a many-valued logic, excepting the case of the 4-valued logic of Dunn and Belnap (see [Belnap, 1977]), but this logic is also rather a meta-theoretical framework for the theory of computation than an effective system.

10. Dedication and Personal Recollections

I am very glad to dedicate this paper to Alexander Karpenko. I met Alexander for the first time at the 1st World Congress on Paraconsistency which took place in Ghent, Belgium, July 30 – August 2, 1997. At this time our discussion was rather limited because I was not speaking Russian and Alexander was not speaking Swiss. We met again the following year at the Stanislaw Jaśkowski Memorial Symposium July 15-18, 1998 in Torun, Poland.

So our encounter started on a paraconsistent basis. But as shown by the paper presented by Alexander in Torun, entitled "Jaśkowski's criterion and three-valued paraconsistent logics" [Karpenko, 1999], he had an interest for a systematic and universal approach, relating different non-classical logics. At this meeting in Torun I presented a talk in some sense diametrically opposed to his paper, because on the one hand my objective was not to work on Jaśkowski's criterion of maximality, but to find an intuitive basis for Jaśkowski's discussive logic, and on the other hand my solution was not based on many-valued matrices, but on possible world semantics (see [Béziau, 2006a]). I started to work on 3 and 4 matrix semantics and paraconsistent logic only later on (see [Béziau, 2011]).

Our 3rd meeting was at Smirnov's Readings — 3rd International Conference — May 24-27, 2001, in Moscow, Russia, which was my first visit to Russia.

And our further meetings were also all in Russia:

- 4th Smirnov's Reading, May 28-31, 2003, Moscow;
- 6th international conference Logic Today: Developments and Perspectives, June 20-22, 2004, Saint Petersburg;
- 6th Smirnov's Readings, June 17-19, 2009, Moscow;
- Nikolai Vasiliev's Logical Legacy and Modern Logic, October 24-25, 2012, Moscow;
- The 12th international conference Logic Today: Developments and Perspectives, June 22-24, 2016, Saint Petersburg.

We had discussion not limited to logic *stricto sensu*. Alexander had interest for many topics including arts, in particular poetry, and our friendship developed in the framework of this general perspective. Alexander was the Head of the Department of Logic at the Institute of Philosophy, Russian Academy of Science from 2000 until his death. I think it is important in Russia and elsewhere to have researchers and in particular directors of research like Alexander who have a general perspective and vision.

Acknowledgements. Thanks for useful comments by an anonymous referee.

References

Asenjo, 1966 – Asenjo, F.G. "A calculus of antinomies", *Notre Dame Journal of Formal Logic*, 1966, Vol. 7, No. 103, pp. 103–105.

Avron et al., 2018 – Avron, A., Arieli, O., Zamansky, A. *Theory of effective propositional paraconsistent logics*, College Publication, London, 2018.

Belnap, 1977 – Belnap, N.D. "A useful four-valued logic", in: *Modern uses of multiple-valued logic*, ed. by M. Dunn. Reidel, Boston, 1977, pp. 8–37.

Béziau, 1994 – Béziau, J.-Y. "Universal logic", in: *Logica'94 – Proceedings of the 8th International Symposium*, T. Childers and O. Majer (eds.), Prague, 1994, pp. 73–93.

Béziau, 1997 – Béziau, J.-Y. "What is many-valued logic ?", in *Proceedings of the 27th International Symposium on Multiple-Valued Logic*, IEEE Computer Society, Los Alamitos, 1997, pp. 117–121.

Béziau, 2000 – Béziau, J.-Y. "What is paraconsistent logic?", in: *Frontiers of paraconsistent logic*, Research Studies Press, Baldock, 2000, pp. 95–111.

Béziau, 2004 – Béziau, J.-Y. "Non truth-functional many-valued semantics", in *Aspects of Universal Logic*, J.-Y. Béziau, A. Costa-Leite and A. Facchini (eds.), University of Neuchâtel, Neuchâtel, 2004, pp. 199–218.

Béziau, 2006a – Béziau, J.-Y. "The paraconsistent logic Z – A possible solution to Jaśkowski's problem", *Logic and Logical Philosophy*, 2006, Vol. 15, pp. 99–111.

Béziau, 2006b – Béziau, J.-Y. "13 Questions about universal logic", *Bulletin of the Section of Logic*, 2006, Vol. 35, pp. 133–150.

Béziau, 2010 – Béziau, J.-Y. "Logic is not logic", *Abstracta*, 2010, Vol. 6, pp. 73–102.

Béziau, 2011 – Béziau, J.-Y. "A new four-valued approach to modal logic", *Logique et Analyse*, 2011, Vol. 54, pp. 109–121.

Béziau, 2012a – Béziau, J.-Y. "History of truth-values", in: *Handbook of the History of Logic, Vol. 11 – Logic: a history of its central concepts*, D.M. Gabbay, F.J. Pelletier and J. Woods (eds.), Elsevier, Amsterdam, 2012, pp. 233–305.

Béziau, 2012b – Béziau, J.-Y. "Paralogics and the theory of valuation", in: *Universal Logic: an Anthology – From Paul Hertz to Dov Gabbay*, Birkhäuser, Basel, 2012, pp. 361–372.

Béziau, 2012c – *Universal Logic: an Anthology – From Paul Hertz to Dov Gabbay*, ed. by Béziau J.-Y., Birkhäuser, Basel, 2012.

Béziau, 2012d – *Logica Universalis*, Special Double Issue on the Hexagon of Opposition, ed. by Béziau J.-Y., 2012, Vol. 6, Issue 1–2.

Béziau, 2013 – Béziau, J.-Y. "The metalogical hexagon of opposition", *Argumentos*, 2013, Vol. 10, pp. 111–122.

Béziau, 2016a – Béziau, J.-Y. "Trivial dialetheism and the logic of paradox", *Logic and Logical Philosophy*, 2016, Vol. 25, pp. 51–56.

Béziau, 2016b – Béziau, J.-Y. "Two Genuine 3-Valued Paraconsistent Logics", in: *Towards Paraconsistent Engineering*, ed. by S.Akama, Springer International Publishing, Cham, 2016, pp. 35–47.

Béziau, 2018a – Béziau, J.-Y. "The Pyramid of Meaning", in: *A Coat of Many Colours — D60*, J. Ceuppens, H. Smessaert, J. van Craenenbroeck and G. Vanden Wyngaerd (eds.), Brussels, 2018.

Béziau, 2018b – Béziau, J.-Y. "Universal Logic: Evolution of a Project", *Logica Universalis*, 2018, Vol. 12, pp. 1–8.

Béziau, 2019 – Béziau, J.-Y. "Cats that are not cats", in: *Natural Arguments — A Tribute to John Woods*, D. Gabbay, L. Magnani, W. Park, and A.V. Pietarinen (eds.), College Publications, London, 2019, pp. 49–71.

Béziau, 2020 – Béziau, J.-Y. "Metalogic, Schopenhauer and Universal Logic", in: *Mathematics, Logic and Language in Schopenhauer*, ed. by Jens Lemanski, Birkhäuser, Basel, 2020.

Béziau, Franceschetto, 2016 – Béziau, J.-Y., Franceschetto, A. "Strong three-valued paraconsistent logics", in: *New Directions in Paraconsistent Logic*, J.-Y. Béziau, M. Chakraborty and S. Dutta (eds.), Springer, New Delhi, 2016, pp. 131–145.

Blanché, 1966 – Blanché, R. *Structures intellectuelles. Essai sur l'organisation systématique des concepts*, Vrin, Paris, 1966.

da Costa et al., 1996 – da Costa, N.C.A., Béziau, J.-Y., Bueno, O.A.S. "Malinowski and Suszko on many-valued logics: On the reduction of many-valuedness to two-valuedness", *Modern Logic*, 1996, Vol. 6, pp. 272–299.

Cobreros et al., 2014 – Cobreros, P., Egré, P., Ripley D., van Rooij, R. "Foreword: Three-valued logics and their applications", *Journal of Applied Non-Classical Logics*, 2014, Vol. 24, pp. 1–11.

D'Ottaviano, da Costa, 1970 – D'Ottaviano, I.M.L., da Costa, N.C.A. "Sur un probléme de Jaśkowski", *Comptes Rendus de l'Académie des Sciences de Paris*, 1970, Vol. 270, pp. 1349–1353.

Dugundji, 1940 – Dugundji, J. "Note on a property of matrices for Lewis and Langford's calculi of propositions", *The Journal of Symbolic Logic*, 1940, Vol. 5, pp. 150–151.

Epstein, 1993 – Epstein, G. *Multiple-Valued Logic Design*, Institute of Physics Publishing, Bristol, 1993.

Février, 1937 – Février, P. "Les relations d'incertitude d'Heisenberg et la logique", in: *Travaux du IXe Congrès International de Philosophie, Vol. 6*, Hermann, Paris, 1937, pp. 88–94,

Karpenko, 1999 – Karpenko, A. "Jaśkowski's criterion and three-valued paraconsistent logics", *Logc and Logical Philosophy*, 1999, Vol. 7, pp. 81–86.

Karpenko, 2006 – Karpenko, A. *Łukasiewicz Logics and Prime Numbers*, Luniver Press, Beckington, 2006.

Łukasiewicz, 1920 – Łukasiewicz, J. "O logice trj́wartosciowej", *Ruch Filozoficny*, 1920, Vol. 5, pp. 170–171.

Łukasiewicz, 1953 – Łukasiewicz, J. "A system of modal logic", *Journal of Computing Systems*, 1953, Vol. 1, pp. 111–149.

Łukasiewicz, Tarski, 1930 – Łukasiewicz, J., Tarski, A. "Untersuchungen über den Aussagenkalkül", *Comptes Rendus des Séances de la Société des Sciences et des Lettres des Varsovie*, Classe III, 1930, Vol. 23, pp. 30–50.

Malinowski, 1990 – Malinowski, G. "Q-consequence operation", *Reports on mathematical logic*, 1990, Vol. 24, pp. 49–54.

Malinowski, 2007 – Malinowski, G. "Many-valued logic and its philosophy", in: *Handbook of the history of logic, Volume 8*, D.M. Gabbay, J. Woods (eds.), North-Holland, Amsterdam, 2007, pp. 13–94.

Priest, 1979 – Priest, G. "The logic of paradox", *Journal of Philosophical Logic*, 1979, Vol. 8, pp. 219–241.

Quine, 1960 – Quine, W.V.O. *Word and Object*, MIT Press, Cambridge MA, 1960.

Rescher, 1969 – Rescher, N. *Many-Valued Logic*, McGraw Hill, New York, 1969.

Shramko, Wansing, 2011 – Shramko, Y., Wansing, H. *Truth and Falsehood – An Inquiry inot Generalized Logical Values*, Springer, Dordrecht, 2011.

Логические исследования
2020. Т. 26. № 1. С. 91–109
УДК 16+510.64

Logical Investigations
2020, Vol. 26, No. 1, pp. 91–109
DOI: 10.21146/2074-1472-2020-26-1-91-109

Bruno R. Mendonça, Walter A. Carnielli

New semantics for urn logics: taming the enduring scandal of deduction

Bruno R. Mendonça
University of Campinas,
R. Sérgio Buarque de Holanda, 251 - Cidade Universitária, Campinas - SP, 13083-859, Brazil.
E-mail: bruno.ramos.mendonca@gmail.com

Walter A. Carnielli
University of Campinas,
R. Sérgio Buarque de Holanda, 251 - Cidade Universitária, Campinas - SP, 13083-859, Brazil.
E-mail: walter.carnielli@cle.unicamp.br

Abstract: The traditional theory of semantic information, originally proposed by[Bar-Hillel, Carnap, 1953], provides a versatile and pretty plausible conception of what kind of thing semantic information is. It embodies, however, the so-called "scandal of deduction", a thesis according to which logical truths are informationally empty. The scandal of deduction is problematic because it contradicts the fact that ordinary reasoners often do not know whether or not a given sentence is a logical truth. Hence, it is plausible to say, at least from the epistemological standpoint of those reasoners, that such logical sentences are really informative. In order to improve over traditional theory, we can replace its classical metatheory by the so-called *urn logics*, non-standard systems of logic (described in detail below) that better describe the epistemological standpoint of ordinary reasoners. Notwithstanding, the application of such systems to the problem of semantic information faces some challenges: first, we must define truth-conditional semantics for these systems. Secondly, we need to precisely distinguish two systems of urn logic, namely, perfect and imperfect urn logics. Finally, we need to prove characterization theorems for both systems of urn logic. In this paper we offer original (and hopefully, elegant) solutions for all such problems.

Keywords: semantic information, scandal of deduction, logical knowledge, urn logics, truth-conditional semantics, characterization theorems, Hintikka normal forms, non-classical logics

For citation: Mendonça R., Carnielli A. "New semantics for urn logics: taming the enduring scandal of deduction", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 91–109. DOI: 10.21146/2074-1472-2020-26-1-91-109

Dedicated to the memory of Professor Alexander Karpenko

1. Introduction

There is a wide debate in the current logical literature on the nature of semantic information and how it can be measured. Traditional theory of semantic information, originally proposed by [Bar-Hillel, Carnap, 1953], equates the semantic information of a sentence with the number of different models that falsify it. A problematic consequence of this approach is the so-called "scandal of deduction" [Hintikka, 1970a]: according to this thesis, logical truths are informationally empty, due to the absence of falsifying models.

The scandal of deduction seems to be in conflict with the recognizable fact that reasoners often do not know whether or not a given sentence is a logical truth. For instance, beginning students of set theory may find trouble to recognize that Russell's paradox — a problematic consequence of naive set theory — involves a contradiction. In more precise terms, these individuals ignore that the sentence $\neg\exists x\forall y(y \in x \leftrightarrow y \notin y)$ is logically valid. So, it is plausible to assume that, since there are logical truths whose validity is prima facie ignored by ordinary reasoners, at least from the epistemological standpoint of these individuals, such logical sentences are not uninformative.

Mainstream attempts to solve the scandal of deduction generally propose replacing traditional theory by more or less distorted alternatives (as, for instance, in [Floridi, 2004]). However, some of these proposals seem to throw the baby out with the bathwater. Traditional theory offers a flexible and persuasive conception of the kind of information associated with the meaning of sentences, and should be preserved. Fortunately, it is possible to simultaneously save traditional theory and block the scandal of deduction by revising the logical metatheory on which our theory of semantic information is based on.

As mentioned above, the scandal of deduction contradicts the fact that at least some logical truths are really informative when considered from the epistemological standpoint of ordinary reasoners. So, we could block this thesis by adopting some logical theory that better describes the semantic knowledge (i.e., the knowledge of truth-conditions, by assuming a truth-conditional approach to semantics) associated with the semantic competence of ordinary reasoners. This new logical theory should be able to support the idea that, for some logical truths, an ordinary reasoner does not know its truth-conditions; more specifically, it should point out that the epistemological stance of an ordinary reasoner cannot exclude some impossible models that falsify the logical truth in question.

Urn logics, a generalization of first-order logic introduced in [Rantala, 1975], are good candidates to fulfill such an explanatory role. Drawing balls from an urn with or without replacement is a classical problem in probability and statistics. An urn problem is a mental experiment in which colored balls that represent events are withdrawn and whose probabilities should be determined. An urn model is a probability distribution, or a family of probability distributions associated with urn problems.

Now, the intuition behind urn logic is to think of the domain of quantification in a way analogous to urns of balls used in probability theory. The metaphor here is to think of quantifying over objects in a domain as being analogous to drawing balls from an urn. In the standard models of first-order logic the domain of quantification stays fixed as we 'draw' elements from the domain. But we can think of models where the domain of quantification changes in the course of evaluating a formula, as probability urns where balls can be replaced, or new balls can be added.

Basically, the fragments of classical logic introduced by this new procedure lose the capacity to express dependence relations between nested quantifiers in a given formula.[1] Consequently, these systems can formalize the epistemological standpoint of someone who does not know that some first-order valid sentence ϕ is a logical truth because this individual is not aware of the dependence relations holding between nested quantifiers occurring in ϕ.

The application of urn logics to the problem of semantic information, however, faces some technical challenges. First, whereas traditional theory presupposes truth-conditional semantics, urn logics were originally defined in game-theoretic terms. Secondly, there are at least two different systems of urn logic which generate different versions of the traditional theory of semantic information, but the literature does not precisely differentiate such systems (in this paper these systems are called *perfect* and *imperfect* urn logics, for reasons that will become clear soon). Finally, it is not clear exactly which logical truths are semantically informative in the context of urn logics. In order to make this point clear, we need to characterize the set of logical truths of urn logics.

To solve these challenges, we define in this paper a truth-conditional semantics for urn logics that is able to precisely separate their perfect and imperfect versions (section 2). Further, we present a characterization theorem for the set of validities of perfect urn logic, and draw some remarks on how it is possible to obtain a corresponding result for imperfect urn logic (section 4). Such a theorem relies on an important auxiliary result to be presented in section 3, namely, the existence of Hintikka normal forms for urn logics. All of these

[1] Urn semantics should not be confused with *IF-logic* [Tulenheimo, 2018] and related systems. Whilst IF-logic extends first order logic with the capacity of expressing independence / dependence relations between nested quantifiers, urn semantics deals with a fragment of classical logic.

results are original contributions to the study of urn logics — in particular, the characterization theorems introduced in section 4 provide a full generalization of a partial result presented in [Rantala, 1975, pp. 470–472].

2. Truth-conditional semantics for urn logics

Generally speaking, the idea behind urn logics is that two nested quantifiers Q and K occurring in some formula may have different quantificational domains and there are not enough logical resources in these systems to express dependence relations between them. Thus, for instance, in urn logics the quantifiers in $\forall x \exists y(x < y)$ may have different domains and there is no way to force dependence relations between them.

We can formalize this idea in the following way. In what follows, let M be a non-empty set. For any elements $a_0, \ldots, a_n$ of M, let $B(a_0, \ldots, a_n)$ be some non-empty subset of M. We call such sets $B(a_0, \ldots, a_n)$ *choice sets* and there is no specific restrictions on how they are supposed to be concretely defined.

For every $n \in \mathbb{N}$, a *set of choice n-sequences* $\mathfrak{M}_n$ of M is such that the following holds:

1. $\mathfrak{M}_0$ is a set of unary sequences of M;
2. For each $\langle a_0, \ldots, a_n \rangle \in \mathfrak{M}_n$, consider a choice set $B(a_0, \ldots, a_n)$ of M. Then, $\mathfrak{M}_{n+1} = \{\langle a_0, \ldots, a_n, b\rangle : \langle a_0, \ldots, a_n \rangle \in \mathfrak{M}_n$, for every $b \in B(a_0, \ldots, a_n)\}$.

Further, let $\mathfrak{M} = \bigcup\limits_{n \in \mathbb{N}} \mathfrak{M}_n$ be the *set of choice sequences* of M. In order to better visualize the meaning of these concepts, consider the following example: assume M is the set of natural numbers and let $\mathfrak{M}_0 = \{1\}$ and, for any sequence of natural numbers $a_0, \ldots, a_n$, $B(a_0, \ldots, a_n) = \{2a_n\}$. Then $\mathfrak{M}$ collects the initial segments of the geometric progression starting with 1 and with common ratio 2.

Based on those notions, [Rantala, 1975] defines a system of urn logic game-theoretically.

Definition 1. Let M be a classic structure, $\mathfrak{M}$ a set of choice sequences of M and ϕ a formula of the language of M. A urn game semantics G(M, $\mathfrak{M}$, ϕ) is similar to classic game semantics except in the following clauses. Assume that ϕ' is the subformula of ϕ considered in the i-th round of a match p of G(M, $\mathfrak{M}$, ϕ) and the choice k-sequence $\langle a_0, \ldots, a_k \rangle$ has been chosen in the previous rounds:

1. If ϕ' is $\exists x \psi$, then the player who holds ϕ' in the i-th round of p holds $\psi(b/x)$ in the $i+1$-th round of p, for some b of his own choice and such that $\langle a_0, \ldots, a_k, b \rangle \in \mathfrak{M}_{k+1}$;

2. If ϕ' is $\forall x\psi$, then the player who holds ϕ' in the i-th round of p holds $\psi(b/x)$ in the $i+1$-th round of p, for some b of other player's choice and such that $\langle a_0, \ldots, a_k, b\rangle \in \mathfrak{M}_{k+1}$;

The difference between classic and urn game semantics lies essentially in the behavior of quantifiers: whilst in the classic game a player who chooses witnesses in the model might choose them looking in the entire domain, in urn game semantics the player must limit herself to choice sequences only.

Definition 1 describes a game of perfect information, i.e., in order to choose witnesses in the model, the player must verify which witnesses were chosen in the previous rounds of the match. We can define a variant game of imperfect information by adding the condition that, for any $n \in \mathbb{N}$, for any two choice n-sequences $\bar{a}$ and $\bar{b}$, $B(\bar{a}) = B(\bar{b})$. In this case, the player does not need anymore to consider the particular sequence of witnesses that have appeared previously in the game because any such sequence determines the same choice set. To avoid confusion, we call the logical systems which are defined by these two different urn game semantics perfect and imperfect urn logics, respectively.

The greatest obstacle for the construction of a truth-conditional semantics for urn logics is the non-compositional character of these systems [Cresswell, 1982, pp. 128–129]. We can solve this issue by relativizing the satisfaction of formulas in the following way.

Definition 2. Let M be a classic structure with a set of choice sequences $\mathfrak{M}$. For every $n \in \mathbb{N}$ and for any formula ϕ of the language of M, the pair (M, $\mathfrak{M}$) n-satisfies ϕ with respect to $\langle a_0, \ldots, a_{n-1}\rangle$, in symbols $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi$, if the following holds:

- If ϕ is atomic formula, $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi \Leftrightarrow$ M classically satisfies ϕ;
- If ϕ is $\psi \wedge \gamma$, $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi$ if and only if $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \psi$ and $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \gamma$ (an analogous clause holds for disjunction);
- If ϕ is $\neg\psi$, $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi \Leftrightarrow \mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \not\models \psi$;
- If ϕ is $\exists x\psi$, $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi \Leftrightarrow \mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1}, b \models \psi(b/x)$, for some b such that $\langle a_0, \ldots, a_{n-1}, b\rangle \in \mathfrak{M}_n$.
- If ϕ is $\forall x\psi$, $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \phi \Leftrightarrow \mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1}, b \models \psi(b/x)$, for every b such that $\langle a_0, \ldots, a_{n-1}, b\rangle \in \mathfrak{M}_n$.[2]

[2]Strictly speaking, we do not need to present a specific definition of satisfaction for universal formulas since $\forall$ is definable in urn logics as $\neg\,\exists\,\neg$. But, since in the following we will extensively consider universal formulas, this specific clause shows itself to be important.

Finally, (M, $\mathfrak{M}$) satisfies ϕ, in symbols $\mathrm{M}_{\mathfrak{M}} \models \phi$, if and only if (M, $\mathfrak{M}$) 0-satisfies ϕ with respect to $\emptyset$.

In the next theorem, following the usual conventions of the literature on game theory, we refer to the players of an urn game semantics as *Abelard* and *Eloise*. Eloise is the player that initially holds the formula ϕ in a match of the game G(M, $\mathfrak{M}$, ϕ).

Theorem 1. *Let* M *be a classic structure with a set of choice sequences* $\mathfrak{M}$*, and let* ϕ *be a formula of the language of* M*. The following are equivalent:*

1. $\mathrm{M}_{\mathfrak{M}} \models \phi$*;*
2. *Eloise has a winning strategy in* G(M, $\mathfrak{M}, \phi$)*.*

Proof. (1⇒2) Assume 1. We can define a winning strategy σ for Eloise in G(M, $\mathfrak{M}, \phi$) as follows. Suppose that, for some $n \in \mathbb{N}$, we have defined σ in the first i-th rounds of a match p of G(M, $\mathfrak{M}, \phi$). Moreover, consider that the choice k-sequence $\langle a_0, \ldots, a_k \rangle$ has been chosen in those rounds of p:

- Assume Abelard holds $\psi \wedge \gamma$ in the i-th round of p. If $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_k \not\models \psi$, then Eloise demands Abelard to hold ψ in the $i+1$-th round of p; otherwise, Eloise demands Abelard to hold γ;
- Assume Eloise holds $\psi \vee \gamma$ in the i-th round of p. If $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_k \models \psi$, then Eloise demands herself to hold ψ in the $i+1$-th round of p; otherwise, Eloise demands herself to hold γ;
- Assume Abelard holds $\forall x\psi$ in the i-th round of p. If there is some b such that $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_k, b \not\models \psi(b/x)$ such that $\langle a_0, \ldots, a_k, b \rangle \in \mathfrak{M}_{k+1}$, then Eloise demands Abelard to hold $\psi(b/x)$; otherwise, Eloise demands Abelard to hold $\psi(c/x)$, for some arbitrary c such that $\langle a_0, \ldots, a_k, c \rangle \in \mathfrak{M}_{k+1}$;
- Assume Eloise holds $\exists x\psi$ in the i-th round of p. If there is some b such that $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_k, b \models \psi(b/x)$ such that $\langle a_0, \ldots, a_k, b \rangle \in \mathfrak{M}_{k+1}$, then Eloise demands herself to hold $\psi(b/x)$; otherwise, Eloise demands herself to hold $\psi(c/x)$, for some arbitrary c such that $\langle a_0, \ldots, a_k, c \rangle \in \mathfrak{M}_{k+1}$.

By induction on the indexes of the rounds of any match p in which Eloise follows strategy σ, we can verify that p is a winning case for her. Consequently, σ is a winning strategy for Eloise.

(2⇒1) Assuming $\mathrm{M}_{\mathfrak{M}} \not\models \phi$, we can define a winning strategy σ' for Abelard considering a set of directives dual to those defining σ. By a similar argument it is possible to prove that Eloise loses any match in which Abelard follows σ'. Consequently, Eloise has no winning strategy in G(M, $\mathfrak{M}, \phi$). ■

3. Hintikka normal forms for urn logics

The existence of Hintikka normal forms is a fundamental property of classical logic with a variety of applications: Hintikka normal forms characterize tableau-like deductive systems [Hintikka, 1965], generate truth tables for quantificational logic [Freire, 2015] and provide an essential tool for the construction of theories of semantic information [Hintikka, 1970b]. Particularly in the case of urn logics, it is possible to obtain characterization theorems for these systems based on the syntactic structure of special kinds of Hintikka normal forms, as we will show in section 4 below.

Let an *unnested formula* be a formula whose terms have complexity at most 1. For a finite set of formulas $\Psi = \{\psi_0, \ldots, \psi_n\}$, let $\bigwedge \Psi$ denote $\bigwedge\limits_{i \leq n} \psi_i$ and $\bigvee \Psi$ denote $\bigvee\limits_{i \leq n} \psi_i$. Let $\neg\Psi = \{\neg\psi : \psi \in \Psi\}$ and $\exists x \Psi = \{\exists x \psi : \psi \in \Psi\}$. Moreover, let the *quantifier rank* of ϕ, $qr(\phi)$, be as follows:

- For every atomic formula ϕ, $qr(\phi) = 0$;
- If ϕ is $\neg\psi$, then $qr(\phi) = qr(\psi)$;
- If ϕ is either $\psi \wedge \gamma$ or $\psi \vee \gamma$, then $qr(\phi) = \max\{qr(\psi), qr(\gamma)\}$;
- If ϕ is either $\forall x \psi$ or $\exists x \psi$, then $qr(\phi) = qr(\psi) + 1$.

Finally, let $\Phi(x_0, \ldots, x_n)$ be the set of unnested atomic formulas of the considered language with variables within $x_0, \ldots, x_n$.

Definition 3. Let $\mathcal{L}$ be a finite first order language. For every $n, m \in \mathbb{N}$, a state description $\theta_m(x_0, \ldots, x_n)$ of $\mathcal{L}$ is such that:

- $\theta_0(x_0, \ldots, x_n)$ is

  $$\bigwedge \Sigma \wedge \bigwedge \neg[\Phi(x_0, \ldots, x_n) - \Sigma],$$

  for some $\Sigma \subseteq \Phi(x_0, \ldots, x_n)$;
- Now, assume that $m > 0$ and the set Θ of state descriptions of $\mathcal{L}$ with quantifier rank $m-1$ and free-variables $x_0, \ldots, x_n, y_m$ has been defined: the state description $\theta_m(x_0, \ldots, x_n)$ is $(\bigwedge \exists y_m \Gamma) \wedge (\forall y_m \bigvee \Gamma)$, for some set $\Gamma \subseteq \Theta$.

For any formula ϕ of $\mathcal{L}$ with quantifier rank m and free-variables $x_0, \ldots, x_n$, ϕ is a Hintikka normal form if and only if ϕ is $\bigvee \Gamma$, for some set Γ of state descriptions with quantifier rank m and free-variables $x_0, \ldots, x_n$.

In classical logic, we can prove the existence of Hintikka normal forms by showing that satisfiable state descriptions define *partial isomorphisms* [Hodges, 1997, pp. 84–85]. We can generalize this strategy for the case of urn logics.

Definition 4. Let M and N be classic structures with sets of choice sequences $\mathfrak{M}$ and $\mathfrak{N}$, respectively. There is a partial isomorphism $\mathrm{M}_{\mathfrak{M}} \simeq \mathrm{N}_{\mathfrak{N}}$ if and only if there is a poset H of isomorphisms between finite parts of M and N such that the following holds:

1. The first element ι_0 in H is $\emptyset$;
2. Assume that, for some $n \in \mathbb{N}$, there is an isomorphism ι_n between substructures $\{a_0, \ldots, a_{n-1}\} \subseteq \mathrm{M}$ and $\{b_0, \ldots, b_{n-1}\} \subseteq \mathrm{N}$ such that $\iota_n(a_i) = b_i$, $\langle a_0, \ldots, a_{n-1}\rangle \in \mathfrak{M}$ and $\langle b_0, \ldots, b_{n-1}\rangle \in \mathfrak{N}$.
    - For any $a \in \mathrm{M}$ such that $\langle a_0, \ldots, a_{n-1}, a\rangle \in \mathfrak{M}$, there is $b \in \mathrm{N}$ such that $\langle b_0, \ldots, b_{n-1}, b\rangle \in \mathfrak{N}$ and there is an isomorphism ι_{n+1} between $\{a_0, \ldots, a_{n-1}, a\}$ and $\{b_0, \ldots, b_{n-1}, b\}$ such that $\iota_n \subseteq \iota_{n+1} \in H$;
    - For any $b \in \mathrm{N}$ such that $\langle b_0, \ldots, b_{n-1}, b\rangle \in \mathfrak{N}$, there is $a \in \mathrm{M}$ such that $\langle a_0, \ldots, a_{n-1}, a\rangle \in \mathfrak{M}$ and there is an isomorphism ι_{n+1} between $\{a_0, \ldots, a_{n-1}, a\}$ and $\{b_0, \ldots, b_{n-1}, b\}$ such that $\iota_n \subseteq \iota_{n+1} \in H$;
3. H is the smallest poset satisfying conditions 1 and 2.

Moreover, we say that there is a k-partial isomorphism $\mathrm{M}_{\mathfrak{M}} \simeq_k \mathrm{N}_{\mathfrak{N}}$ if and only if condition 2 holds at least for every $n < k$.

Lemma 1. *Let* M *be a classic structure of a finite language and let* $\mathfrak{M}$ *be a set of choice sequences of* M*. Then, for every* $m, n, k \in \mathbb{N}$*, for any elements* $a_0, \ldots, a_n \in \mathrm{M}$*, there is a* unique *state description* $\theta_m(x_0, \ldots, x_n)$ *such that* $\mathrm{M}_{\mathfrak{M}} \models \theta_m(a_0, \ldots, a_n)$.

Proof. Proof by induction on $m \in \mathbb{N}$. In the base case, $\theta_0(x_0, \ldots, x_n)$ is

$$\bigwedge \Sigma \wedge \bigwedge \neg[\Phi(x_0, \ldots, x_n) - \Sigma],$$

for $\Sigma = \{\phi \in \Phi(x_0, \ldots, x_n) : \mathrm{M}_{\mathfrak{M}} \models \phi(a_0, \ldots, a_n)\}$. In the inductive step, consider $\Gamma = \{\theta_m(x_0, \ldots, x_n, y) : \mathrm{M}_{\mathfrak{M}} \models \theta_m(a_0, \ldots, a_n, b),$ for some b such that $\langle a_0, \ldots, a_n, b\rangle \in \mathfrak{M}\}$. Then, $\bigwedge \exists y \Gamma \wedge \forall y \bigvee \Gamma$ is the relevant state description. ■

Lemma 2. *Let* $\mathcal{L}$ *be a finite language,* M *and* N *be classic structures of* $\mathcal{L}$*. Consider* $\mathfrak{M}$ *and* $\mathfrak{N}$ *sets of choice sequences of* M *and* N*, respectively. For any* $k \in \mathbb{N}$*, the following are equivalent:*

1. $\mathrm{M}_{\mathfrak{M}} \simeq_k \mathrm{N}_{\mathfrak{N}}$;

2. *For every* $m \leq k$, *there is a unique state description* θ_m *of* $\mathcal{L}$ *that is satisfied by* $(\mathrm{M}, \mathfrak{M})$ *and* $(\mathrm{N}, \mathfrak{N})$.

Proof. ($1 \Rightarrow 2$) Assume 1. Item 2 is a consequence of the following fact:

($*$) For any m and n such that $m + n \leq k$, for any state description $\theta_m(x_0, \ldots, x_{n-1})$ and for any $\langle a_0, \ldots, a_{n-1}\rangle \in \mathfrak{M}$,

$$\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \theta_m(a_0, \ldots, a_{n-1}) \Leftrightarrow$$

$$\mathrm{N}_{\mathfrak{N}}, \iota_n(a_0), \ldots, \iota_n(a_{n-1}) \models \theta_m(\iota_n(a_0), \ldots, \iota_n(a_{n-1}))$$

for ι_n in the poset H of isomoprhisms that defines $\mathrm{M}_{\mathfrak{M}} \simeq_k \mathrm{N}_{\mathfrak{N}}$.

The proof of ($*$) follows by induction on m. Assume that the property holds for some $m < k$. For any n such that $m + n < k$, consider a state description $\theta_m(x_0, \ldots, x_{n-1}, y)$. We can show that $(\mathrm{M}, \mathfrak{M})$ and $(\mathrm{N}, \mathfrak{N})$ agree in all formulas of the form $\exists y\ \theta_m(x_0, \ldots, x_{n-1}, y)$ in the following way.

Suppose that $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1} \models \exists y\ \theta_m(a_0, \ldots, a_{n-1}, y)$ (the proof in the other direction works in exactly the same way). Therefore,

$$\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1}, a \models \theta_m(a_0, \ldots, a_{n-1}, a)$$

for some a such that $\langle a_0, \ldots, a_{n-1}, a\rangle \in \mathfrak{M}$.

By inductive hypothesis and 1, there is $\iota_{n+1} \in \mathrm{H}$ such that

$$\mathrm{N}_{\mathfrak{N}}, \iota_{n+1}(a_0), \ldots, \iota_{n+1}(a_{n-1}), \iota_{n+1}(a) \models \theta_m(\iota_{n+1}(a_0), \ldots, \iota_{n+1}(a_{n-1}), \iota_{n+1}(a)).$$

So, $\mathrm{N}_{\mathfrak{N}}, \iota_{n+1}(a_0), \ldots, \iota_{n+1}(a_{n-1}) \models \exists y\ \theta_m(\iota_{n+1}(a_0), \ldots, \iota_{n+1}(a_{n-1}), y)$.

Consider the set

$$\Gamma = \{\theta_m(x_0, \ldots, x_{n-1}, y) : \mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{n-1}, a \models \theta_m(a_0, \ldots, a_{n-1}, a),$$

$$\text{for some } a \text{ such that } \langle a_0, \ldots, a_{n-1}, a\rangle \in \mathfrak{M}\}.$$

$(\mathrm{M}, \mathfrak{M})$ and $(\mathrm{N}, \mathfrak{N})$ both satisfy $(\bigwedge \exists y \Gamma) \wedge (\forall y \bigvee \Gamma)$. Moreover, by Lemma 6 this is the only state description that we must consider. This completes the proof of ($*$), 2 being a subcase of it.

($2 \Rightarrow 1$) Assume 2. We can construct $\mathrm{M}_{\mathfrak{M}} \simeq_k \mathrm{N}_{\mathfrak{N}}$ in the following way. For some $m < k$, suppose $\mathrm{M}_{\mathfrak{M}} \simeq_m \mathrm{N}_{\mathfrak{N}}$ has been defined. Let ι_m be one of the isomorphisms which define $\mathrm{M}_{\mathfrak{M}} \simeq_m \mathrm{N}_{\mathfrak{N}}$ and let the domain of ι_m be $\{a_0, \ldots, a_{m-1}\}$ such that $\langle a_0, \ldots, a_{m-1}\rangle \in \mathfrak{M}$.

Without loss of generality, fix some element a such that $\langle a_0, \ldots, a_{m-1}, a\rangle \in \mathfrak{M}$ and consider the state description $\theta_{k-(m+1)}(x_0, \ldots, x_{m-1}, y)$ such that

$$\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{m-1}, a \models \theta_{k-(m+1)}(a_0, \ldots, a_{m-1}, a).$$

By 2 and the definition of $\mathrm{M}_{\mathfrak{M}} \simeq_m \mathrm{N}_{\mathfrak{N}}$,

$$\mathrm{N}_{\mathfrak{N}}, \iota_m(a_0), \ldots, \iota_m(a_{m-1}), b \models \theta_{k-(m+1)}(\iota_m(a_0), \ldots, \iota_m(a_{m-1}), b),$$

for some b such that $\langle \iota_m(a_0), \ldots, \iota_m(a_{m-1}), b\rangle \in \mathfrak{N}$.

Consider the function $\iota^* \supset \iota_m$ with domain $\{a_0, \ldots, a_{m-1}, a\}$ such that $\iota^*(a) = b$. Now, note that for any constant $t \in \mathcal{L}$,

$$t^{\mathrm{M}} = a \Leftrightarrow t = y \text{ occurs in } \theta_{k-(m+1)}(x_0, \ldots, x_{m-1}, y) \Leftrightarrow t^{\mathrm{N}} = b.$$

Moreover, for any n-ary relation $R \in \mathcal{L}$, for every $\bar{a} \in \{a_0, \ldots, a_{m-1}, a\}^n$ and $\bar{x} \in \{x_0, \ldots, x_{m-1}, y\}^n$,

$$\bar{a} \in R^{\mathrm{M}} \Leftrightarrow R(\bar{x}) \text{ occurs in } \theta_{k-(m+1)}(x_0, \ldots, x_{m-1}, y) \Leftrightarrow \iota^*(\bar{a}) \in R^{\mathrm{N}}.$$

The case of functions can be reduced to the case of relations. So, ι^* is an isomorphism. ■

Based on Lemma 2, we can prove that in urn logics every unnested formula has a set of Hintikka normal forms.

Theorem 2. *Let ϕ be an unnested formula with quantifier rank k and free-variables $x_0, \ldots, x_n$. Then, for every $q \geq k$, there is a Hintikka normal form ψ equivalent to ϕ and with quantifier rank q.*

Proof. Proof by induction on ϕ. If ϕ is atomic formula, then ψ is $\bigvee \Gamma$, for

$$\Gamma = \{\theta_q(x_0, \ldots, x_n) : \mathrm{M}_{\mathfrak{M}} \models \theta_q(a_0, \ldots, a_n) \wedge \phi,$$

$$\text{for some M, } \mathfrak{M} \text{ and } a_0, \ldots, a_n \in \mathrm{M}\}.$$

To verify this, first assume that, for some structure M and a set of choice sequences $\mathfrak{M}$, $\mathrm{M}_{\mathfrak{M}} \models \phi(a_0, \ldots, a_n)$. By Lemma 6, there is some state description θ_q such that $\mathrm{M}_{\mathfrak{M}} \models \theta_q(a_0, \ldots, a_n)$. Hence, $\mathrm{M}_{\mathfrak{M}} \models \bigvee \Gamma(a_0, \ldots, a_n)$, by definition of Γ.

On the other hand, assume $\mathrm{M}_{\mathfrak{M}} \not\models \phi(a_0, \ldots, a_n)$. Moreover, suppose that $\mathrm{M}_{\mathfrak{M}} \models \theta_q(a_0, \ldots, a_n)$, for some $\theta_q \in \Gamma$. Then, by definition of Γ, there is some structure N and some set of choice sequences $\mathfrak{N}$ such that $\mathrm{N}_{\mathfrak{N}} \models \theta_q(b_0, \ldots, b_n) \wedge \phi$. However, by Lemma 7, θ_q defines an isomorphism ι;

$\{a_0, \dots, a_n\} \to \{b_0, \dots, b_n\}$ and, consequently, $M_{\mathfrak{M}} \models \phi(a_0, \dots, a_n)$, what contradicts our original hypothesis. So, $\phi \equiv \bigvee \Gamma$.

For the inductive step, fix some $q \geq k$ and assume ϕ is $\psi \wedge \gamma$. By inductive hypothesis, there are Hintikka normal forms $\bigvee \Psi$ and $\bigvee \Gamma$ of ψ and γ, respectively, with quantifier rank q. So, $\bigvee(\Psi \cap \Gamma)$ is the Hintikka normal form of ϕ. A similar reasoning shows that $\bigvee(\Psi \cup \Gamma)$ and $\bigvee(\bar{\Psi})$ are the Hintikka normal forms of $\psi \vee \gamma$ and $\neg\psi$, respectively, for $\bar{\Psi} = \{\theta_q(x_0, \dots, x_n) : \theta_q(x_0, \dots, x_n) \notin \Psi\}$.

Finally, suppose that ϕ is $\exists y \psi$. By inductive hypothesis, ψ has a Hintikka normal form $\bigvee \Gamma$ with quantifier rank $q - 1$. So, $\exists y \psi$ is equivalent to $\exists y \bigvee \Gamma$. Let $\Delta_0, \dots, \Delta_m$ be all the non-empty elements of the power set of Γ. For every $i \leq m$, the formula $\bigwedge \exists y \Delta_i \wedge \forall y \bigvee \Delta_i$ is a state description with quantifier rank q. Further, $\bigvee \Gamma$ is equivalent to $\bigvee\limits_{i \leq m} (\bigwedge \exists y \Delta_i \wedge \forall y \bigvee \Delta_i)$. Therefore, this is a Hintikka normal form of ϕ. ■

Theorem 2 establishes that in urn logics at least unnested formulas have Hintikka normal forms. What about nested formulas? The next result shows that in urn logics it is possible to weakly generalize Theorem 2 for all formulas.

In urn logics, every formula is *equisatisfiable* with some unnested translation of it. We will not prove this fact here (see [Mendonça, 2018, p. 50] for more details), but the following example gives an illustration. For instance, consider a pair $(M, \mathfrak{M})$ satisfying the formula $f(g(x)) = c$ and let $g(x)^{M} = a$ and $f(a)^{M} = b$, for some $a, b \in M$. Let $\mathfrak{M}'$ be a variant of $\mathfrak{M}$ such that $\langle a, b \rangle \in \mathfrak{M}'$. Then, $M_{\mathfrak{M}'} \models \exists y \exists w (g(x) = y \wedge f(y) = w \wedge c = w)$. So, $f(g(x)) = c$ is equisatisfiable with the formula $\exists y \exists w (g(x) = y \wedge f(y) = w \wedge c = w)$.

Corollary 1. Let ϕ be a formula with quantifier rank k. Then, for every $q \geq k$, there is a Hintikka normal form ψ with quantifier rank q equisatisfiable with ϕ.

Proof. Consider a pair $(M, \mathfrak{M})$ that satisfies ϕ. There is a variant $\mathfrak{M}'$ of $\mathfrak{M}$ and an unnested translation ϕ' of ϕ such that $M_{\mathfrak{M}'} \models \phi'$. By Theorem 2, for any $q \geq k$, there is some Hintikka normal form ψ equivalent to ϕ' with quantifier rank q. So, ψ is equisatisfiable with ϕ. ■

4. Characterization theorems

In this section we finally show that a special class of Hintikka normal forms determines a characterization theorem for perfect urn logic. Moreover, we indicate how we can obtain an equivalent result for imperfect urn logic.

In what follows, for any atomic formula ϕ, we say that the *positive literal* ϕ occurs in some state description θ if and only if ϕ occurs in θ and $\neg\phi$ does not occur in θ.

Definition 5. A state description θ is p-consistent if and only if the following holds:

- For any term t that occurs in θ, the positive literal $(t = t)$ occurs in θ;
- For any state description θ' with quantifier rank 0 occurring in θ, for any terms $t_0, \ldots, t_n, s_0, \ldots, s_n$ such that the positive literal $(t_i = s_i)$ occurs in θ', and for every atomic formula ϕ with free-variables within $x_0, \ldots, x_n$, the positive literal $\phi(t_0, \ldots, t_n/x_0, \ldots, x_n)$ occurs in θ' if and only if the positive literal $\phi(s_0, \ldots, s_n/x_0, \ldots, x_n)$ occurs in θ';
- For every atomic formula ϕ whose terms are free in θ, the positive literal ϕ occurs in θ if and only if $\neg\phi$ does not occur in θ.

We say that a Hintikka normal form is p-consistent if and only if every state description occurring in it is p-consistent.

In classical logic, we can prove that consistent formulas are satisfiable by showing that they define *Hintikka sets* [Hodges, 1997, pp. 40–42]. We can explore a similar strategy to show that, in perfect urn logic, all and only p-consistent Hintikka normal forms are satisfiable.

In what follows, let a *subformula chain* be a sequence of formulas $\langle \phi_0, \ldots, \phi_n \rangle$ such that ϕ_j is *subformula* of ϕ_i, for every $0 \leq i < j \leq n$. We rely here on a "relaxed" concept of subformula: in particular, consider that $\neg\phi$ and $\neg\psi$ are subformulas of both $\neg(\phi \wedge \psi)$ and $\neg(\phi \vee \psi)$. Moreover, let $\phi(t/x)$ be a subformula of both $\exists x\phi$ and $\forall x\phi$, and let $\neg\phi(t/x)$ be a subformula of $\neg\exists x\phi$ and $\neg\forall x\phi$, for any term t of the considered language.

For any quantifier Q, for any subformula chain T of the form

$$\langle \phi_0, \ldots, Qx\psi, \psi(t/x), \ldots, \phi_n \rangle,$$

t is called a *witness* of T. The *sequence of witnesses* $\langle t_0, \ldots, t_n \rangle$ of T is the collection of witnesses of T such that t_i occurs first in T than t_j, for every $0 \leq i < j \leq n$. Finally, the quantifier rank of a subformula chain T, in symbols $QR(\mathrm{T})$, is the total number of formulas of the form $Qx\psi$ in T.

Definition 6. For a language $\mathcal{L}$, consider a collection of sets $\{\Delta_n\}_{n \in \mathbb{N}}$ such that the following holds:

1. $(t = t) \in \Delta_0$, for every term t of $\mathcal{L}$ with complexity at most 2;
2. For every atomic formula ϕ, either $\phi \notin \Delta_0$ or $\neg\phi \notin \Delta_0$;

3. For every $n \in \mathbb{N}$, If $\phi \wedge \psi \in \Delta_n$, then $\phi, \psi \in \Delta_n$; If $\neg(\phi \wedge \psi) \in \Delta_n$, then either $\neg\phi$ or $\neg\psi$ are in Δ_n (Dual clauses can be defined for $\phi \vee \psi$ and $\neg(\phi \vee \psi)$);

4. For every $n \in \mathbb{N}$, if $\neg\neg\phi \in \Delta_n$, then $\phi \in \Delta_n$;

5. For every $n \in \mathbb{N}$, for every atomic formula ϕ with free-variables $x_0, \ldots, x_m$, if there are terms $t_0, \ldots, t_m, s_0, \ldots, s_m$ such that $(t_i = s_i) \in \bigcup_{j \leq n} \Delta_j$, then
$$\phi(t_0, \ldots, t_m/x_0, \ldots, x_m) \in \Delta_n \Leftrightarrow \phi(s_0, \ldots, s_m/x_0, \ldots, x_m) \in \Delta_n;$$

6. For every formula ϕ that is either atomic or the negation of an atomic formula, if $\phi \in \Delta_n$, then $\phi \in \Delta_{n+1}$;

7. Assume $\exists x\phi \in \Delta_n$ and $\{\mathrm{T}_0, \ldots, \mathrm{T}_i, \ldots\}$ is the set of all subformula chains defined in $\bigcup_{j \leq n} \Delta_j$ such that $\exists x\phi$ is the last element in every T_i and $QR(\mathrm{T}_i) = n+1$. Then, for each T_i, there is a constant $c_{\phi,i}$ of $\mathcal{L}$ such that $\phi(c_{\phi,i}/x) \in \Delta_{n+1}$ (A dual condition holds for $\neg\forall x\phi$);

8. Assume $\forall x\phi \in \Delta_n$. For every formula ψ of $\mathcal{L}$ and for every non-empty set of subformula chains $\{\mathrm{T}_0, \ldots, \mathrm{T}_i, \ldots\}$ defined in $\bigcup_{j \leq n} \Delta_j$ such that $\exists x\psi \in \Delta_n$ is the last element of each T_i, $QR(\mathrm{T}_i) = n+1$, if there is some T_i with the same sequence of witnesses as that of some subformula chain T defined in $\bigcup_{j \leq n} \Delta_j$ with $\forall x\phi$ as its last element, then $\phi(c_{\psi,i}/x) \in \Delta_{n+1}$, for the constant $c_{\psi,i}$ of $\mathcal{L}$ defined as in clause 7 (A dual condition holds for $\neg\exists x\phi$).

A collection of sets $\bigcup_{j \leq n} \Delta_j$ such that Δ_n has only quantifier-free formulas is a p-set of formulas of $\mathcal{L}$.

The following lemma shows that p-sets are satisfiable in perfect urn logic.

Lemma 3. *For any p-set of sentences $\Delta = \bigcup_{j \leq n} \Delta_j$ of a language $\mathcal{L}$, there is a classic structure* M *and a set of choice sequences $\mathfrak{M}$ such that, for every $j \leq n$, for every $\phi \in \Delta_j$ and for some sequence $\langle a_0, \ldots, a_{j-1} \rangle \in \mathfrak{M}$, it is the case that* $\mathrm{M}_{\mathfrak{M}}, a_0, \ldots, a_{j-1} \models \phi$.

Proof. We will build a canonical model M and a set of choice sequences $\mathfrak{M}$ for Δ. Consider the set C of closed terms of $\mathcal{L}$ and a partition $\Pi(\mathrm{C})$ such that, for every $[\![t]\!] \in \Pi(\mathrm{C})$, $s \in [\![t]\!]$ if and only if $(t = s) \in \Delta$.

Let $^{\mathrm{M}}$ be the following interpretation function:

- For every constant $c \in \mathcal{L}$, $c^{\mathrm{M}} = [\![c]\!]$;
- For every m-ary relation $R \in \mathcal{L}$, $R^{\mathrm{M}} = \{\langle [\![t_0]\!], \ldots, [\![t_{m-1}]\!] \rangle :$ $R(t_0, \ldots, t_{m-1}) \in \Delta\}$ (An analogous condition holds for m-ary functions of $\mathcal{L}$).

Let M= $(\Pi(\mathrm{C}),^{\mathrm{M}})$. $\mathfrak{M}$ can be defined as follows:

- $\mathfrak{M}_0 = \{[\![t]\!] :$ for some formula ϕ of $\mathcal{L}$, $\exists x\phi \in \Delta_0$ and $\phi(t/x) \in \Delta_1\}$;

Now, remember that, by definition of p-set, for every formula of the form $\exists x\phi \in \Delta_j$, the set of subformula chains $\{\mathrm{T}_0, \ldots, \mathrm{T}_i, \ldots\}$ defined in $\bigcup_{k \leq j} \Delta_k$ such that $QR(\mathrm{T}_i) = j+1$ which have $\exists x\phi$ as their last element generates a set of constants $\{c_{\phi,0}, \ldots, c_{\phi,i}, \ldots\}$ such that $\phi(c_{\phi,i}/x) \in \Delta_{j+1}$. Hence:

- For any $[\![\bar{a}]\!] \in \mathfrak{M}_j$, let $B([\![\bar{a}]\!]) = \{[\![c_{\phi,i}]\!] : \bar{a}$ is the sequence of witnesses of some subformula chain T_i defined in $\bigcup_{k \leq j} \Delta_k$ with $\exists x\phi$ as its last element, for some formula ϕ of $\mathcal{L}\}$. Based on this, $\mathfrak{M}_{j+1}$ is straightforwardly defined.

The lemma is a consequence of the following claim:

(**) For every $j \leq n$, for any formula ϕ with free-variables $x_0, \ldots, x_m$ and for any closed terms $t_0, \ldots, t_m$ of $\mathcal{L}$, the following holds:

- If $\phi(t_0, \ldots, t_m/x_0, \ldots, x_m) \in \Delta_j$, then $\mathrm{M}_{\mathfrak{M}}, [\![\bar{s}]\!] \models \phi([\![t_0]\!], \ldots, [\![t_m]\!])$, for every $[\![\bar{s}]\!] \in \mathfrak{M}_{j-1}$ such that $\bar{s}$ is the sequence of witnesses of some subformula chain T defined in $\bigcup_{k \leq j} \Delta_k$ with $\phi(t_0, \ldots, t_m/x_0, \ldots, x_m)$ as its last element;
- If $\neg\phi(t_0, \ldots, t_m/x_0, \ldots, x_m) \in \Delta_j$, then $\mathrm{M}_{\mathfrak{M}}, [\![\bar{s}]\!] \not\models \phi([\![t_0]\!], \ldots, [\![t_m]\!])$, for every $[\![\bar{s}]\!] \in \mathfrak{M}_{j-1}$ such that $\bar{s}$ is the sequence of witnesses of some subformula chain T defined in $\bigcup_{k \leq j} \Delta_k$ with $\neg\phi(t_0, \ldots, t_m/x_0, \ldots, x_m)$ as its last element.

We can prove (**) by induction on ϕ. If ϕ is atomic, then the property follows by definition of M and by clauses 2 and 5 of Definition 6.

Assume ϕ is $\psi \wedge \gamma$ and $\psi \wedge \gamma \in \Delta_j$. Moreover, assume that $\bar{s}$ is the sequence of witnesses of a subformula chain T defined in $\bigcup_{k \leq j} \Delta_k$ with ϕ as its last element. Then, $\psi, \gamma \in \Delta_j$ and both formulas extend T to two other subformula chains that have the same sequence of witnesses. Therefore, by inductive hypothesis,

$M_{\mathfrak{M}}, [\![\bar{s}]\!] \models \psi \wedge \gamma([\![t_0]\!], \dots, [\![t_m]\!])$. On the other hand, assume $\neg(\psi \wedge \gamma) \in \Delta_j$. Then, either $\neg\psi$ or $\neg\gamma$ are in Δ_j. Without loss of generality, suppose $\neg\psi \in \Delta_j$. So, by inductive hypothesis, $M_{\mathfrak{M}}, [\![\bar{s}]\!] \models \neg(\psi \wedge \gamma)([\![t_0]\!], \dots, [\![t_m]\!])$. Based on a similar argument we can verify that the property holds for $\psi \vee \gamma$ and $\neg(\psi \vee \gamma)$.

Assume ϕ is $\neg\psi$. If $\phi \in \Delta_j$, then, by inductive hypothesis,

$$M_{\mathfrak{M}}, [\![\bar{s}]\!] \models \phi([\![t_0]\!], \dots, [\![t_m]\!]).$$

On the other hand, if $\neg\phi \in \Delta_j$, then $\psi \in \Delta_j$ and, by inductive hypothesis, $M_{\mathfrak{M}}, [\![\bar{s}]\!] \not\models \phi([\![t_0]\!], \dots, [\![t_m]\!])$.

Assume ϕ is $\exists x\psi$ and $\phi \in \Delta_j$. Then, $\psi(c_{\psi,i}/x) \in \Delta_{j+1}$. By inductive hypothesis, $M_{\mathfrak{M}}, [\![\bar{s}]\!], [\![c_{\psi,i}]\!] \not\models \psi([\![t_0]\!], \dots, [\![t_m]\!], [\![c_{\psi,i}]\!])$. Given that $[\![c_{\psi,i}]\!] \in B([\![\bar{s}]\!])$, then $M_{\mathfrak{M}}, [\![\bar{s}]\!] \models \exists x\psi([\![t_0]\!], \dots, [\![t_m]\!])$.

Finally, assume ϕ is $\forall x\psi$ and $\phi \in \Delta_j$. Fix some formula γ such that there is some subformula chain T defined in $\bigcup_{k\leq j} \Delta_k$ with $\exists x\gamma$ as its last element and whose sequence of witnesses is $\bar{s}$. Then, $\psi(c_{\gamma,i}/x) \in \Delta_{j+1}$. By inductive hypothesis,

$$M_{\mathfrak{M}}, [\![\bar{s}]\!], [\![c_{\gamma,i}]\!] \models \psi([\![t_0]\!], \dots, [\![t_m]\!], [\![c_{\gamma,i}]\!]).$$

By definition of $B([\![\bar{s}]\!])$, these are all the terms that need to be considered in order to conclude that $M_{\mathfrak{M}}, [\![\bar{s}]\!] \models \forall x\psi([\![t_0]\!], \dots, [\![t_m]\!])$. Based on a similar argument we can verify that the property holds for $\neg\forall x\psi$ and $\neg\exists x\psi$. ■

Lemma 4. *For every p-consistent state description θ with quantifier rank $n+1$ there is a p-set $\Delta = \bigcup_{j\leq n+1} \Delta_j$ such that $\theta \in \Delta_0$.*

Proof. Let $\mathcal{L}$ be a finite language of θ. By definition, θ is of the form

$$(\bigwedge \exists x_0 \Gamma) \wedge (\forall x_0 \bigvee \Gamma),$$

for some set Γ of game normal forms. So, let Δ_0 be the set that collects all the conjuncts of θ plus all identities $t = t$ such that t is a term of complexity at most 2 of $\mathcal{L}$.

In order to define Δ_1, consider an extension $\mathcal{L}_1 = \mathcal{L} \cup \{c_\phi : \phi \in \Gamma\}$. Based on this extension, let

$$\Delta_1 = \{\phi(c_\phi/x_0) : \phi \in \Gamma\} \cup \{\bigvee \Gamma(c_\phi/x_0) : \phi \in \Gamma\} \cup$$

$$\cup \{t = t : t = t \in \Delta_0\} \cup \{\psi : \psi \text{ is either a conjunct of } \phi(c_\phi/x_0) \text{ or}$$

$$\text{is a disjunct of} \bigvee \Gamma(c_\phi/x_0), \text{ for some } \phi \in \Gamma\}.$$

For any $\phi' \in \Gamma$, every $\phi(c_{\phi'}/x_0) \in \Delta_1$ is a game normal form of the form $(\bigwedge \exists x_1 \Gamma_\phi) \wedge (\forall x_1 \bigvee \Gamma_\phi)$. Assume $\phi_0, \dots, \phi_m$ are all such formulas. In order to define Δ_2, consider a new extension $\mathcal{L}_2 = \mathcal{L}_1 \cup \{c_\psi : \psi \in \Gamma_{\phi_i},\ 0 \leq i \leq m\}$.

Let $\Sigma' = \{\psi(c_\psi/x_1) : \psi \in \Gamma_{\phi_i},\ 0 \leq i \leq m\}$. For every $0 \leq i \leq m$, let T_i be the subformula chain defined in $\Delta_0 \cup \Delta_1$ with $\forall x_1 \bigvee \Gamma_{\phi_i}$ as its last element and let $\Sigma'' = \{\bigvee \Gamma_{\phi_i}(c_\psi/x_1)$: the subformula chain T defined in $\Delta_0 \cup \Delta_1$ with $\exists x_1 \psi$ as its last element has the same sequence of witnesses as $\mathrm{T}_i\}$. Based on this, consider that

$$\Delta_2 = \Sigma' \cup \Sigma'' \cup \{t = t : t = t \in \Delta_1\} \cup$$

$$\cup \{\psi : \psi \text{ is either a conjunct of some element of } \Sigma' \text{ or a disjunct of some element of } \Sigma''\}.$$

The reiteration of this process $n+1$-times generates an adequate p-set Δ.■

Theorem 3. *For every state description θ, θ is satisfiable in perfect urn logic if and only if θ is p-consistent.*

Proof. (Sufficiency) Assume θ is not p-consistent. Then, we must consider the following three cases:

- Case 1: for some term t of the considered language, $t \neq t$ occurs in θ. In this case, in every game $\mathrm{G}(\mathrm{M}, \mathfrak{M}, \theta)$ there is a winning strategy for Abelard, namely, to force Eloise to hold $t \neq t$.

- Case 2: in some state description θ' occurring in θ with $qr(\theta') = 0$, for some terms $t_0, \ldots, t_n, s_0, \ldots, s_n$ such that the positive literal $t_i = s_i$ occurs in θ', for some atomic formula ϕ with free-variables $x_0, \ldots, x_n$, the positive literal $\phi(t_0, \ldots, t_n/x_0, \ldots, x_n)$ occurs in θ' but the positive literal $\phi(s_0, \ldots, s_n/x_0, \ldots, x_n)$ does not occur in θ'. In this case, Abelard has the following winning strategy in a game $\mathrm{G}(\mathrm{M}, \mathfrak{M}, \theta)$. Consider that Eloise substitutes $t_0, \ldots, t_n, s_0, \ldots, s_n$ by some $a_0, \ldots, a_n \in \mathrm{M}$ during a match of $\mathrm{G}(\mathrm{M}, \mathfrak{M}, \theta)$. If $\mathrm{M}_\mathfrak{M}, a_0, \ldots, a_n \models \phi$, then Abelard enforces Eloise to hold the formula $\neg\phi(s_0, \ldots, s_n/x_0, \ldots, x_n)$; Otherwise, Abelard enforces Eloise to hold the formula $\phi(t_0, \ldots, t_n/x_0, \ldots, x_n)$.

- Case 3: for some atomic formula ϕ whose terms $t_0, \ldots, t_n$ are not bounded in θ, both ϕ and $\neg\phi$ occur in θ. In this case, Abelard has the following winning strategy in any game $\mathrm{G}(\mathrm{M}, \mathfrak{M}, \theta)$: if $\mathrm{M}_\mathfrak{M} \models \phi$, then Abelard enforces Eloise to hold $\neg\phi$; Otherwise, Abelard enforces Eloise to hold ϕ.

(Necessity) The proof is immediate from Lemmas 3 and 4 . ■

Corollary 2. [Characterization theorem for perfect urn logic] A formula ϕ is satisfiable in perfect urn logic if and only if ϕ has p-consistent Hintikka normal forms.

Imperfect urn logic describes a stricter notion of satisfiability than perfect urn logic. So, finally, let us a present a quick remark on how we can obtain a similar characterization theorem for the former logical system.

Essentially, in imperfect urn logic different choice n-sequences $\bar{a}$ and $\bar{b}$ have always the same choice sets, that is, $B(\bar{a}) = B(\bar{b})$. Now, if the concept of choice sequence is syntactically grasped by the notion of a sequence of witnesses, then what we need to do is to define a subclass of the set of p-consistent state descriptions such that, for any θ in this subclass, any two existential formulas $\exists x\phi$ and $\exists x\psi$ occurring in θ with the same quantifier rank should accept the same witnesses. This subset of formulas defines a notion of consistency for imperfect urn logic (for simplicity, we call this notion *i-consistency*). If we generalize this property for Hintikka normal forms (by saying that for any two state descriptions θ and θ' occurring in some Hintikka normal form, any two existential formulas $\exists x\phi$ and $\exists x\psi$ occurring in them with the same quantifier rank accept the same witnesses), then, appealing to a strategy similar to the one used above, we could show that these are all and the only formulas which are satisfiable in imperfect urn logic. Since the proof of this result is very similar to the one presented above, we let to the reader the task of proving this theorem – the reader can get more information about it in [Mendonça, 2018].

5. How have we avoided the scandal of deduction?

In this paper we have shown that we can define truth-conditional semantics for both perfect and imperfect urn logics. Furthermore, we have presented a full characterization of the sets of logical validities of these systems. In the case of perfect urn logic, every formula satisfiable in this system is equisatisfiable with a p-consistent Hintikka normal form. By its turn, in the case of imperfect urn logic, we have indicated that the satisfiable formulas are all equisatisfiable with i-consistent Hintikka normal forms. Now, we need to go back and ask again: what happens to the scandal of deduction under this new logical framework?

It is not difficult to see that traditional theory of semantic information based on urn logics in fact blocks the scandal of deduction. We can see this by noticing that classical logic defines a proper subclass of structures of urn logic, i.e., we can conceive of a classic structure M as a model whose set of choice sequences $\mathfrak{M}$ is equal to the set of all sequences of M. Now, given the fact that, in this sense, classical logic is a semantic fragment of urn logics, some formulas which are unsatisfiable in the former system become satisfiable in the latter ones. So, many classical logical truths turn out to be not valid in urn logics and, consequently, have its real informativeness acknowledged in this new theoretical context.

However, this is too much theoretical. The reader might want to see some concrete examples of classical logical validities which show themselves to be really informative in this new logical framework. In order to see one such example, we can consider again Russell's paradox, the anomalous consequence of naive set theory. As we have said before, beginning students of set theory often find it difficult to see that the notion of a set of all sets that do not contain themselves as elements is contradictory. Hence, these reasoners ignore that the sentence $\neg\exists x\forall y(y \in x \leftrightarrow y \notin y)$ is logically valid. In urn logics, we can formalize the epistemological standpoint of a person who ignores the logical validity of that sentence in the following way. Consider a classic structure Q and a set of choice sequences $\mathfrak{Q}$ such that $\mathfrak{Q}_0 = \{a\}$, $\mathfrak{Q}_1 = \{b\}$ and Q classically satisfies $b \in a$ and $b \notin b$. It is easy to see that $Q_{\mathfrak{Q}} \models \exists x\forall y(y \in x \leftrightarrow y \notin y)$. So, based on urn logics, the epistemological standpoint of an individual who does not know that the sentence $\neg\exists x\forall y(y \in x \leftrightarrow y \notin y)$ is logically valid can be explained by the idea that her knowledge of truth-conditions does not exclude impossible models such as $Q_{\mathfrak{Q}}$.

Finally, let us make a few more general remarks on what more can we expect to obtain from traditional theory of semantic information in this new logical framework. Since urn logics are decidable formal systems [Olin, 1978, p. 357], with this replacement we obtain a decidable theory of semantic information. This is a theoretical advantage for the following reason: given that here we associated semantic information with the epistemological standpoint of ordinary reasoners, this conception presupposes (at least from an internalist approach to epistemology) that ordinary reasoners are able to effectively measure the semantic information of a given sentence. This presupposition is entirely met by the decidability of urn logics.

However, one question remains unanswered: given that there are at least two different systems of urn logic, which one provides the best framework for the development of a theory of semantic information? This is a difficult question. Perhaps, in order to solve this issue, we should do an empirical research on the actual conditions of the epistemological standpoint of ordinary reasoners towards logical knowledge or, even, of the knowledge associated with the semantic competence of linguistic users. Anyway, this discussion surpasses the limits of the present exposition. Provisionally, we could state that perfect urn logic, given its generality (that is, given the fact that this system semantically includes both classical logic and imperfect urn logic), is a quite appropriate framework for a theory of semantic information.

Acknowledgements. This paper is based on the PhD Thesis of the first author under the supervision of the second [Mendonça, 2018], which received the 2019 CAPES-Brazil award for Best Thesis in Philosophy. The first author is supported by a postdoc scholarship from CAPES Brazil (project 88887.475401/2020-00). The second author has been supported by an individual research grant from CNPq Brazil (grant 307376/2018-4).

References

Bar-Hillel, Carnap, 1953 – Bar-Hillel, Y., Carnap, R. "Semantic information", *The British Journal for the Philosophy of Science*, 1953, Vol. 4, No. 14, pp. 147–157.

Cresswell, 1982 – Cresswell, M.J. "Urn models: A classical exposition", *Studia Logica*, 1982, Vol. 41, No. 2–3, pp. 109–130.

Floridi, 2004 – Floridi, L. "Outline of a theory of strongly semantic information", *Minds and machines*, 2004, Vol. 14, No. 2, pp. 197–221.

Freire, 2015 – Freire, R.A. "First-order logic and first-order functions", *Logica Universalis*, 2015, Vol. 9, No. 3, pp. 281–329.

Hintikka, 1965 – Hintikka, J. "Distributive normal forms in first-order logic", in: *Studies in Logic and the Foundations of Mathematics*, Elsevier, 1965, Vol. 40, pp. 48–91.

Hintikka, 1970a – Hintikka, J. "Information, deduction, and the a priori", *Nous*, 1970, pp. 135–152.

Hintikka, 1970b – Hintikka, J. "Surface information and depth information", in: *Information and inference*, Springer, 1970, pp. 263–297.

Hodges, 1997 – Hodges, W. *A shorter model theory.* Cambridge university press, 1997.

Mendonça, 2018 – Mendonça, B.R. *Traditional theory of semantic information without scandal of deduction*, PhD Thesis, Unicamp, 2018. `http://repositorio.unicamp.br/handle/REPOSIP/332305`.

Olin, 1978 – Olin, P. "Urn models and categoricity", *Journal of Philosophical Logic*, 1978, Vol. 7, No. 1, pp. 331–345.

Rantala, 1975 – Rantala, V. "Urn models: a new kind of non-standard model for first-order logic", in: *Game-Theoretical Semantics*, Springer, 1975, pp. 347–366.

Tulenheimo, 2018 – Tulenheimo, T. "Independence friendly logic", in: *The Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta, Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018.

Логические исследования
2020. Т. 26. № 1. С. 110–123
УДК 510.64

Logical Investigations
2020, Vol. 26, No. 1, pp. 110–123
DOI: 10.21146/2074-1472-2020-26-1-110-123

Janusz Ciuciura

A lattice of the paracomplete calculi

Janusz Ciuciura
University of Łódź,
Faculty of Philosophy and History,
Institute of Philosophy,
Lindleya 3/5, 90–131 Łódź, Poland.
E-mail: `janusz.ciuciura@uni.lodz.pl`

Abstract: Paracomplete logic is intended to cope with the problem of vagueness, or uncertain and incomplete data. It deals with the situation when some propositions and their negations are allowed to be simultaneously false, which is obviously impossible in the classical and many non–classical propositional logics. In paracomplete logic, such classical laws as tertium non datur or consequentia mirabilis are not generally accepted. This implies that the logic is defined negatively.

In this paper, we introduce a family of the paracomplete calculi that will be defined in a Hilbert-style formalization. We propose the so-called bi–valuational semantics and prove the key metatheorems for the calculi. We also discuss a generalization of the paracomplete calculus QD^1 to the hierarchy of related calculi.

Keywords: paracomplete logic, paracompleteness, the law of exluded middle, tertium non datur, consequentia mirabilis, weakly–intuitionistic logic, literal–paracomplete logic, paraconsistent logic

For citation: Ciuciura J. "A lattice of the paracomplete calculi", *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 110–123. DOI: 10.21146/2074-1472-2020-26-1-110-123

1. Introduction

Let *var* denote a (non-empty) denumerable set of all propositional variables. The set of formulas $\mathcal{F}$ is inductively defined as follows:

$$\varphi ::= p \mid \neg\alpha \mid \alpha \vee \alpha \mid \alpha \wedge \alpha \mid \alpha \to \alpha,$$

where $p \in var$, $\alpha \in \mathcal{F}$ and the symbols $\neg$, $\vee$, $\wedge$, $\to$ denote negation, disjunction, conjunction and implication, respectively. The connective of equivalence, $\alpha \leftrightarrow \beta$, is treated as an abbreviation for $(\alpha \to \beta) \wedge (\beta \to \alpha)$.

Paracomplete logic can be defined in various ways, for instance,

Definition 1. A logic $\langle \mathcal{L}, \vdash \rangle$ is said to be paracomplete if, and only if
(1) $\{\beta \to \alpha, \neg\beta \to \alpha\} \nvdash \alpha$, for some $\alpha, \beta \in \mathcal{F}$; or
(2) $\emptyset \nvdash \alpha \vee \neg\alpha$, for some $\alpha \in \mathcal{F}$; or
(3) $\emptyset \nvdash (\neg\alpha \to \alpha) \to \alpha$, for some $\alpha \in \mathcal{F}$; or
(4) $\emptyset \nvdash (\alpha \to \neg\alpha) \to \neg\alpha$, for some $\alpha \in \mathcal{F}$.[1]

It is noticeable that paracomplete logic is specified negatively: any logic is paracomplete if it meets at least one of the criteria listed above. The definitions may seem too general at first sight; in particular, they may suggest some logics which have nothing in common with *paracompleteness*. Suffice it to note that Łukasiewicz's three–valued logic meets the four requirements. It is not by accident, however, that the example has been cited here. From philosophical perspective, paracomplete calculi are expected to cope with the problem of vagueness,[2] or uncertain and incomplete information.[3] Seen from this viewpoint, Łukasiewicz's logic is a good example of how to interpret uncertainty in relation to the issue of determinism or fatalism, whereas paracomplete calculi – with regard to the dynamic character of information or knowledge. Metaphorically speaking, in paracomplete logic, the dilemma of 'Tomorrow's sea fight' has been reduced to 'Today's communication'.

The paracomplete calculi are expected to deal with the situation when some propositions and their negations are allowed to be simultaneously false, which is impossible in the classical and many non–classical propositional logics. The calculi are also viewed as being *dual* to their paraconsistent counterparts, in a sense that "(...) a logic is paraconsistent if it can be the underlying logic of theories containing contradictory theorems which are both true. (...) a logical system is paracomplete if it can function as the underlying logic of theories in which there are (closed) formulas such that these formulas and their negations are simultaneously false" [Loparić, da Costa, 1984, p. 119].

In what follows, we will consider axiomatic propositional calculi in a Hilbert-style formalization with the sole rule of inference (MP): $\alpha \to \beta,\ \alpha\ /\ \beta$. Such a calculus $\mathcal{C}$, identified with the triple $\langle \mathcal{F}, Ax_{\mathcal{C}}, \vdash_{\mathcal{C}} \rangle$, is determined by its set of axioms $Ax_{\mathcal{C}}$ which is included in $\mathcal{F}$. We will require for each paracomplete calculus that it contains all axiom schemas of the positive fragment of Classical Propositional Calculus ($\mathcal{CPC}^+$, for short), that is, all instances of the following schemas:

[1] *Cit. per* [Petrukhin, 2018, pp. 425–426]. Some interesting examples of the paracomplete calculi are given in [Batens et all, 1999; Bolotov et all, 2018; Ciuciura, 2015; Karpenko, Tomova, 2017; Loparić, da Costa, 1984; Popov, 2002; Sette, Carnielli, 1995].

[2] See [Arruda, Alves, 1979; Arruda, Alves, 1979] and [Beall, 2017, Section 4.1], for details.

[3] See [Bolotov et all, 2018], for details.

$(A1)$ $\alpha \to (\beta \to \alpha)$
$(A2)$ $(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$
$(A3)$ $((\alpha \to \beta) \to \alpha) \to \alpha$
$(A4)$ $(\alpha \wedge \beta) \to \alpha$
$(A5)$ $(\alpha \wedge \beta) \to \beta$
$(A6)$ $\alpha \to (\beta \to (\alpha \wedge \beta))$
$(A7)$ $\alpha \to (\alpha \vee \beta)$
$(A8)$ $\beta \to (\alpha \vee \beta)$
$(A9)$ $(\alpha \to \gamma) \to ((\beta \to \gamma) \to (\alpha \vee \beta \to \gamma))$,

admits the rule (MP), and fulfils all the criteria listed in *Definition 1.* To put it more accurately:

Definition 2. A calculus $\langle \mathcal{F}, Ax_{\mathcal{C}}, \vdash_{\mathcal{C}} \rangle$ is said to be paracomplete if, and only if it contains $\mathcal{CPC}^{+}$, admits (MP) and cumulatively meets the conditions:
(1) $\{\beta \to \alpha, \neg\beta \to \alpha\} \nvdash \alpha$, for some $\alpha, \beta \in \mathcal{F}$
(2) $\emptyset \nvdash \alpha \vee \neg\alpha$, for some $\alpha \in \mathcal{F}$
(3) $\emptyset \nvdash (\neg\alpha \to \alpha) \to \alpha$, for some $\alpha \in \mathcal{F}$
(4) $\emptyset \nvdash (\alpha \to \neg\alpha) \to \neg\alpha$, for some $\alpha \in \mathcal{F}$.

Observe that many non-classical logics, esp. Intuitionistic and Łukasiewicz's three–valued logic, do not come within the scope of *paracompleteness*.

Definition 3. For $\mathcal{C}$, any $\alpha \in \mathcal{F}$ and any $\Gamma \subseteq \mathcal{F}$, we say that α *is provable from* Γ *within* $\mathcal{C}$ (in symbols: $\Gamma \vdash_{\mathcal{C}} \alpha$) iff there is a finite sequence of formulas, $\beta_1, \beta_2, \ldots, \beta_n$ such that $\beta_n = \alpha$ and for each $i \leqslant n$, either $\beta_i \in \Gamma$, or $\beta_i \in Ax_{\mathcal{C}}$, or for some $j, k \leqslant i$ we have $\beta_k = \beta_j \to \beta_i$. A formula α is a thesis of $\mathcal{C}$ iff α is provable from $\emptyset$ within $\mathcal{C}$ (in symbols: $\emptyset \vdash_{\mathcal{C}} \alpha$).

Definition 4. Let $\mathcal{T}(\mathcal{C})$ be the set of all theses of $\mathcal{C}$. For any calculi $\mathcal{C}$ and $\mathcal{C}\star$ in $\mathcal{F}$, we say that $\mathcal{C}$ is an extension of $\mathcal{C}\star$ if, and only if $\mathcal{T}(\mathcal{C}\star) \subseteq \mathcal{T}(\mathcal{C})$. We say that $\mathcal{C}\star$ is a *proper subsystem* of $\mathcal{C}$ (in symbols: $\mathcal{C}\star \sqsubset \mathcal{C}$) if, and only if $\mathcal{T}(\mathcal{C}\star) \subseteq \mathcal{T}(\mathcal{C})$ and $\mathcal{T}(\mathcal{C}) \nsubseteq \mathcal{T}(\mathcal{C}\star)$.

Let us recall a few well–known facts about $\mathcal{C}$, where $\mathcal{C} = \mathcal{CPC}^{+} + \text{(MP)}$.

Theorem 1. *Deduction theorem holds for* $\mathcal{C}$*.*

Proof. This follows from the fact that $\mathcal{C}$ includes $(A1)$ and $(A2)$, and the sole rule of inference in $\mathcal{C}$ is (MP). ■

Lemma 1. *Let* $\Gamma, \Delta \subseteq \mathcal{F}$ *and* α, β, $\gamma \in \mathcal{F}$.

(1) If $\alpha \in \Gamma$, *then* $\Gamma \vdash_{\mathcal{C}} \alpha$
(2) If $\Gamma \subseteq \Delta$ *and* $\Gamma \vdash_{\mathcal{C}} \alpha$, *then* $\Delta \vdash_{\mathcal{C}} \alpha$
(3) $\Gamma \vdash_{\mathcal{C}} \alpha$ *iff for some finite* $\Delta \subseteq \Gamma$, $\Delta \vdash_{\mathcal{C}} \alpha$
(4) If $\Delta \vdash_{\mathcal{C}} \alpha$ *and, for every* $\beta \in \Delta$ *it is true that* $\Gamma \vdash_{\mathcal{C}} \beta$, *then* $\Gamma \vdash_{\mathcal{C}} \alpha$
(5) If $\Gamma \cup \{\alpha\} \vdash_{\mathcal{C}} \gamma$ *and* $\Gamma \cup \{\beta\} \vdash_{\mathcal{C}} \gamma$, *then* $\Gamma \cup \{\alpha \vee \beta\} \vdash_{\mathcal{C}} \gamma$
(6) If $\Gamma \cup \{\alpha\} \vdash_{\mathcal{C}} \beta$ *and* $\Delta \vdash_{\mathcal{C}} \alpha$, *then* $\Gamma \cup \Delta \vdash_{\mathcal{C}} \beta$
(in particular, if $\Gamma \cup \{\alpha\} \vdash_{\mathcal{C}} \beta$ *and* $\emptyset \vdash_{\mathcal{C}} \alpha$, *then* $\Gamma \vdash_{\mathcal{C}} \beta$*)*

Proof. We refer the interested reader to [Wójcicki, 1988] and [Pogorzelski, Wojtylak, 2008] for details. ■

Remark 1. The relation $\vdash_{\mathcal{C}}$ is a finitary consequence relation satisfying Tarskian properties (reflexivity, monotonicity, transitivity).

2. Paracomplete calculi. Axioms

The *basic* paracomplete calculus discussed in this section is $CLaN$. $CLaN$, as introduced in [Batens et all, 1999], is defined by (MP), CPC^+ and the law of explosion (DS): $\alpha \to (\neg\alpha \to \beta)$. In the succeeding paragraphs, we consider some extensions of $CLaN$. They are obtained from $CLaN$ by adding to it at least one of the schemas:

(ExM^2) $\alpha \vee \neg\alpha \vee \neg\neg\alpha$
($NN^\star$) $\alpha \to \neg\neg\alpha$.

As a result, we obtain three such extensions, namely,

$D_{min} = CLaN + (NN^\star)$
$Q^1 = CLaN + (ExM^2)$
$QD^1 = CLaN + (ExM^2) + (NN^\star)$.

The calculus Q^1 was introduced in [Ciuciura, 2019]; D_{min} was briefly discussed in [Carnielli, Marcos, 1999]; QD^1 seems to be pretty new. Notice that the calculi (incl. QD^1) are proper subsystems of I^1. The propositional calculus I^1 was originally defined by (MP), ($A1$), ($A2$),

($I1$) $(\neg\neg\alpha \to \neg\beta) \to ((\neg\neg\alpha \to \beta) \to \neg\alpha)$
($I2$) $\neg\neg(\alpha \to \beta) \to (\alpha \to \beta)$.[4]

The connectives of $\neg$ and $\to$ are taken as primitives. Conjunction, disjunction and equivalence are useful abbreviations. They can be introduced *via* the definitions:

[4][Sette, Carnielli, 1995, pp. 182–183].

$\alpha \wedge \beta =_{df} \neg(((\alpha \to \alpha) \to \alpha) \to \neg((\beta \to \beta) \to \beta))$
$\alpha \vee \beta =_{df} (\neg(\beta \to \beta) \to \beta) \to ((\alpha \to \alpha) \to \alpha)$
$\alpha \leftrightarrow \beta =_{df} (\alpha \to \beta) \wedge (\beta \to \alpha)$.[5]

It is noteworthy that I^1 gave an impulse for further research and several alternative axiomatizations for the calculus were proposed. In [*Ciuciura*, 2015], for instance, the *consequentia mirabilis* (*cf. Defition 1, (3)*) plays the key role; in [Fernández, Coniglio, 2003], the role is taken by the *tertium non datur* (*cf. Defition 1, (2)*) which suggests that the connective of disjunction (and conjunction) formally appears in formulas. Indeed, Fernández–Coniglio's axiomatization consists of $(A1)$, $(A2)$, $(A4)$–$(A9)$, $(NN^\star)$ and

(nC) $\neg(\alpha \wedge \neg\alpha)$
$(NI^\star)$ $(\alpha \vee \neg\alpha) \to ((\alpha \to \beta) \to ((\alpha \to \neg\beta) \to \neg\alpha))$
$(ExM^\neg)$ $\neg\alpha \vee \neg\neg\alpha$
$(ExM^\ddagger)$ $(\alpha \ddagger \beta) \vee \neg(\alpha \ddagger \beta)$, where $\ddagger \in \{\wedge, \vee, \to\}$.

The sole rule of inference is (MP).

We prove now that $CLaN, D_{min}, Q^1$ and QD^1 meet the criteria mentioned in *Definition 2*; let $\mathcal{C} \in \{CLaN, D_{min}, Q^1, QD^1\}$, for the sake of brevity.

Remark 2. (1) The formulas
(ExM) $p \vee \neg p$
$(CM1)$ $(p \to \neg p) \to \neg p$
$(CM2)$ $(\neg p \to p) \to p$
(NN) $\neg\neg p \to p$
are not provable in $\mathcal{C}$.

(2) Neither (a) $\{\beta \to \alpha, \neg\beta \to \alpha\} \vdash_{\mathcal{C}} \alpha$, nor (b) $\{\neg(\alpha \to \beta)\} \vdash_{\mathcal{C}} \neg\beta$, nor (c) $\{\alpha \to \neg\beta, \alpha \to \beta\} \vdash_{\mathcal{C}} \neg\alpha$ hold, for any $\alpha, \beta \in \mathcal{F}$.

Proof. Apply the matrix $\mathcal{M}_I = \langle\{1, 2, 0\}, \{1\}, \neg, \wedge, \vee, \to\rangle$, where $\{1, 2, 0\}$ is the set of logical values, 1 is the designated truth value in $\mathcal{M}_I$ and the connectives $\neg, \wedge, \vee, \to$ are defined in the same way as it is done in [Sette, Carnielli, 1995] (see pp. 190, 199), that is,

$\to$	1	2	0
1	1	0	0
2	1	1	1
0	1	1	1

	$\neg$
1	0
2	0
0	1

[5] See *Ibid.*, p. 199.; see also [Karpenko, Tomova, 2017, p. 14].

$\wedge$	1	2	0
1	1	0	0
2	0	0	0
0	0	0	0

$\vee$	1	2	0
1	1	1	1
2	1	0	0
0	1	0	0

Observe that $(A1)$–$(A9)$, (DS), (ExM^2), $(NN^\star)$ are valid in $\mathcal{M}_I$ and (MP) preserves validity. To demonstrate that (ExM), $(CM1)$, $(CM2)$ and (NN) are unprovable in $\mathcal{C}$, it suffices to assign 2 to p in the formulas $p \vee \neg p$, $(p \to \neg p) \to \neg p$, $(\neg p \to p) \to p$ and $\neg\neg p \to p$, respectively. This shows that the claim *(1)* holds. For *(2)*, assign 2 to α and β in *(a)*; 1 to α and 2 to β in *(b)*; 2 to α and 0 to β in *(c)*. ■

Remark 3. The calculus QD^1 can be defined, in a Hilbert-style formalization, by the axiom schemas of CPC^+, (DS), $(ExM^\neg)$ $\neg\alpha \vee \neg\neg\alpha$ and (MP).

Proof. We need to show that *(1)* $(ExM^\neg)$ is a thesis of QD^1, and *(2)* (ExM^2) and $(NN^\star)$ are provable in $QD^1_\star$, where $QD^1_\star$ is defined by CPC^+, (DS), $(ExM^\neg)$ and (MP). *(1)*: This can be easily done by means of (ExM^2), $(NN^\star)$, the thesis of CPC^+ $(\alpha \vee \beta \vee \gamma) \to ((\alpha \to \gamma) \to (\beta \vee \gamma))$ and (MP). *(2)*: Assume that α (by the deduction theorem). Then, we obtain $\neg\alpha \to \neg\neg\alpha$ by (DS), the assumption and (MP). Notice that $\emptyset \vdash_{QD^1_\star} (\neg\alpha \to \neg\neg\alpha) \to \neg\neg\alpha$ by $(ExM^\neg)$, the thesis of CPC^+ $(\alpha \vee \beta) \to ((\alpha \to \beta) \to \beta)$ and (MP). If $\neg\alpha \to \neg\neg\alpha$ and $(\neg\alpha \to \neg\neg\alpha) \to \neg\neg\alpha$, then $\neg\neg\alpha$, and finally $\emptyset \vdash_{QD^1_\star} \alpha \to \neg\neg\alpha$ by the deduction theorem. To prove that (ExM^2) is a thesis of $QD^1_\star$, it suffices to apply (MP) to $(A8)$ and $(ExM^\neg)$. ■

Remark 4. $CLaN \sqsubset Q^1 \sqsubset QD^1$ and $CLaN \sqsubset D_{min} \sqsubset QD^1$.

Proof. It is clear that Q^1 and D_{min} are the extensions of $CLaN$. A proof that $CLaN$ is a proper subsystem of Q^1 immediately follows from the classical truth tables for implication, conjunction and disjunction plus the following one for negation:

	$\neg$
1	0
0	0

The designated value is 1. As expected, $(A1)$–$(A9)$, (DS) are valid under the interpretation and (MP) preserves validity. Now, assign 0 to p in $p \vee \neg p \vee \neg\neg p$ to demonstrate that there is a thesis of Q^1 which is unprovable in $CLaN$.

A proof that $CLaN$ is a proper subsystem of D_{min} basically follows from the fact that $p \to \neg\neg p$ is not provable in $CLaN$. This can be shown by modifying the matrix $\mathcal{M}_I$ appropriately, that is, by replacing the truth table for negation with the so-called rotary negation:

	$\neg$
1	2
2	0
0	1

and assigning 1 to p in $p \to \neg\neg p$. Let $\mathcal{M}_3$ denote the resulting matrix, henceforth.

It is obvious that QD^1 is an extension of Q^1 and D_{min}. Now, we show that the formula $\neg p \vee \neg\neg p$ is provable neither in Q^1 nor D_{min}. Case Q^1: apply $\mathcal{M}_3$ and assign 1 to p in $\neg p \vee \neg\neg p$. Case D_{min}: consider the matrix $\mathcal{M}_{3\star} = \langle \{1, 2, 0\}, \{1\}, \neg, \wedge, \vee, \to \rangle$, where the connectives $\wedge, \vee, \to$ are defined in the same way as in $\mathcal{M}_I$, but the truth table for negation is as follows:

	$\neg$
1	0
2	2
0	1

The axiom schemas of D_{min} are valid in the matrix and (MP) preserves validity; to *falsify* $\neg p \vee \neg\neg p$, it is enough to assign 2 to p. ■

Remark 5. (1) $D_{min} \not\sqsubset Q^1$
(2) $Q^1 \not\sqsubset D_{min}$.

Proof. *(1)*: Apply the matrix $\mathcal{M}_{3\star}$ and assign 2 to p in $p \vee \neg p \vee \neg\neg p$, to show that the formula $p \vee \neg p \vee \neg\neg p$ is unprovable in D_{min}. *(2)*: Use the matrix $\mathcal{M}_3$ and assign 1 to p in $p \to \neg\neg p$, to demonstrate that $p \to \neg\neg p$ is unprovable in Q^1. ■

Remark 6. $QD^1 \sqsubset I^1 \sqsubset CPC$, where CPC denotes the classical propositional calculus.

Proof. It is known that $I^1 \sqsubset CPC$.[6] All we have to do is to prove that $QD^1 \sqsubset I^1$. Since (MP) is the sole rule of inference of both calculi and each axiom schema of QD^1 is provable in I^1, then I^1 is an extension of QD^1. Now, we prove that (nC^p) $\neg(p \wedge \neg p)$ is not a thesis of QD^1 (*cf.* Fernández–Coniglio's axiomatization of I^1). For this purpose, consider the matrix $\mathcal{M}_{3\star\star} = \langle \{1, 2, 0\}, \{1\}, \neg, \wedge, \vee, \to \rangle$, where 1 is the only designated value in $\mathcal{M}_{3\star\star}$, the connectives of negation and implication are specified in the same way as in $\mathcal{M}_I$, but conjunction and disjunction are defined as follows:

[6] See [Sette, Carnielli, 1995; Karpenko, Tomova, 2017; Ciuciura, 2015], for details.

$\wedge$	1	2	0
1	1	2	2
2	2	2	2
0	2	2	0

$\vee$	1	2	0
1	1	1	1
2	1	2	2
0	1	2	0

Each axiom schema of QD^1 is valid in $\mathcal{M}_{3\star\star}$ and the rule of detachment preserves validity. To show that (nC^p) is unprovable in QD^1, it is enough to assign 2 to p in $\neg(p \wedge \neg p)$. ■

Since $CLaN$, D_{min}, Q^1 and QD^1 are proper subsystems of I^1, I^1 is the *strongest* calculus among the paracomplete calculi that have been discussed so far. Moreover, I^1 is *maximal* in the sense that if we enrich the calculus with any classical tautology, which is not valid in I^1, the resulting calculus collapses into CPC. It means that there is no *structural* proper subsystem of CPC stronger than I^1. But 'Is $CLaN$ the *weakest* paracomplete calculus?', or: 'Is there a proper subsystem of $CLaN$ admitting CPC^+ and (MP)?' Some results supporting a positive answer were suggested in *Section 7* of [Nowak, 1998]. The requested calculus, denoted as $\vdash_{Cl1}$, is defined by CPC^+, (MP) and $(DS^\star)$ $\alpha \to (\neg\alpha \to \neg\beta)$.

Remark 7. $CPC^+ \sqsubset \vdash_{Cl1} \sqsubset CLaN$.

Proof. It is obvious that $CPC^+ \sqsubset \vdash_{Cl1}$. For $\vdash_{Cl1} \sqsubset CLaN$, note that $(DS^\star)$ is an instance of (DS). Thus all the axiom schemas of $\vdash_{Cl1}$ are theses of $CLaN$. To show that $p \to (\neg p \to q)$ is unprovable in $\vdash_{Cl1}$, apply the classical truth tables for implication, conjunction and disjunction plus the following one for negation (1 is the designated value):

	$\neg$
1	1
0	1

■

Let us summarize that the lattice relationships between the calculi can be represented by the structure of Figure 1.

3. Paracomplete calculi. Semantics

A Kripke-type semantics for $\vdash_{Cl1}$ was given in [Nowak, 1998, p. 98]; a valuation semantics for $CLaN$ was introduced in [Batens et all, 1999, p. 32]; and a three-valued semantics for I^1 was proposed in [Sette, Carnielli, 1995, p. 190]; an alternative semantics for I^1 was discussed in [Fernández, Coniglio, 2003]. In this section, we propose a bi-valuational semantics for the calculi Q^1 and QD^1; let $\mathcal{C}\star \in \{Q^1, QD^1\}$, for the sake of brevity.

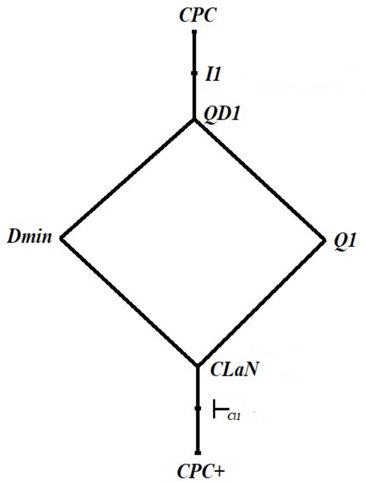


Fig. 1. A lattice of the paracomplete calculi.

Definition 5. A $\mathcal{C}\star$-valuation is any function $v : \mathcal{F} \longrightarrow \{1, 0\}$ that satisfies, for any $\alpha, \beta \in \mathcal{F}$, the following conditions:

($\vee$) $v(\alpha \vee \beta)$=1 iff $v(\alpha)$=1 or $v(\beta)$=1
($\wedge$) $v(\alpha \wedge \beta)$=1 iff $v(\alpha)$=1 and $v(\beta)$=1
($\rightarrow$) $v(\alpha \rightarrow \beta)$=1 iff $v(\alpha)$=0 or $v(\beta)$=1
($\neg$) if $v(\neg\alpha)$=1, then $v(\alpha)$=0,

and additionally,

($\neg\neg$) if $v(\neg\neg\alpha)$=0, then ($v(\alpha)$=1 or $v(\neg\alpha)$=1), for $\mathcal{C}\star = Q^1$
($\neg\neg$) if $v(\neg\neg\alpha)$=0, then $v(\neg\alpha)$=1, for $\mathcal{C}\star = QD^1$.

Definition 6. A formula α is a $\mathcal{C}\star$-tautology if, and only if for every $\mathcal{C}\star$-valuation v, $v(\alpha) = 1$. For any $\alpha \in \mathcal{F}$ and $\Gamma \subseteq \mathcal{F}$, α is a semantic consequence of Γ ($\Gamma \models_{C\star} \alpha$, in symbols) iff for any $\mathcal{C}\star$-valuation v: if $v(\beta) = 1$ for any $\beta \in \Gamma$, then $v(\alpha) = 1$.

The proof of soundness can be obtained in the standard way, by induction on the length of a derivation in $C\star$.

Theorem 2. *For every $\Gamma \subseteq \mathcal{F}$ and $\alpha \in \mathcal{F}$, we have if $\Gamma \vdash_{C\star} \alpha$, then $\Gamma \models_{C\star} \alpha$.*

For the proof of completeness, we apply the method which is based on the notion of maximal non-trivial sets of formulas. We use the technique proposed in [Carnielli, Coniglio, 2016, Section 2.2]. Before going further, let us recall some important definitions and results. Let $\mathcal{C} = \langle \mathcal{F}, Ax_{\mathcal{C}}, \vdash_{\mathcal{C}} \rangle$ be a calculus (satisfying Tarskian properties) and $\Delta \subseteq \mathcal{F}$.

Definition 7. We say that Δ is a closed theory of $\mathcal{C}$ if, and only if for any $\beta \in \mathcal{F}$: $\Delta \vdash_{\mathcal{C}} \beta$ *iff* $\beta \in \Delta$. We say that Δ is maximal non-trivial with respect to $\alpha \in \mathcal{F}$ in $\mathcal{C}$, if, and only if (i) $\Delta \nvdash_{\mathcal{C}} \alpha$, and (ii) for every $\beta \in \mathcal{F}$, if $\beta \notin \Delta$ then $\Delta \cup \{\beta\} \vdash_{\mathcal{C}} \alpha$.

Lemma 2 ([Carnielli, Coniglio, 2016], Lemma 2.2.5)**.** *Every maximal non-trivial set with respect to some formula is a closed theory.*

Observe that the lemma holds for $\mathcal{C}\star$. Moreover, we have:

Lemma 3. *For any maximal non-trivial set Δ with respect to α in $\mathcal{C}\star$ the mapping $v : \mathcal{F} \longrightarrow \{1, 0\}$ defined, for any $\delta \in \mathcal{F}$, as ($\star$): $v(\delta) = 1$ if and only if $\delta \in \Delta$, is a $\mathcal{C}\star$-valuation.*

Proof. We only prove the clauses for negation. The rest of the proof is similar to that of *Theorem 2.2.7* in [Carnielli, Coniglio, 2016].

Assume, for a contradiction, that $v(\neg\beta) = 1$ and $v(\beta) = 1$. Thus we have $\neg\beta \in \Delta$ and $\beta \in \Delta$ by ($\star$). This implies, by *Lemma 1(1)*, that $\Delta \vdash_{\mathcal{C}\star} \neg\beta$ and $\Delta \vdash_{\mathcal{C}\star} \beta$. But, if $\Delta \vdash_{\mathcal{C}\star} \neg\beta$ and $\Delta \vdash_{\mathcal{C}\star} \beta$, then $\Delta \vdash_{\mathcal{C}\star} \{\neg\beta, \beta\}$. Since $\emptyset \vdash_{\mathcal{C}\star} \beta \to (\neg\beta \to \gamma)$, thus $\{\beta, \neg\beta\} \vdash_{\mathcal{C}\star} \gamma$, by the deduction theorem. The relation $\vdash_{\mathcal{C}\star}$ is transitive, so $\Delta \vdash_{\mathcal{C}\star} \gamma$. Notice that Δ is a closed theory, so $\alpha \in \Delta$. But $\alpha \notin \Delta$ (by the main assumption). This yields a contradiction.

If $\mathcal{C}\star = Q^1$, we need to show that the mapping v satisfies the following clause: if $v(\neg\neg\beta) = 0$ then ($v(\beta) = 1$ or $v(\neg\beta) = 1$), for any $\beta \in \mathcal{F}$. Assume, for a contradiction, that $v(\neg\neg\beta) = 0$ and $v(\neg\beta) = v(\beta) = 0$. Thus we have $\neg\neg\beta \notin \Delta$, $\neg\beta \notin \Delta$ and $\beta \notin \Delta$ by ($\star$). Since Δ is a maximal non-trivial set with respect to α, $\Delta \cup \{\beta\} \vdash_{Q^1} \alpha$, $\Delta \cup \{\neg\beta\} \vdash_{Q^1} \alpha$ and $\Delta \cup \{\neg\neg\beta\} \vdash_{Q^1} \alpha$. Consequently, $\Delta \cup \{\beta \vee \neg\beta \vee \neg\neg\beta\} \vdash_{Q^1} \alpha$, by *Lemma 1 (5)*. Note that $\emptyset \vdash_{Q^1} \beta \vee \neg\beta \vee \neg\neg\beta$, so $\Delta \vdash_{Q^1} \alpha$, by *Lemma 1 (6)*. Since Δ is a closed theory, then $\alpha \in \Delta$. But $\alpha \notin \Delta$. This yields a contradiction.

If $\mathcal{C}\star = QD^1$, we have to prove that the mapping v satisfies the clause: if $v(\neg\neg\beta) = 0$ then $v(\neg\beta) = 1$, for any $\beta \in \mathcal{F}$. Assume, for a contradiction, that $v(\neg\neg\beta) = 0$ and $v(\neg\beta) = 0$. Then we have $\neg\neg\beta \notin \Delta$ and $\neg\beta \notin \Delta$ by ($\star$). Since Δ is a maximal non-trivial set with respect to α, then $\Delta \cup \{\neg\neg\beta\} \vdash_{QD^1} \alpha$ and $\Delta \cup \{\neg\beta\} \vdash_{QD^1} \alpha$. Consequently, $\Delta \cup \{\neg\beta \vee \neg\neg\beta\} \vdash_{QD^1} \alpha$, by *Lemma 1 (6)*. It is known that $\emptyset \vdash_{QD^1} \neg\beta \vee \neg\neg\beta$, so $\Delta \vdash_{QD^1} \alpha$, by *Lemma 1 (6)*. Since Δ is a closed theory, then $\alpha \in \Delta$. But $\alpha \notin \Delta$. This yields a contradiction. ∎

Note that the so-called Lindenbaum–Łoś' theorem holds, for any finitary calculus $\mathcal{C} = \langle \mathcal{F}, Ax_{\mathcal{C}}, \vdash_{\mathcal{C}} \rangle$.

Lemma 4 ([Pogorzelski, Wojtylak, 2008], Theorem 3.31; [Carnielli, Coniglio, 2016], Theorem 2.2.6)**.** *For any $\Gamma \subseteq \mathcal{F}$ and $\alpha \in \mathcal{F}$ such that $\Gamma \nvdash_{\mathcal{C}} \alpha$, there is a maximal non-trivial set Δ with respect to α in $\mathcal{C}$ such that $\Gamma \subseteq \Delta$.*

Thus, the completeness of $\mathcal{C}\star$ follows:

Theorem 3. *For all* $\Gamma \subseteq \mathcal{F}$ *and* $\alpha \in \mathcal{F}$*: if* $\Gamma \models_{\mathcal{C}\star} \alpha$*, then* $\Gamma \vdash_{\mathcal{C}\star} \alpha$*.*

Proof. Assume that $\Gamma \nvdash_{\mathcal{C}\star} \alpha$ and Δ be a maximal non-trivial set with respect to α in $\mathcal{C}\star$ such that $\Gamma \subseteq \Delta$. Then $\alpha \notin \Delta$. Because *Lemma 3* holds, there is a valuation v such that $v(\alpha) = 0$ and $v(\beta) = 1$, for any $\beta \in \Gamma$. Hence $\Gamma \not\models_{\mathcal{C}\star} \alpha$. ■

4. A hierarchy of the paracomplete calculi

The $\vdash_{Cl1}$, $CLaN$, D_{min}, Q^1, QD^1 and I^1 are not the only paracomplete calculi that satisfy the criteria specified in *Definition 2*. In fact, there are infinitely many such calculi, for example, Q^1, Q^2, ..., Q^n; or QD^1, QD^2,... QD^n. The hierarchy of Q^n–calculi, $n \in \mathbb{N}$, was considered in [Ciuciura, 2019]. In the subsequent paragraphs we will discuss the hierarchy of QD^n–calculi. The hierarchy is obtained by replacing $(ExM^{\neg})$ with a more general schema, that is,

$(ExM^{\neg n})$ $\neg^n\alpha \vee \neg^{n+1}\alpha$,

where $n \in \mathbb{N}$ and $\neg^n\alpha$ is an abbreviation for $\overbrace{\neg\neg ... \neg}^{n}\alpha$. To put it more precisely, for each $n \in \mathbb{N}$, let QD^n be obtained from CPC^+ (and (MP)) by adding to it the axiom schemas:

(DS) $\alpha \to (\neg\alpha \to \beta)$
$(ExM^{\neg n})$ $\neg^n\alpha \vee \neg^{n+1}\alpha$.[7]

For each $n \in \mathbb{N}$, the semantics for QD^n results from replacing the evaluation condition for $(\neg\neg)$ with a more general one, *i.e.*

$(\neg^{n+1})$ if $v(\neg^{n+1}\alpha)$=0, then $v(\neg^n\alpha)$=1.

The semantic clauses for $(\vee)$, $(\wedge)$, $(\to)$ and $(\neg)$ remain unchanged, *i.e.*

Definition 8. A QD^n-valuation is any function $v : \mathcal{F} \longrightarrow \{1, 0\}$ that satisfies, for any $\alpha, \beta \in \mathcal{F}$, the conditions:
$(\vee)$ $v(\alpha \vee \beta)$=1 iff $v(\alpha)$=1 or $v(\beta)$=1
$(\wedge)$ $v(\alpha \wedge \beta)$=1 iff $v(\alpha) = 1$ and $v(\beta)$=1

[7] If $n = 0$, then $QD^0 = CPC$. Some other examples of the hierarchies are known in the logical literature. For instance, the hierarchy of I^n–calculi is proposed in [Sette, Carnielli, 1995] and [Fernández, Coniglio, 2003]. There are also interesting hierarchies in a Newton da Costa-style presentation, *e.g.* da Costa and Marconi's hierarchy of paracomplete calculi P_n, see [da Costa, Marconi, 1986]; or Arruda–Alves' logic of vagueness, see [Arruda, Alves, 1979] and [Arruda, Alves, 1979].

($\to$) $v(\alpha \to \beta)$=1 iff $v(\alpha)$=0 or $v(\beta)$=1
($\neg$) if $v(\neg\alpha)$=1, then $v(\alpha)$=0,
($\neg^{n+1}$) if $v(\neg^{n+1}\alpha)$=0, then $v(\neg^{n}\alpha)$=1, where $n \in \mathbb{N}$.

The definition of QD^n-tautology (and semantic consequence $\models_{QD^n}$) is analogous to that of *Definition 6.*

Theorem 4. *For every* $\Gamma \subseteq \mathcal{F}$ *and* $\alpha \in \mathcal{F}$, $\Gamma \vdash_{QD^n} \alpha$ *iff* $\Gamma \models_{QD^n} \alpha$, $n \in \mathbb{N}$.

Proof. Proceed analogously to the proof of *Theorems 2 and 3.* ■

At the end of this section, we state a few simple facts about the QD^n–calculi.

Remark 8. If $n > 1$, then the formula $p \to \neg\neg p$ is not provable in QD^n.

Proof. This follows from the completeness of QD^n–calculi. ■

Remark 9. If $n > 1$, then
(1) $D_{min} \not\sqsubseteq QD^n$
(2) $QD^n \not\sqsubseteq D_{min}$.

Proof. *(1)*: Although $(ExM^{\neg n})$ is an axiom schema of QD^n, the formula $\neg^n p \vee \neg^{n+1} p$ is not provable in D_{min} (it is enough to apply the semantics and completeness theorem for D_{min}, *cf.* [Carnielli, Marcos, 1999], *Proposition 6.2*). *(2)*: This is a consequence of *Remark 8* and the fact that $(NN^{\star})$ is an axiom schema of D_{min}. ■

Remark 10. For any r, $m \in \mathbb{N}$ such that $r > m$, we have $QD^r \sqsubset QD^m$.

Proof. The proof follows from the completeness of QD^n–calculi. ■

Remark 11. Enriching the set of axiom schemas of any QD^n-calculus ($n \in \mathbb{N}$) with the formula (NN) $\neg\neg\alpha \to \alpha$, results in obtaining the axiom system of CPC.

Proof. This follows from the fact that the axiom schemas $(ExM^{\neg n})$ and (NN) are equivalent to (ExM) $\alpha \vee \neg\alpha$ in CPC. ■

Remark 12. Enriching the set of axiom schemas of any QD^n-calculus (for $n > 1$) with the formula $\neg\neg\neg\alpha \to \neg\alpha$, results in obtaining the calculus QD^1.

Proof. Notice that *(1)* $\neg\neg\neg\alpha \to \neg\alpha$ is a thesis of QD^1, and *(2)* $\neg\neg\neg p \to \neg p$ is not provable in any QD^n-calculus that is weaker than QD^1. Now it suffices to show that $(ExM^{\neg n})$, where $n > 1$, and $\neg\neg\neg\alpha \to \neg\alpha$ are equivalent to $(ExM^{\neg})$ in QD^1. ■

Acknowledgements. I am very grateful to anonymous reviewers for their helpful comments on an earlier draft of this paper.

References

Arruda, Alves, 1979 – Arruda, A.I., Alves, E.H. "Some remarks on the logic of vagueness", *Bulletin of the Section of Logic*, 1979, Vol. 8, No. 3, pp. 133–138.

Arruda, Alves, 1979 – Arruda A.I., Alves, E.H. "A semantical study of some systems of vagueness logic", *Bulletin of the Section of Logic*, 1979, Vol. 8, No. 3, pp. 139–144.

Batens et all, 1999 – Batens, D., De Clercq, K., Kurtonina, N. "Embedding and interpolation for some paralogics. The propositional case", *Reports on Mathematical Logic*, Vol. 33, pp. 29–44.

Beall, 2017 – Beall, Jc. "There is no logical negation: true, false, both and neither", *Australasian Journal of Logic*, 2017, Vol. 14, No. 1, pp. 1–29.

Bolotov et all, 2018 – Bolotov, A., Kozhemiachenko, D., Shangin, V. "Paracomplete Logic Kl – Natural Deduction, its Automation, Complexity and Applications", *Journal of Applied Logics*, 2018, Vol. 5, No. 1, pp. 221–261.

Carnielli, Coniglio, 2016 – Carnielli, W., Coniglio, M.E., *Paraconsistent Logic: Consistency, Contradiction and Negation.* Logic, Epistemology, and the Unity of Science. Vol. 40. Springer International Publishing, 2016.

Carnielli, Marcos, 1999 – Carnielli, W., Marcos, J., "Limits for paraconsistent calculi", *Notre Dame Journal of Formal Logi*, 1999, Vol. 40, No. 3, pp. 375–390.

Ciuciura, 2015 – Ciuciura, J., "A Weakly-Intuitionistic Logic I1", *Logical Investigations*, 2015, Vol. 21, No. 2, pp. 53—60.

Ciuciura, 2019 – Ciuciura, J. "Paraconsistency and Paracompleteness", *Logical Investigations*, 2019, Vol. 25, No. 2, pp. 46–60.

da Costa, Marconi, 1986 – da Costa, N.C.A., Marconi, D., "A note on paracomplete logic", in: *Atti della Accademia Nazionale dei Lincei*, Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, 1986, Vol. 80, No. 7–12, pp. 504–509.

Fernández, Coniglio, 2003 – Fernández, V.L., Coniglio, M.E., "Combining Valuations with Society Semantics", *Journal of Applied Non-Classical Logics*, 2003, Vol. 13, No. 1, pp. 21–46.

Karpenko, Tomova, 2017 – Karpenko, A., Tomova, N., "Bochvar's three-valued logic and literal paralogics: Their lattice and functional equivalence", *Logic and Logical Philosophy*, 2017, Vol. 26, No. 2, pp. 207–235.

Loparić, da Costa, 1984 – Loparić, A., da Costa, N.C.A., "Paraconsistency, Paracompleteness, and Valuations", *Logique et Analyse*, 1984, Vol. 27, No. 106, pp. 119–131.

Nowak, 1998 – Nowak, M. "Kripke semantics for some paraconsistent logics", *Logica Trianguli*, 1998, Vol. 2, pp. 87–101.

Petrukhin, 2018 – Petrukhin, Y. "Generalized Correspondence Analysis for Three-Valued Logics", *Logica Universalis*, 2018, Vol. 12, No. 3–4, pp. 423–460.

Pogorzelski, Wojtylak, 2008 – Pogorzelski, W.A., Wojtylak, P., *Completeness Theory for Propositional Logics*, Studies in Universal Logic, Birkhäuser Basel, 2008.

Popov, 2002 – Popov, V.M. "On a three-valued paracomplete logic", *Logical Investigations*, 2002, Vol. 9, pp. 175–178. (In Russian)

Sette, Carnielli, 1995 – Sette, A.M., Carnielli, W.A. “Maximal weakly-intuitionistic logics”, *Studia Logica*, 1995, Vol. 55, No. 1, pp. 181–203.

Wójcicki, 1988 – Wójcicki, R. *Theory of Logical Calculi. Basic Theory of Consequence Operations*, Synthese Library, Vol. 199, Springer Netherlands, 1988.

Логические исследования
2020. Т. 26. № 1. С. 124–143
УДК 162.3+510.643

Logical Investigations
2020, Vol. 26, No. 1, pp. 124–143
DOI: 10.21146/2074-1472-2020-26-1-124-143

Прикладная логика
Applied logic

В.К. Финн

О логиках эмпирических модальностей

Виктор Константинович Финн
ФИЦ УИ РАН.
Российская Федерация, 119333, г. Москва, ул. Вавилова, 44, корп. 2.
Российский государственный гуманитарный университет.
Российская Федерация, 125993, г. Москва, Миусская пл., 6.
E-mail: `ira.finn@gmail.com`

Аннотация: В статье рассматривается класс ERA-логик с эмпирическими модальностями $\Box$ (необходимость) и $\Diamond$ (возможность), которые характеризуют, соответственно, высказывания, представляющие эмпирические законы и эмпирические тенденции, т. е. эмпирические закономерности. Эмпирические закономерности являются результатом ДСМ-рассуждений, которые образованы взаимодействием правил индуктивного вывода и правил вывода по аналогии, а также процедурами абдуктивного принятия гипотез. Рассматриваемые ERA-логики являются пропозициональной имитацией ДСМ-рассуждений, применимых к последовательностям расширяемых баз фактов интеллектуальных систем. Характерной особенностью ERA-логик является применение двух концепций истины – когерентной и корреспондентной. Применение когерентной концепции истины обусловлено порождением гипотез посредством правил индуктивного вывода и вывода по аналогии. Применение же корреспондентной концепции истины обусловлено применением абдуктивного вывода, принятие результатов которого использует верификацию гипотез о предсказаниях. С этой целью ERA-логики применяют оператор T: «истинно, что...». В заключении статьи обсуждаются нефинитные расширения ERA-логик, а также их отличия как логик эмпирических модальностей от логики M логических модальностей Г.Х. фон Вригта.

Ключевые слова: ДСМ-рассуждения, правила индуктивного вывода, правила вывода по аналогии, абдукция, эмпирические закономерности, эмпирический закон, эмпирическая тенденция, эмпирические модальности, логические модальности, оператор «истинно, что...», номологические высказывания

Для цитирования: *Финн В.К.* О логиках эмпирических модальностей // Логические исследования / Logical Investigations. 2020. Т. 26. № 1. С. 124–143. DOI: 10.21146/2074-1472-2020-26-1-124-143

Универсальность двузначной логики связана как с её простотой, порождаемой законом исключенного третьего и законом противоречия, а также с тем фактом, что в некоторых «интересных» неклассических логиках (например, B_3 Д.А. Бочвара, $Ł_n$ Я. Лукасевича, модальных логиках) либо сохраняются логические связки двузначной логики на ограничении истинностных значений «истина» и «ложь», либо предполагается использование двузначной логики в соответствующих фрагментах этих неклассических логик.

Однако имеются и логики для специфических рассуждений, порожденных ориентацией на решение соответствующих проблем. Интуиционистская логика формализует конструктивность доказательств, трехзначная логика B_3 применима для анализа парадоксов, логики $Ł_n$ связаны с простыми числами [Бочвар, 1938; Łukasiewicz, 1920], четырехзначные логики аргументации [Финн, 2006] используются для формализации социологических опросов.

ДСМ-метод автоматизированной поддержки исследований (ДСМ-метод АПИ) реализует ДСМ-рассуждения и ДСМ-исследования, которые посредством применения ДСМ-рассуждений к расширяемым последовательностям баз фактов («возможным мирам») порождают эмпирические закономерности (ЭЗК) – эмпирические законы и эмпирические тенденции [Аншаков, 2009; Финн, Шестерникова, 2018; Финн, 2019; Финн, 2020a].

Этап применения ДСМ-метода АПИ к расширяемым последовательностям баз фактов образует ДСМ-исследования, результатом которых является поддержка и расширение открытых эмпирических теорий (квазиаксиоматических теорий [Финн, 2019]).

Квазиаксиоматические теории образованы множествами фактов, открытым и пополняемым множеством аксиом и правилами вывода (правдоподобными и дедуктивными). Правдоподобными выводами являются правила индуктивного вывода и вывода по аналогии. Взаимодействие этих правил, принятие порожденных гипотез о причине (результат индукции) и гипотез о предсказании (результат аналогии) посредством абдукции $1^{\underline{\text{ого}}}$ рода образуют ДСМ-рассуждения, применяемые к базе фактов («возможному миру»). Заметим, что правила индуктивного вывода ДСМ-рассуждений являются формализацией и усилением известных канонов индукции Д.С. Милля [Финн, 2020b].[1]

Продолжение применений ДСМ-рассуждений к последовательностям расширяемых баз фактов, представляющих истории возможных миров HPW_h, образуют **ДСМ-исследование**, завершаемое абдуктивным выво-

[1] В Приложении в [Финн, 2020b] представлены формализации индуктивных канонов Д.С. Милля.

дом гипотезы о причине с соответствующим модальным оператором (результат абдукции $2^{\underline{\text{ого}}}$ рода [Финн, 2019; Финн, 2020a]).[2]

Правила индуктивного вывода ДСМ-рассуждений образуют дистрибутивную решетку [Финн, 2014; Финн, 2016], представляющую возможные стратегии ДСМ-рассуждений $Str_{x,y}$. Каждая $Str_{x,y}$ из множества всех стратегий $\overline{Str}$ применяется к множеству всех историй возможных миров $\overline{HPW}$ для обнаружения **эмпирических закономерностей**. Эмпирической закономерностью является сохранение гипотез о причинах и гипотез о предсказаниях в историях возможных миров HPW_h из $\overline{HPW}$ таких, что в последней базе фактов этих историй возможных миров гипотезы о предсказаниях **верифицируются**. Эта верификация используется в абдуктивном выводе абдукции $2^{\underline{\text{ого}}}$ рода, посредством которой принимается гипотеза о причине, сохраняемая в историях возможных миров. К этой гипотезе и применим соответствующий модальный оператор[3].

Порождаемые ДСМ-исследованием эмпирические закономерности являются **эмпирическими номологическими высказываниями**, понятие номологических высказываний было предложено Гансом Рейхенбахом в [Reichenbach, 1947; Reichenbach, 1954]. Посредством номологических высказываний он определял физические модальности. Однако номологические высказывания Г. Рейхенбаха выражали как физические закономерности, так и логические законы.

В [Финн, Шестерникова, 2018; Финн, 2019; Финн, 2020a] были определены эмпирические модальности, порожденные ДСМ-исследованиями и соответствующие пропозициональные модальные логики семейства ERA – логики эмпирических закономерностей (ER) и абдукции (A).

В [Финн, Шестерникова, 2018; Финн, 2019] были определены четырнадцать модальностей, соответствующих эмпирическим законам (8 модальностей), эмпирическим тенденциям (4 модальности) и слабым эмпирическим тенденциям (2 модальности). Факторизация частично упорядоченного множества модальностей порождает два варианта модальных логик – модальные логики трех модальностей (необходимость $\Box$, возможность $\Diamond$ и слабая возможность $\triangledown$), соответствующие эмпирическим законам, тенденциям и слабым тенденциям [Финн, Шестерникова, 2018] и модальные логики двух модальностей ($\Box$, $\Diamond$).

[2] В [Фейс, 1965] Р. Фейс замечает, что модальности могут применяться для описания физического мира и что модальности могут быть использованы для анализа причинности [Фейс, 1965, с. 24]. Эта идея реализована в ДСМ-методе АПИ в компьютерных системах интеллектуального анализа данных [Аншаков, 2009; Финн, 2019; Финн, 2020a]

[3] Множество модальных операторов частично упорядочено и имеет наибольший и наименьший элементы [Финн, Шестерникова, 2018].

В этих логиках импликация $\to$ истолковывается как «если p – причина эффекта, то q – предсказание эффекта». $\Box(p \to q)$, $\Diamond(p \to q)$, $\nabla(p \to q)$ истолковываются как представление ЭЗК типа «эмпирический закон», «эмпирическая тенденция» и «слабая эмпирическая тенденция», соответственно. Tq означает верификацию q, а T – оператор «истинно, что ...», который аналогичен оператору T Г. фон Вригта [фон Вригт, 1971], но отличен от него.

В [Финн, 2019] представлена модальная логика двух модальностей $(\Box, \Diamond)$ $ERA_{0.1}$, имеющая аксиомы, выражающие абдукцию $2^{\underline{\text{ого}}}$ рода [Финн, Шестерникова, 2018; Финн, 2019; Финн, 2020a]: $((\Box(p \to q) \& Tq) \to \Box p)$ и $((\Diamond(p \to q) \& Tq) \to \Diamond p)$.

Важно отметить, что в [Финн, Шестерникова, 2018; Финн, 2019] эмпирические закономерности характеризуются соответствующими регулярными кодами $Cd(j,h)$ такими, что для возможных миров $j = 0, 1, \ldots, s$ и их историй $h = 1, \ldots, (s+1)!$ имеются последовательности $\nu \ldots \nu$, где $\nu = 1, -1$ и $\tau \ldots \tau\ \nu \ldots \nu$ длины s такие, что $\nu \ldots \nu$ характеризует эмпирический закон, а $\tau \ldots \tau\ \nu \ldots \nu$ – эмпирическая тенденция, где $1, -1, \tau$ – типы истинностных значений гипотез – фактическая истина, фактическая ложь и неопределенность, соответственно. Коды $\nu \ldots \nu$ и $\tau \ldots \tau\ \nu \ldots \nu$ есть значения пропозициональных переменных, представляющих ЭЗК, любой отличный от них код является нерегулярным и соответствует отсутствию ЭЗК. Регулярные коды $\nu \ldots \nu$ и $\tau \ldots \tau\ \nu \ldots \nu$ делают истинными пропозициональные переменные $p, q, \ldots$, а нерегулярные коды делают истинными отрицания пропозициональных переменных $\neg p, \neg q, \ldots$[4]

Итерации модальностей в логиках семейства ERA означают вид расширений возможных миров – сохранение гипотез в возможных мирах, что представимо кодами типа $\nu \ldots \nu$, где $\nu = 1, -1$ (только истина или только ложь) и кодами типа $\tau \ldots \tau\ \nu \ldots \nu$, где начало кода – подпоследовательность неопределенностей (τ), выражающая эмпирическую тенденцию.

Рассмотрим возможные итерации модальностей: $\Box\Box p, \Box\Diamond p, \Diamond\Box p$ и $\Diamond\Diamond p$. Имеются два способа (направления) расширений возможных миров: справа налево (от переменной $p :\leftarrow p$) и слева направо (от левой модальности $M : M \to$, где M есть $\Box, \Diamond$).

В [Финн, 2019] $ERA_0, ERA_{0.1}, ERA_1$ используют расширения возможных миров справа налево, что представимо аксиомами $\Box\Diamond p \to \Diamond p$, $\Diamond\Box p \to \neg p$, $\Box\Box p \to \Box p, \Diamond\Diamond p \to \neg p$.

[4]Нерегулярные коды ЭЗК могут содержать вхождение 0 – тип истинностных значений «фактическое противоречие» [Финн, Шестерникова, 2018].

В настоящей статье рассмотрим исчисление ERA_{0^*}, $ERA_{0.2}$ и ERA_2 такие, что итерации модальностей рассматриваются слева направо (в направлении от левой модальности к переменной p).

Таким образом, $\Box\Diamond p$ имеет следующий код ЭЗК $\nu\ldots\nu\ \tau\ldots\tau\ \nu\ldots\nu$ такой, что он является нерегулярным, а, следовательно, характеризует отсутствие ЭЗК, а потому истинна аксиома $\Box\Diamond p \to \neg p$, так как результирующий код $\Box\Diamond p$ есть $\nu\ldots\nu\ \tau\ldots\tau\ \nu\ldots\nu$ такой, что он нерегулярный. Аналогично получим истинную формулу $\Diamond\Box p \to \Diamond p$, соответствующую регулярному коду $\tau\ldots\tau\ \nu\ldots\nu$.

Ниже сформулируем исчисления ERA_{0^*}, $ERA_{0.2}$ и ERA_2.

Алфавит:

$p, q, r, \ldots,$ (быть может с нижними индексами) – пропозициональные переменные;

$\neg, \&, \vee, \to, \Box, \Diamond, T$ – логические связки;

$(,)$ – скобки.

Посредством букв греческого алфавита будем обозначать метасимволы для формул.

Определение 1. Определение формулы

1^0. $p, q, r, \ldots$ – формулы;

2^0. $\neg p, \neg q, \neg r, \ldots$ – формулы;

3^0. если φ, ψ – формулы, то $(\varphi \,\&\, \psi), (\varphi \vee \psi), (\varphi \to \psi)$ – формулы;

4^0. если φ – формула, то $\Box\varphi, \Diamond\varphi$ – формулы;

5^0. если φ – формула, то $T\varphi$ – формула;

6^0. если φ – формула, то $\neg\varphi$ – формула;

7^0. других формул нет.

$$ERA_{0^*}.$$

(1) Аксиомы двузначной логики L_2.

(2) Аксиомы ERA_{0^*}

($\Box 2$) $\Box p \to p$

($\Diamond 2$) $\Diamond p \to p$

($\Box 3$) $\Box(p \,\&\, q) \leftrightarrow (\Box p \,\&\, \Box q)$

($\Box 4$) $\Box(p \vee q) \leftrightarrow (\Box p \vee \Box q)$

($\Diamond 3$) $\Diamond(p \,\&\, q) \leftrightarrow (\Diamond p \,\&\, \Diamond q)$

($\Diamond 4$) $\Diamond(p \vee q) \leftrightarrow (\Diamond p \vee \Diamond q)$

($\neg\Box$) $\neg\Box p \to (\Diamond p \vee \neg p)$

($\neg\Diamond$) $\neg\Diamond p \to (\Box p \vee \neg p)$

($\Box\Box$) $\Box\Box p \to \Box p$

$(\Box\Diamond)_2$ $\Box\Diamond p \to \neg p$

$(\Diamond\Box p)_2$ $\Diamond\Box p \to \Diamond p$

($\Diamond\Diamond$) $\Diamond\Diamond p \to \neg p$

($\Box \,\&\, \Diamond$) $\neg(\Box p \,\&\, \Diamond p)$

($\Box \,\&\, \neg$) $\neg(\Box p \,\&\, \neg p)$

($\Diamond \,\&\, \neg$) $\neg(\Diamond p \,\&\, \neg p)$

($\Box\neg$) $\neg\Box\neg p$

($\Diamond\neg$) $\neg\Diamond\neg p$

($\Box\neg\Diamond$) $\Box p \to \neg\Diamond p$

$\varphi \leftrightarrow \psi \rightleftharpoons (\varphi \to \psi) \,\&\, (\varphi \to \psi)$[5]

$f \rightleftharpoons p \,\&\, \neg p$

$t \rightleftharpoons \neg f$

(3) Правила вывода ERA_{0^*}:

R1. $\varphi, (\varphi \to \psi) \vdash \psi$, где «$\vdash$» есть метасимвол отношения выводимости;

R2. $\varphi(p) \vdash \varphi(q)$, $\varphi(\chi) = \int_p^{\chi} \varphi(p)|$ – правило подстановки (Sub);

R3. $\Box\varphi, \Box(\varphi \to \psi) \vdash \Box\psi$;

R4. $\Diamond\varphi, \Diamond(\varphi \to \psi) \vdash \Diamond\psi$;

R5. $\varphi(\chi) \vdash \varphi(\chi_1)$, где $(\chi \leftrightarrow \chi_1)$, – правило замены эквивалентных формул.

Определение доказуемой формулы (обозначение: $\vdash \varphi$) стандартно.

[5] Символ «$\rightleftharpoons$» означает равенство по определению.

Утверждение 1. ERA_{0^*} является противоречивым.

Доказательство. $\neg\Box p \to (\Diamond p \vee \neg p), \Box p \vee \Diamond p \vee \neg p, \Box\neg p \vee \Diamond\neg p \vee \neg\neg p$, $\Box\neg p \leftrightarrow f$, $\Diamond\neg p \leftrightarrow f$ $[(\Box\neg), (\Diamond\neg), \neg t \leftrightarrow f]$; $f \vee f \vee p, p$.

Таким образом, $Sub\,\neg p$ в $\Box p \vee \Diamond p \vee \neg p$ порождает $\vdash p$ – противоречивость в смысле Э. Поста. Поэтому в силу Sub получаем $\vdash \varphi$ для любой φ, то есть, абсолютную (или тривиальную) противоречивость ERA_{0^*} ■

Получим тогда исчисление $ERA_{0.2}$ посредством ограничения R2 следующим образом: $\mathrm{R}_2^*\, \varphi(p) \vdash \varphi(\chi), \varphi(\chi) = \int_p^\chi \varphi(p) \mid$, где $\chi \in [\{(p\,\&\,q), (p \vee q)\}]$, а $[\{(p\,\&\,q), (p \vee q)\}]$ – замыкание $\{(p\,\&\,q), (p \vee q)\}$, соответствующее подмножеству множества монотонных булевских функций.

Таким образом, исчисление $ERA_{0.2}$ имеет аксиомы ERA_{0^*} и правила вывода R1, R_2^*, R3, R4 и R5 (то есть, ограниченную Sub). Это ограничение вызвано тем, что $ERA_{0.2}$ и ERA_2, формулируемое ниже, пропозициональными средствами имитируют рассуждения относительно эмпирических закономерностей, которые являются значениями пропозициональных переменных p таких, что им соответствуют регулярные коды ЭЗК.

Заметим, что $ERA_{0.2}$ получено из $ERA_{0.1}$ [Финн, 2019] заменой $\Box\Diamond p \to \Diamond p$ на $\Box\Diamond p \to \neg p$ и заменой $\Diamond\Box p \to \neg p$ на $\Diamond\Box p \to \Diamond p$.

В [Финн, 2019] была установлена непротиворечивость $ERA_{0.1}$ относительно семантики историй возможных миров HPW_h из $\overline{HPW}$.

Расширим исчисление $ERA_{0.2}$, добавив фрагмент TM_2, формулируемый ниже, и получим исчисление ERA_2.

$$TM_2$$

T1. $Tp \to \neg T\neg p$

T2. $T(Tp) \leftrightarrow Tp$

T3. $Tp \to p$

T4. $T\neg p \to \neg p$

T5. $T(p^{\sigma_1}\,\&\,q^{\sigma_2}) \leftrightarrow (Tp^{\sigma_1}\,\&\,Tp^{\sigma_2})$

T6. $T(p^{\sigma_1} \vee p^{\sigma_2}) \leftrightarrow (Tp^{\sigma_1} \vee Tp^{\sigma_2})$

$$p^\sigma = \begin{cases} p, & \text{если } \sigma = 1 \\ \neg p, & \text{если } \sigma = 0 \end{cases}$$

T7. $T\neg(p\,\&\,q) \leftrightarrow (T\neg p \vee T\neg q)$

T8. $T\neg(p \vee q) \leftrightarrow (T\neg p \,\&\, T\neg q)$

T9. $T(p \to q) \leftrightarrow (Tp \to Tq)$

T10. $T\Box p \leftrightarrow (\Box p \,\&\, Tp)$

T11. $T\Diamond p \leftrightarrow (\Diamond p \,\&\, Tp)$

T12. $((\Box(p \to q) \,\&\, Tq) \to T\Box p)$

T13. $((\Diamond(p \to q) \,\&\, Tq) \to T\Diamond p)$

R6. $T\varphi, T(\varphi \to \psi) \vdash T\psi$

TM_2 может быть расширено добавлением
(1) конечного множества аксиом $T_n = \{Tq_1, \dots, Tq_n\}$,
(2) бесконечного множества $T = \{Tq_1, \dots, Tq_n, \dots\}$.
Тогда получим два варианта исчислений: $TM_2^{(1)}$ и $TM_2^{(2)}$.

Соответственно, получим исчисления $ERA_2^{(1)}$ и $ERA_2^{(2)}$.

(3) Третий вариант исчисления ERA_2 получим посредством добавления к ERA_2 базисного фрагмента TM_2, тогда формулы вида Tq будут применяться как **предположения** для выводов с их использованием.

Таким образом, имеем исчисления ERA_2, $ERA_2^{(1)}$ и $ERA_2^{(2)}$, правилами вывода которых являются R1, R*2, R3, R4, R5 и R6, где ERA_2 есть $ERA_{0.2}$ с добавлением TM_2.

Замечание 1. В исчислениях ERA_2, $ERA_2^{(1)}$ и $ERA_2^{(2)}$ имеются аксиомы абдукции T12 $((\Box(p \to q) \,\&\, Tq) \to T\Box p)$ и T13 $((\Diamond(p \to q) \,\&\, Tq) \to T\Diamond p)$ такие, что они являются усилением аксиом A10 $((\Box(p) \to q) \,\&\, Tq)) \to \Box p)$ и A11 $((\Diamond(p) \to q) \,\&\, Tq)) \to \Diamond p)$ из [Финн, 2019]. Заметим, что A10 и A11 доказуемы в указанных исчислениях.

Замечание 2. Семантика ДСМ-исследований основана на применении двух концепций истины – **когерентной** [Rescher, 1973; Вейнгартен, 2000] и **корреспондентной** [Вейнгартен, 2000; Tarski, 1956]. Когерентная концепция использует соответствие оценки высказывания и некоторого множества непротиворечивых знаний, а корреспондентная – соответствие высказывания и «положения дел», к которому оно относится.

ДСМ-рассуждение, представляющее взаимодействие индукции, аналогии и абдукции $1^{\underline{\text{ого}}}$ рода [Финн, 2019; Финн, 2020a] основано на **принятии** результатов посредством локальных вынуждений, порождающих их оценки посредством правил индуктивного вывода для гипотез о причинах эффектов, и посредством каузальных вынуждений посредством правил вывода по аналогии для гипотез о предсказаниях. Принятие результатов ДСМ-рассуждения завершается применением абдукции $1^{\underline{\text{ого}}}$ рода, реализующей степень объяснения баз фактов («возможных миров») посредством порожденных гипотез о причинах исследуемых эффектов.

Принятие же результатов ДСМ-исследований осуществляется посредством **подтвержденных** ДСМ-рассуждений, применяемых к множеству всех возможных миров с использованием **верификации** гипотез о предсказании исследуемых эффектов и принятия гипотез о причинах посредством абдукции $2^{\underline{\text{ого}}}$ рода, что означает применение **корреспондентной** концепции истины.

Из Замечания 2 следует необходимость определения оценки результатов ДСМ-исследования посредством совместного применения двух типов оценки, соответствующих когерентной и корреспондентной концепцям истины. Применение двух концепций истины отображается в аксиомах абдукции в исчислениях типа ERA для модальностей необходимости и возможности $((\Box(p \to q) \,\&\, Tq) \to T\Box p)$ и $((\Diamond(p \to q) \,\&\, Tq) \to T\Diamond p)$, в которых подформулы $\Box(p \to q)$ и $\Diamond(p \to q)$ оцениваются когерентными истинностными значениями, а подформулы Tq – корреспондентными истинностными значениями, соответствующими верификации гипотез о предсказании исследуемых эффектов. Подформулы $T\Box p$ и $T\Diamond p$ оцениваются **одновременно** когерентной и корреспондентной концепциями истины. В связи с чем в [Финн, 2020a] было введено понятие **косвенной** корреспондентной истины, ибо Tq представляет корреспондентно истинное предсказание (то есть, верифицированное), а так как p выражает причину q и когерентно истинно $\Box(p \to q)$ и $\Diamond(p \to q)$, то косвенно корреспондентно истинно Tp (то есть, верифицируемо), но в силу $\Box(p \to q) \,\&\, Tq$ и $\Diamond(p \to q) \,\&\, Tq$ имеет место когерентная истина $\Box p$ и $\Diamond p$, соответственно. Следовательно, истинна $Tp \,\&\, \Box p$ и $T\Box p$, $Tp \,\&\, \Diamond p$ и $T\Diamond p$, соответственно, а потому истинны $T\Box p$ и $T\Diamond p$, соответственно.

Семантическими основаниями логики ERA_2 (как и логики ERA_1 [Финн, 2019]) является конечное множество конечных историй возможных миров $\overline{HPW}$ такое, что возможным миром является база фактов интеллектуальных систем, а историями возможных миров HPW_h являются конструктивно порождаемые последовательности вложенных возможных миров. Если число расширений баз фактов есть s, то число всех возможных историй возможных миров HPW_h есть $\mid \overline{HPW} \mid = (s+1)!$ [Финн, 2019; Финн, 2020a].

Будем использовать метасимвол $\models$ для обозначения утверждения «φ истинно в HPW_h»: $HPW_h \models \varphi$, где φ – формула ERA_2, HPW_h – история возможных миров длины s, а $1 \leq h \leq (s+1)!$.

Оценка формул, не содержащих оператора T, реализует **когерентную** концепцию истины [Rescher, 1973; Вейнгартен, 2000], конструктивно реализуемую ДСМ-рассуждениями. Они формализуют **принятие** гипотез о причинах (посредством индукции) и гипотез о предсказаниях (посредством аналогии), а также формализуют абдукцию $1^{\underline{\text{ого}}}$ рода, которая завершает принятие порожденных гипотез посредством **объяснения** баз фактов.

Оценка формул, содержащих оператор T «истинно, что …» выражает акт **прямой верификации** гипотез о предсказаниях и акт **косвенной верификации** гипотез о причинах, которые когерентно истинны посредством интегральных каузальных вынуждений [Финн, 2019], реализуемых во всех HPW_h из $\overline{HPW}$. Эти оценки выражают корреспондентную концепцию истины [Tarski, 1956].

Базисом оценок корреспондентной истины является задание множества корреспондентно истинных элементарных формул $\mathbf{T}$, где $\mathbf{T}$ является средством семантики ERA_2, $ERA_2^{(1)}$ и $ERA_2^{(2)}$, но может быть добавлением соответствующих аксиом для $ERA_2^{(1)}$ и $ERA_2^{(2)}$.

Таким образом, $HPW_h \vDash Tq$, если и только если $q \in \mathbf{T}$.

Определение 2. Определение истинности формул в HPW_h

1°. $HPW_h \vDash p$, если и только если $Cd(p, h) = \nu \ldots \nu$ или $Cd(p, h) = \tau \ldots \tau\ \nu \ldots \nu$, где $\nu = 1, -1$ («1» и «-1» – типы истинностных значений «фактически истинно» и «фактически ложно», соответственно, а «τ» – тип истинностного значения «неопределенно»), $Cd(p, h)$ – код эмпирической закономерности, образованной последовательностями типов истинностных значений гипотез, порожденных ДСМ-рассуждениями в соответствующих базах фактов, h – номер истории возможных миров, $1 \leq h \leq (s+1)!$.

2°. $HPW_h \vDash \neg p$, если и только если неверно, что $HPW_h \vDash p$, то есть: $Cd(p, h) \neq \nu \ldots \nu$ и $Cd(p, h) \neq \tau \ldots \tau\nu \ldots \nu$; где $\nu \ldots \nu$ и $\tau \ldots \tau\nu \ldots \nu$ – регулярные коды ЭЗК [Финн, Шестерникова, 2018; Финн, 2019].

3°. $HPW_h \vDash Tp$, если и только если $p \in \mathbf{T}$.

4°. $HPW_h \vDash (\varphi \,\&\, \psi)$, если и только если $HPW_h \vDash \varphi$ и $HPW_h \vDash \psi$.

5°. $HPW_h \vDash (\varphi \vee \psi)$, если и только если $HPW_h \vDash \varphi$ или $HPW_h \vDash \psi$.

6°. $HPW_h \vDash (\varphi \to \psi)$, если и только если «если $HPW_h \vDash \varphi$, то $HPW_h \vDash \psi$».

7°. $HPW_h \vDash \Box\varphi$, если и только если $HPW_j \vDash \varphi$ для всех HPW_j, $HPW_j \in \overline{HPW}$.

8°. $HPW_h \vDash \Diamond\varphi$, если и только если существует HPW_h такая, что $Cd(\varphi, h) = \tau \ldots \tau\nu \ldots \nu$ и для всех HPW_j, $HPW_j \in \overline{HPW} Cd(\varphi, j) = \tau \ldots \tau\nu \ldots \nu$ или $Cd(\varphi, j) = \nu \ldots \nu$.

9°. $HPW_h \vDash \Box(\varphi \,\&\, \psi)$, если и только если $HPW_h \vDash \Box\varphi$ и и $HPW_h \vDash \Box\psi$, где $HPW_h \in \overline{HPW}$.

10°. $HPW_h \models \Box(\varphi \vee \psi)$, если и только если $HPW_h \models \Box\varphi$ или $HPW_h \models \Box\varphi$, $HPW_h \in \overline{HPW}$.

11°. $HPW_h \models \Diamond(\varphi \,\&\, \psi)$, если и только если $HPW_h \models \Diamond\varphi$ и $HPW_h \models \Diamond\psi, HPW_h \in \overline{HPW}$.

12°. $HPW_h \models \Diamond(\varphi \vee \psi)$, если и только если $HPW_h \models \Diamond\varphi$ или $HPW_h \models \Diamond\psi, HPW_h \in \overline{HPW}$.

13°. $HPW_h \models T\Box\varphi$, если и только если $HPW_h \models T\varphi$ и $HPW_h \models \Box\varphi, HPW_h \in \overline{HPW}$.

14°. $HPW_h \models \neg\varphi$, если и только если неверно, что $HPW_h \models \varphi$, $HPW_h \in \overline{HPW}$.

Аксиомы ERA_2 $(\Box\Diamond)_2$ $\Box\Diamond p \to \neg p$ и $(\Diamond\Box)_2$ $\Diamond\Box p \to \Diamond p$ сохраняют истинность относительно Определения 2 при применении R^*2.

Приведем некоторые теоремы ERA_2: $p \vee \neg p$, $(\Box p \vee \Diamond p \neg p)$, $(p \leftrightarrow (\Diamond p \vee \Box p))$, $(\neg\Box p \leftrightarrow (\Diamond p \vee \neg p))$, $(\neg\Diamond p \leftrightarrow (\Box p \vee \neg p))$, $((\Box(p \to q) \,\&\, Tq) \to \Box p)$, $((\Box(p \to q) \,\&\, Tq) \to Tp)$.

Доказательство $\neg\Box p \leftrightarrow (\Diamond p \vee \neg p)$: $\Box p \to \neg\Diamond p, \Diamond p \to \neg\Box p; \neg\Box p \to (\Diamond p \vee \neg p)\,(\neg\Box); \Box p \to p\,(\Box 2), \neg p \to \neg\Box p; \Diamond p \to \neg\Box p, \neg p \to \neg\Box p \vdash (\Diamond p \vee \neg p) \to \neg\Box p$, из $\neg\Box p \to (\Diamond p \vee \neg p)$ и $(\Diamond p \vee \neg p) \to \neg\Box p$ следует $\neg\Box p \leftrightarrow (\Diamond p \vee \neg p)$.

Добавим к исчислению ERA_2 в качестве допустимого правила вывода R7 теорему дедукции:[6]

$$\frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash (\varphi \to \psi)}.$$

Из аксиом T12 и T13 выведем производные правила вывода R8 $\Box(p \to q)$, $Tq \vdash \Box p$ и R9 $\Diamond(p \to q)$, $Tq \vdash \Diamond p$.

Тогда получим, применяя R8 и R7 $\Box(p \to q)$, $Tq \vdash \Box p$; и Tq, $\Box(p \to q)$, $\Box p \vdash \Box q$; и Tq, $\Box(p \to q) \vdash \Box p \to \Box q$; и $Tq \vdash \Box(p \to q) \to (\Box p \to \Box q)$.

Это означает, что из эмпирического предположения Tq выводима дистрибутивность $\Box$ относительно $\to$. Аналогично получим, используя R9 $Tq \vdash \Diamond(p \to q) \to (\Diamond p \to \Diamond q)$.

Если же добавить R7 к исчислениям $ERA_2^{(1)}$ и $ERA_2^{(2)}$, то для их аксиом Tq получим доказуемость дистрибутивности $\Box$ и $\Diamond$ относительно $\to$: $\vdash (\Box(p \to q) \to (\Box p \to \Box q))$, $\vdash (\Diamond(p \to q) \to (\Diamond p \to \Diamond q))$.

Исчисления с правилом R7 обозначим посредством ERA_{2^*}, $ERA_{2^*}^{(1)}$ и $ERA_{2^*}^{(2)}$.

Обратим внимание на интересный факт: дистрибутивность $\Box$ и $\Diamond$ относительно $\to$ в рассмотренных исчислениях связана с аксиомами абдукции T12, T13 и с эмпирическими аксиомами Tq, а, следовательно, она зависит

[6]Правило вывода называют допустимым, если его добавление к исчислению не порождает противоречий.

и от двух концепций истины – **когерентной** и **корреспондентной**, ибо логики типа ERA являются логиками двух концепций истины.

Рассмотрим теперь аксиомы ERA_2, выражающие итерации модальностей $(\Box\Box)$, $(\Box\Diamond)_2$, $(\Diamond\Box)_2$ и $(\Diamond\Diamond)$. Они имитируют пропозициональными средствами расширение баз фактов (возможных миров) для продолжения ДСМ-рассуждений в ДСМ-исследованиях, порождающих эмпирические закономерности относительно множества историй возможных миров $\overline{HPW}$ [Финн, 2019].

В силу ограничения правила подстановки R2 посредством правила R^*2 из $\Box\Box p \to \Box p$ не выводимы $\underbrace{\Box\ldots\Box}_{k} p \to \Box p$, из $\Box\Diamond p \to \neg p$ не выводимы $\underbrace{\Box\ldots\Box}_{k}\underbrace{\Diamond\ldots\Diamond}_{l} p \to \neg p$, из $\Diamond\Diamond p \to \neg p$ не выводимы $\underbrace{\Diamond\ldots\Diamond}_{k} p \to \neg p$, а из $\Diamond\Box p \to \Diamond p$ не выводимы $\Diamond\underbrace{\Box\ldots\Box}_{k} p \to \Diamond p$, где k и l – числа повторений модальных операторов.

Легко показать, что приведенные выше формулы с k и l итерациями модальностей истинны в семантике историй возможных миров.

В самом деле, рассмотрим $\Box\ldots\Box p \to \Box p$. Так как код ЭЗК для $\Box\ldots\Box p$ есть $\nu\ldots\nu\ldots\nu\ldots\nu$, то $HPW_h \vDash \Box\ldots\Box p$ и для всех HPW_j, $HPW_j \vDash \Box\ldots\Box p$, но и $HPW_j \vDash \Box p$ для всех HPW_j, так как код $\Box p$ есть $\nu\ldots\nu$, но и $\nu\ldots\nu\ldots\nu\ldots\nu = \nu\ldots\nu$.

Таким образом, $\Box\ldots\Box p \to \Box p$ истинно относительно $\overline{HPW}$. Откуда следует неполнота ERA_2, а также $ERA_2^{(1)}$ и $ERA_2^{(2)}$.

Аналогичные рассуждения имеют место для формул $\Box\ldots\Box\ \Diamond\ldots\Diamond \to \neg p$, $\Diamond\ldots\Diamond p \to \neg p$, $\Diamond\Box\ldots\Box p \to \Diamond p$, $\Box\ldots\Box p \to p$, соответствующей аксиоме $(\Box 2)$ $\Box p \to p$.

Сформулируем ниже нефинитное исчисление ERA_2 такое, что к его аксиомам добавим $\underbrace{\Box\ldots\Box}_{k} p \to p$, $\underbrace{\Box\ldots\Box}_{k} p \to \Box p$, $\underbrace{\Box\ldots\Box}_{k}\underbrace{\Diamond\ldots\Diamond}_{l} p \to \neg p$, $\underbrace{\Diamond\ldots\Diamond}_{k} p \to \neg p$, $\Diamond\underbrace{\Box\ldots\Box}_{k} p \to \Diamond p$ для любых целых положительных k и l.

Аналогичные исчисления получим для $ERA_2^{(1)}$ и $ERA_2^{(2)}$, а соответствующие исчисления обозначим посредством $\overline{ERA}_2$, $\overline{ERA}_2^{(1)}$ и $\overline{ERA}_2^{(2)}$. Сформулируем также исчисления $\overline{ERA}_{2^*}$, $\overline{ERA}_{2^*}^{(1)}$ и $\overline{ERA}_{2^*}^{(2)}$ с правилами вывода R1, R^*2, R3–R7.[7]

[7] В [Финн, 2019] логики аргументации, формализованные посредством метода аналитических таблиц, в связи с отсутствием ассоциативности у & и ∨ имеют также нефинитную формализацию.

Замечание 3. В [Финн, 2019] и [Финн, 2020a] решением проблемы индукции средствами ДСМ-метода АПИ для интеллектуальных систем является порождение М-последовательностей модальных операторов ранга r, где r – число периодов длины s, а $r \geq 1$. М-последовательности представляют возможные типы эмпирических закономерностей, являющиеся результатами ДСМ-исследований, применяющих формализованные и усиленные индуктивные каноны Д.С. Милля в качестве правил вывода, порождающих гипотезы о причинах. Нефинитные расширения ERA_2-логик являются упрощенной попыткой пропозициональной имитации рассуждений относительно историй расширяемых возможных миров (баз фактов)[8].

Следствием Замечания 3 является потребность в расширениях ERA_2-логик посредством оператора слабой возможности [Финн, 2019] и оператора N такого, что Np фиксирует существование незакономерности, тогда как $\neg p$ есть отрицание закономерности, а потому $NNp \to Np$, но $\neg\neg p \leftrightarrow p$, а $Np \to \neg p$. Если M-последовательность $\overline{M} = M_1 M_2 \ldots M_{r-1} N$, то ДСМ-исследование ранга r не является закономерностью для периодов повторения ДСМ-рассуждений r раз.

Кроме того, ERA_2-логики могут быть расширены добавлением оператора слабой возможности $\triangledown$ [Финн, Шестерникова, 2018].

Интересно рассмотреть в связи со сказанным нефинитные ERA_2-логики с операторами $\triangledown$ и N.

Так как аксиома $S4$ $\Box p \to \Box\Box p$ истинна в ERA_2, то возможна логика $ERA_{2.4}$ с аксиомой $\Box p \to \Box\Box p$.

Заметим, что аксиома $S5$ $\Diamond p \to \Box\Diamond p$ [Фейс, 1965; Hughes, Cresswell, 1972] не является истинной в ERA, но она истинна в логике ERA_1 [Финн, 2019], поэтому возможно её расширение $ERA_{1.5}$.

Замечание 4. Логика M Г.Х. фон Вригта [Фейс, 1965; Hughes, Cresswell, 1972] образована аксиомами:
M1. $\Box p \to p$
M2. $\Box(p \to q) \to (\Box p \to \Box q)$
и правилами:
RM1. $\varphi, \varphi \to \psi \vdash \psi$
RM2. $\varphi(p) \vdash \varphi(\chi), \varphi(\chi) = \int_p^\chi \varphi(p) \mid$
RM3. $\varphi \vdash \Box\varphi$

Следовательно, в M доказуема любая формула $\Box\varphi$ такая, что φ – тавтология двузначной логики L_2.

[8] Упрощенность обусловлена тем, что рассматриваются не все модальности из [Финн, Шестерникова, 2018; Финн, 2019], а только $\Box$ и $\Diamond$, тогда как операторов $\Box$ имеется восемь, операторов $\Diamond$ имеется четыре, а операторов слабой возможности $\triangledown$ имеется два.

В этом смысле можно говорить, что логика M и её расширения $S4$ и $S5$ являются логиками **логических** модальностей.

Следующие сходства и различия логики M и ERA_2-логик имеют место.

1. В ERA_2-логиках (и ERA_1-логиках [Финн, 2019]) имеется аксиома $\Box p \to p$.

2. В ERA_2 из предположения Tq выводима $\Box(p \to q) \to (\Box p \to \Box q)$: $Tq \vdash \Box(p \to q) \to (\Box p \to \Box q)$.

3. В $ERA_2^{(1)}$ и $ERA_2^{(2)}$, имеющих «эмпирические» аксиомы Tq доказуема $\Box(p \to q) \to (\Box p \to \Box q)$.

4. В возможных мирах логики M не является истинной константа f [Chellas, 1980], а, следовательно, истинна константа t, представляющая тавтологию L_2. Однако истинность и ложность f и t в историях возможных миров HPW_h **не определима**, так как выполнимость в HPW_h предполагает конструктивные вынуждения (forcing) посредством правил вывода ДСМ-рассуждений (индукции и аналогии).

5. Правило вывода RM3 $\varphi \vdash \Box\varphi$, где φ доказуема в двузначной логике L_2, в ERA_i-логиках $(i = 1, 2)$ не имеет места.

Дело в том, что введение $\Box$ возможно только в силу двух правил вывода:

R3 $\Box\varphi, \Box(\varphi \to \psi) \vdash \Box\psi$ и производного правила вывода $Tq, \Box(p \to q) \vdash \Box p$. Следовательно, должна быть для введения $\Box$ уже доказанная формула вида $\Box(\varphi \to \psi)$.

Невыполнимость t (тавтологий L_2) в HPW_h согласуется с неприменимостью правила RM3 в ERA-логиках.

6. В M имеет место $\Diamond p \leftrightarrow \neg\Box\neg p$, то есть, выразимость $\Diamond$ через $\Box$ и $\neg$, тогда как в ERA-логиках модальные операторы $\Box$ и $\Diamond$ независимы, ибо $\Box$ характеризует эмпирические законы (их коды – $\nu \ldots \nu$), а $\Diamond$ – эмпирические тенденции (их коды – $\tau \ldots \tau\ \nu \ldots \nu$).

7. Существенным отличием ERA-логик от M, $S4$ и $S5$ является применение двух концепций истины – когерентной и корреспондентной, наличие в ERA-логиках в связи с этим оператора T и аксиом абдукции T12, T13.

Существенно также, что в $ERA_2^{(1)}$ и $ERA_2^{(2)}$ имеются эмпирические аксиомы Tq, имитирующие использование эмпирических (экспериментальных) данных в ДСМ-рассуждениях.

Рассмотренные особенности ERA_2-логик характеризуют их как логик **эмпирических** (нелогических) модальностей, что согласуется с идеей Р. Фейса о связи модальностей и причинности [Фейс, 1965], ибо источником

ERA_2-логик (ERA_1-логик [Финн, 2019]) являются ДСМ-исследования, которые образованы рассуждениями, порождающими гипотезы о причинах эффектов и гипотезы о предсказаниях этих эффектов (с использованием гипотез о причинах и абдукции для принятия порожденных гипотез).

В [Reichenbach, 1947; Reichenbach, 1954] Г. Рейхенбах определил «физические модальности» следующим образом: пусть «p» – имя высказывания p, тогда p – физически необходимо $\rightleftharpoons$ если «p» есть номологическое высказывание,

> p – физически невозможно $\rightleftharpoons$ если «$\neg p$» есть номологическое высказывание,
>
> p – физически возможно $\rightleftharpoons$ если ни «p», ни «$\neg p$» не являются номологическими высказываниями.

Номологические высказывания у Г. Рейхенбаха определяются посредством специальных условий, выразимых в языке логики предикатов $1^{\underline{\text{ого}}}$ порядка. Таким образом, возможность у Г. Рейхенбаха не представляет эмпирические закономерности. Тогда как в ДСМ-методе АПИ возможность представляет эмпирические тенденции, а эмпирические законы и эмпирические тенденции в ERA-логиках характеризуются следующими аксиомами: $\neg\Box p \to (\Diamond p \vee \neg p)$, $\neg\Diamond p \to (\Box p \vee \neg p)$, которые выражают различие и независимость модальностей $\Box$ и $\Diamond$, характеризующих эмпирические законы и эмпирические тенденции, соответственно.

Таким образом, ERA-логики являются логиками, порожденными ДСМ-рассуждениями, а так как ДСМ-рассуждения являются логическими средствами ДСМ-исследований, порождающими эмпирические закономерности (эмпирические номологические высказывания [Финн, 2019]), то модальности, соответствующие обнаруженным закономерностям, являются **эмпирическими**, отличными от логических модальностей известных модальных логик. Специфические свойства эмпирических модальностей обусловлены их происхождением от средств интеллектуального анализа данных, итогом которого являются эмпирические номологические высказывания, определяющие эмпирические модальности.

Как уже было сказано выше, эпистемологической особенностью ERA-логик является применение двух концепций истины – когерентной (она обусловлена ДСМ-рассуждениями) и корреспондентной (она обусловлена верификацией гипотез о предсказаниях), что вызвало необходимость применения оператора T в ERA-логиках.

Наличие двух типов истинностных оценок породило вопрос: будут ли непротиворечивы (противоречивы) ERA-логики, дополненные условием

$\neg Tp \,\&\, p$?[9] Этот вопрос вызван тем обстоятельством, что когерентные истинностные значения и корреспондентные истинностные значения **независимы**, так как первые порождены ДСМ-рассуждениями, а вторые – верификациями их результатов [Финн, 2019; Финн, 2020a]. Поэтому когерентные истинностные значения можно называть «**внутренними**», а корреспондентные истинностные значения – «**внешними**», используя терминологию Д.А. Бочвара [Бочвар, 1938].

Замечание 5. Система аксиом ERA_2 является **зависимой**, так как доказуемы аксиомы ($\Box 2$), ($\Diamond 2$), ($\Box\neg\Diamond$).

Если добавить к ERA_2 в качестве аксиомы ($\Box p \vee \Diamond p \vee \neg p$), то будут доказуемы ($\neg\Box$) и ($\neg\Diamond$).

Таким образом, получаем экономную формулировку ERA_2, устранив ($\Box 2$), ($\Diamond 2$), ($\Box\neg\Diamond$), ($\neg\Box$) и ($\neg\Diamond$), добавив аксиому ($\Box p \vee \Diamond p \vee \neg p$): любое высказывание p является необходимым, или возможным, или не представляет эмпирическую закономерность.

Литература

Аншаков, 2009 – ДСМ-метод автоматического порождения гипотез. Логические и эпистемологические основания / Под общ. ред. О.М. Аншакова. М.: Книжный дом «Либроком», 2009.

Бочвар, 1938 – *Бочвар Д.А.* Об одном трехзначном исчислении и его применении к анализу парадоксов классического расширенного функционального исчисления // Математический сборник. 1938. Т. 4. Вып. 46. № 2. С. 287–308.

Вейнгартен, 2000 – *Вейнгартен П.* Фундаментальные проблемы истины. М.: Росспэн, 2005. (Weingartner P. Basic question on truth. Dordrecht / Boston / London: Kluwer Academic Publishers, 2000.)

фон Вригт, 1971 – *Вригт фон Г.Х.* Логико-философские исследования. М.: Прогресс, 1986. (Von Wright G.H. Explanation and Understanding. London, 1971.)

Фейс, 1965 – *Фейс Р.* Модальная логика. М.: Наука, 1974; (Feys R. Modal Logics. Louvain / Paris: E. Nauwelaerts / Gauthier – Villars Publishers, 1965.)

Финн, 2006 – *Финн В.К.* Стандартные и нестандартные логики аргументации // Логические исследования. Вып. 13. М.: Наука, 2006. С. 158–189.

Финн, 2014 – Финн В.К. Дистрибутивные решетки индуктивных ДСМ-процедур // Научно-техническая информация. 2014. Сер. 2. № 11. С. 1–30. (Finn V.K. Distributive Lattices of Inductive JSM Procedures. Automatic Documentation and Mathematical Linguistics, Vol. 48, pp. 265–295.)

Финн, 2016 – *Финн В.К.* О классе ДСМ-рассуждений, использующих изоморфизм правил индуктивного вывода // Искусственный интеллект и принятие решений. 2016. № 3. С. 95–108.

[9] Напомним, что $p \leftrightarrow (\Box p \vee \Diamond p)$ соответствует когерентному истинностному значению, а Tp – корреспондентному.

Финн, 2018 – *Финн В.К.* Искусственный интеллект: методология, применения, философия. М.: КРАСАНД, 2018.

Финн, Шестерникова, 2018 – *Финн В.К., Шестерникова О.П.* Эвристика обнаружения эмпирических закономерностей посредством ДСМ-рассуждений // Научно-техническая информация. 2018. Серия 2. № 9. С. 7–42.

Финн, 2019 – *Финн В.К.* Об эвристиках ДСМ-исследований (дополнения к статьям) // Научно-техническая информация. 2019. Серия 2. № 19. С. 1–34.

Финн, 2020a – *Финн В.К.* Точная эпистемология и искусственный интеллект // Научно-техническая информация. 2020. Серия 2. (В печати)

Финн, 2020b – Автоматическое порождение гипотез в интеллектуальных системах / Под общ. ред. *В.К. Финна.* М.: Книжный дом «Либроком», 2020.

Chellas, 1980 – *Chellas B.F.* Modal Logic. An Introduction. Cambridge: Cambridge University Press, 1980.

Hughes, Cresswell, 1972 – *Hughes G.E., Cresswell M.J.* An Introduction to Modal Logic. London EC4: Methuen and Co. LTD. 11 New Fetter Lanc, 1972.

Łukasiewicz, 1920 – *Łukasiewicz J.* O logice trójwartościowej // Ruch Filozoficzny. 1920. Vol. V. P. 169–171.

Reichenbach, 1947 – *Reichenbach H.* Elements of Symbolic Logic. The Macmillan Company, New York, 1947.

Reichenbach, 1954 – *Reichenbach H.* Nomological Statements and Admissible Operations. Amsterdam: North-Holland Publishing Co., 1954.

Rescher, 1973 – *Rescher N.* The coherence theory of truth. Oxford: The Clarendon Press, 1973.

Tarski, 1956 – *Tarski A.* The concept of truth in formalized languages // A Tarski Logic, Semantics, Metamathematics. Oxford: At the Clarendon Press, 1956. P. 152–278.

Victor K. Finn

On the logics of empirical modalities

Victor K. Finn
Federal Research Center "Computer Science and Control", Russian Academy of Sciences,
44/2 Vavilova Str., Moscow, 119333, Russian Federation.
Russian State University for the Humanities,
6 Miusskaya Sq., Moscow, 125993, Russian Federation.
E-mail: ira.finn@gmail.com

Abstract: The article considers a class of ERA-logics with the empirical modalities $\Box$ (necessity) and $\Diamond$ (possibility), which characterize, respectively, statements representing empirical laws and empirical tendencies, i.e., empirical regularities. Empirical regularities are the result of JSM reasoning, which is formed by the interaction of the inductive inference rules and inference rules by analogy, as well as the procedures for abductive acceptance of hypotheses. The ERA-logics under consideration are propositional imitation of JSM reasoning applicable to sequences of extensible fact bases of intelligent systems. A characteristic feature of ERA-logics is the application of two concepts of truth — coherent and correspondent. The application of the coherent concept of truth is due to the generation of hypotheses through the rules of inductive inference and inference by analogy. The application of the correspondent concept of truth is due to the use of an abductive inference, the acceptance of the results of which uses verification of predictions hypotheses. For this purpose, ERA-logics use the operator T: "it is true that…". In the conclusion of the article, non-finite extensions of ERA-logics are discussed, as well as their differences as logics of empirical modalities from G.H. von Wright' M-logic of logical modalities.

Keywords: JSM-reasoning, inductive inference rules, inference rules by analogy, abduction, empirical regularities, empirical law, empirical tendency, empirical modalities, logical modalities, operator "it is true that…", nomological statements

For citation: Finn V.K. "O logikakh empiricheskikh modal'nostei" [On the logics of empirical modalities], *Logicheskie Issledovaniya / Logical Investigations*, 2020, Vol. 26, No. 1, pp. 124–143. DOI: 10.21146/2074-1472-2020-26-1-124-143 (In Russian)

References

Anshakov, 2009 – *DSM-metod avtomaticheskogo porozhdeniya gipotez. Logicheskie i epistemologicheskie osnovaniya* [JSM Method of Automatic Hypotheses Generation: Logical and Epistemological Foundation], ed. by Anshakov O.M. Moscow: Knizhnyi dom "Librokom", 2009. (In Russian)

Bochvar, 1938 – Bochvar, D.A. "Ob odnom trekhznachnom ischislenii i ego primenenii k analizu paradoksov klassicheskogo rasshirennogo funktsional'nogo ischisleniya"

[On a Three-Valued Logical Calculus and its Application to the Analysis of Contradictions], in: *Matematicheskii sbornik*, 1938, Vol. 4, No. 2, pp. 287–308. (In Russian)

Chellas, 1980 – Chellas, B.F. *Modal Logic. An Introduction.* Cambridge: Cambridge University Press, 1980.

Feys, 1965 – Feys, R. *Modal Logics.* Louvain / Paris: E. Nauwelaerts / Gauthier – Villars Publishers, 1965.

Finn, 2006 – Finn, V.K. "Standartnye i nestandartnye logiki argumentatsii" [Standard and non-standard argumentation logics], *Logical Investigations*, 2006, Vol. 13, pp. 158–189. (In Russian)

Finn, 2014 – Finn, V.K. "Distributive Lattices of Inductive JSM Procedures", *Automatic Documentation and Mathematical Linguistics*, 2014, Vol. 48, pp. 265–295.

Finn, 2016 – Finn, V.K. "O klasse DSM-rassuzhdenii, ispol'zuyushchikh izomorfizm pravil induktivnogo vyvoda" [On the Class of JSM reasoning, using the isomorphism of inductive inference rules], *Artificial Intelligence and Decision Making*, 2016, No. 3, pp. 95–108. (In Russian)

Finn, 2018 – Finn, V.K. *Iskusstvennyi intellect: Metodologiya, primeneniya, filosofiya* [Artificial Intelligence: Methodology, Application, Philosophy], Moscow: KRASAND, 2011. (In Russian)

Finn, Shesternikova, 2018 – Finn, V.K., Shesternikova, O.P. "Evristika obnaruzheniya empiricheskikh zakonomernostei posredstvom DSM-rassuzhdenii" [The Heuristics of Empirical Regularities discovering by JSM Reasoning] , *Scientific and Technical Information Processing*, 2018, Ser. 2, No. 9, pp. 7–42. (In Russian)

Finn, 2019 – Finn, V.K. "Ob evristikakh DSM-issledovanii (dopolneniya k stat'yam)" [On JSM researches heuristics (addittions th the articles)], *Scientific and Technical Information Processing*, 2019, Ser. 2, No. 19, pp. 1–34. (In Russian)

Finn, 2020a – Finn, V.K. "Tochnaya epistemologiya i iskusstvennyi intellekt" [Precise epistemology and artificial intelligence], *Scientific and Technical Information Processing*, 2020, Ser. 2. (In print)

Finn, 2020b – *Avtomaticheskoe porozhdenie gipotez v intellectual'nykh sistemakh* [Automatic Hypotheses Generation in Intelligent Systems], ed. by Finn V.K. Moscow: Knizhnyi dom "Librokom", 2020. (In Russian)

Hughes, Cresswell, 1972 – Hughes, G.E., Cresswell, M.J. *An Introduction to Modal Logic.* London EC4: Methuen and Co. LTD. 11 New Fetter Lanc, 1972.

Łukasiewicz, 1920 – Łukasiewicz, J. "O logice trójwartościowej", *Ruch Filozoficzny*, 1920, Vol. V, pp. 169–171.

Reichenbach, 1947 – Reichenbach, H. *Elements of Symbolic Logic.* The Macmillan Company, New York, 1947.

Reichenbach, 1954 – Reichenbach, H. *Nomological Statements and Admissible Operations.* Amsterdam: North-Holland Publishing Co., 1954.

Weingartner, 2000 – Weingartner, P. *Basic question on truth.* Dordrecht / Boston / London: Kluwer Academic Publishers, 2000.

Von Wright, 1971 – Von Wright, G.H. *Explanation and Understanding.* London, 1971.

Rescher, 1973 – Rescher, N. *The coherence theory of truth*. Oxford: The Clarendon Press, 1973.

Tarski, 1956 – Tarski, A. "The concept of truth in formalized languages", in: *A Tarski Logic, Semantics, Metamathematics*, Oxford: At the Clarendon Press, 1956, pp. 152–278.

Список трудов А.С. Карпенко

Alexander S. Karpenko: list of publications

Монографии

1. Фатализм и случайность будущего: Логический анализ. М.: Наука, 1990. 214 с. 2-е исправленное издание — М.: ЛКИ/URSS, 2008. (Рецензия Y.V. Shramko в: From the Logical Point of View. 1992. Vol. 2. P. 102–103).

2. Многозначные логики. Серия «Логика и компьютер». Вып. 4. М.: Наука, 1997. 223 с. (Рецензия A. Adamatzky в: Multiple-Valued Logic. 2000. Vol 5. P. 81–86).

3. Логики Лукасевича и простые числа. М.: Наука, 2000. 319 с. 2-е исправленное издание — М.: ЛКИ/URSS, 2007. 3-е издание — М.: ЛКИ/URSS, 2009. (Рецензия A. Adamatzky в: Journal of Multiple-Valued Logic and Soft Computing. 2004. Vol. 10. № 3. P. 309–314).

4. Łukasiewicz's Logics and Prime Numbers. Luniver Press, 2006. 168 p.

5. Развитие многозначной логики. М.: ЛКИ, 2010. 448 с. (стереотипное изд. в 2014 г.). (Рецензия А. Шумана: Вопросы философии. 2012. № 3. Р. 184–185).

6. Логика и метафизика. Сборник статей. М.-СПб.: ЦГИ, 2017. 432 с.

Статьи в журналах и сборниках

1. Проблема высказываний о будущих событиях в трехзначной логике // Проблемы логики, методологии и истории науки. Вып. 2. М.: Изд-во МГУ, 1978.

2. Актуальные проблемы современной логики // Вопросы философии. 1979. № 4. С. 162–164. (в соавторстве с В.К. Финном)

3. Матричная логика для простых чисел и ее обобщение // Модальные и релевантные логики (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 1). М., 1982. С. 81–90.

4. Aristotle, Łukasiewicz and factor-semantics // Acta Philosophica Fennica. Vol 25. 1982. P. 7–21.

5. Многозначная логика: Применение, развитие, проблемы (Обзор) // Материалы к VII Международному конгрессу по логике, методологии и философии науки. М.: ИНИОН, 1983. С. 37–69.

6. Последовательность предполных логик Лукасевича и графы для простых чисел // Логические исследования (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 2). М., 1983. С. 103–111.

7. Factor-semantics for n-valued logics // Studia Logica. Vol 42. No. 2/3. 1983. P. 179–185.

8. Аристотель, Лукасевич и фактор-семантика // В.А. Смирнов и др. (ред.). Модальные и интенсиональные логики и их применение к проблемам методологии науки. М.: Наука, 1984. С. 107–124.

9. Фактор-семантика для n-значных логик // Смирнов В.А. и др. (ред.). Модальные и интенсиональные логики и их применение к проблемам методологии науки. М.: Наука, 1984. С. 124–140.

10. Паранепротиворечивая структура внутри многозначной логики // Многозначные, релевантные и паранепортиворечивые логики (Труды научно-исследовательского семинара по логике ИФАН. Вып. 3). М., 1984. С. 34–39.

11. Функциональная предполнота и квази-штрих Тюркетта // Многозначные, релевантные и паранепортиворечивые логики (Труды научно-исследовательского семинара по логике ИФАН. Вып. 3). М., 1984. С. 65–69.

12. Characterization of prime numbers by Łukasiewicz's many-valued logic // Bulletin of the Section of Logic. 1984. Vol. 13. № 2. С. 64–68.

13. Фактор-семантика для бесконечнозначной логики Лукасевича // Неклассические логики (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 4). М., 1985. С. 20–26.

14. Paraconsistent structure inside many-valued logics // Synthese. 1986. Vol. 66. № 1. P. 63–69.

15. Интенсиональная логика и логическая структура теорий // Философские науки. 1986. № 6. (в соавторстве с Л.И. Мчедлишвили)

16. Некоторые алгебры в качестве истинностных значений // Нестандартные семантики неклассических логик (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 5). М., 1986. С. 6–13.

17. A hypothesis on the finiteness of graphs for Łukasiewicz's precomplete logics (graphs for prime numbers) // Bulletin of the Section of Logic. 1986. Vol. 15. № 3. P. 102–108.

18. Новый взгляд на логику пропозициональных установок // Неклассические логики и пропозициональные установки (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 6). М., 1987. С. 13–17.

19. Симметрический моноид Гейтинга: модель для бесконечнозначной логики Лукасевича // Неклассические логики и пропозициональные установки (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 6). М., 1987. С. 118–124. (в соавторстве с В.Л. Васюковым)

20. Algebraic structure of the truth-values for $Ł_\omega$ // Bulletin of the Section of Logic. 1988. Vol. 17. № 3/4. P. 127–137.

21. От фактор-семантики к исчислению алгебр // В.А. Смирнов (ред.) Синтаксические и семантические исследования неэкстенсиональных логик. М.: Наука, 1989. С. 93–100.

22. Конечнозначные логики Лукасевича: алгебро-логические свойства простых чисел // В.А. Смирнов (ред.) Синтаксические и семантические исследования неэкстенсиональных логик. М.: Наука, 1989. С. 170–205.

23. О паранепротиворечивой логике // В.А. Смирнов (ред.) Синтаксические и семантические исследования неэкстенсиональных логик. М.: Наука, 1989. С. 261–284. (в соавторстве с А.Т. Ишмуратовым и В.М. Поповым)

24. Максимальная паранепротиворечивая логика и дважды p-алгебры // Неклассические логики и их применения (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 7). М., 1989. С. 69–73.

25. Characterization of prime numbers in Łukasiewicz's logical matrix // Studia Logica. 1989. Vol. 48. № 4. P. 465–478.

26. Логический фатализм и тоталитаризм // Философские основания неклассических логик (Труды научно-исследовательского семинара по логике Института философии АН СССР. Вып. 8). М., 1990. С. 138–151.

27. Две решетки импликативных логик // Вестник Санкт-Петербургского Университета. Серия 6. 1992. Вып. 4. С. 28–29.

28. Логический фатализм // Е.А. Мамчур (ред.) Проблемы методологии постнеклассической науки. М.: ИФАН, 1992. С. 178–186.

29. The problem of extending $H_{\rightarrow}$ to $TV_{\rightarrow}$ // IGPL Newsletter. 1992. Vol. 1. № 2. P. 5–7.

30. Lattices of implicational logics // Bulletin of the Section of Logic. 1992. Vol. 21. № 3. P. 82–91.

31. Матричная логика без неподвижных точек // Логические исследования. Вып. 1. М.: Наука, 1993. С. 181–185.

32. Ян Лукасевич — детерминизм и логика // Логические исследования. Вып. 2. М.: Наука, 1993. С. 206–223.

33. Импликативные логики: Решетки и конструкции // Логические исследования. Вып. 2. М.: Наука, 1993. С. 224–258.

34. Construction of classical propositional logic // Bulletin of the Section of Logic. 1993. Vol. 22. № 3. P. 92–97.

35. Штрих Шеффера для простых чисел // Труды научно-исследовательского семинара логического центра Института философии РАН. 1993. М., 1994. С. 102–106.

36. Матрицы для независимости аксиомы транзитивности в аксиоматизации классической импликации // Труды нучно-исследовательского семинара логического центра Института философии РАН. 1993. М., 1994. С. 107–109. (В соавторстве с С.А. Павловым)

37. Sheffer's stroke for prime numbers // Bulletin of the Section of Logic. 1994. Vol. 23. № 3. P. 126–129.

38. Логика, детерминизм и феномен прошлого (к публикации статьи Яна Лукасевича «О детерминизме») // Вопросы философии. 1995. № 5. С. 72–81.

39. Штрих Шеффера для простых чисел // Логические исследования. Вып. 3. М.: Наука, 1995. С. 292–313.

40. A logic without fixed points // P.I. Bystrov, V. Sadovsky (eds.) Philosophical logic and logical philosophy. Dordrecht: Kluwer, 1996. P. 213–219.

41. The class of precomplete Łukasiewicz's many-valued logics and the law of prime number generation // Bulletin of the Section of Logic. 1996. Vol. 25. № 1. P. 52–57.

42. Классификация пропозициональных логик // Логические исследования. Вып. 4. М.: Наука, 1997. С. 107–133.

43. Учитель // Логические исследования. Вып. 4. М.: Наука, 1997. С. 20–24.

44. Минимальные модели для нечеткой алгебры типа 2 // Труды научно-исследовательского семинара логического центра Института философии РАН. 1996. М.: ИФРАН, 1997. С. 56–63. (В соавторстве с В.И. Шалаком)

45. Новая аксиоматизация импликативного фрагмента бесконечнозначной логики Лукасевича $Ł_\omega$ // Труды научно-исследовательского семинара логического центра Института философии РАН. 1996. М.: ИФРАН, 1997. С. 71–75. (В соавторстве с В.М. Поповым)

46. Библиотечно-библиографическая классификация литературы по логике // Труды научно-исследовательского семинара логического центра Института философии РАН. 1996. М., 1997. С. 201–205.

47. *BCKX* is the axiomatization of implicational fragment of Łukasiewicz's infinite-valued logic $Ł_\omega$ // Bulletin of the Section of Logic. 1997. Vol. 26. № 3. P. 112–117. (with V.M. Popov)

48. Certain logical ideas of V.A. Smirnov and modern logic // Modern Logic. 1997. Vol. 7. № 3/4. P. 282–295.

49. Некоторые логические идеи В.А. Смирнова // Вопросы философии. 1998. № 2. С. 164–172. Переиздано в: Логические исследования. Вып. 5. 1998; Логическое кантоведение-4. Калининград, 1998; Философия не кончается... Из истории отечественной философии. XX век. 1960–80-е годы. М.: РОССПЭН, 1998; Смирнов В.А. Теория логического вывода. М.: РОССПЭН, 1999.

50. Научная деятельность В.А. Смирнова (1931–1996) // Философия науки. Вып. 4. М.: ИФРАН, 1998. С. 5–9.

51. Решение проблемы В.К. Финна о погружении степени простого числа в многозначные логики Лукасевича // Труды научно-исследовательского семинара Логического центра Института философии РАН. 1997. М., 1998. С. 35–40.

52. Note on the V. Smirnov's scientific activity: work and life (1931–1996) // Bulletin of the Section of Logic. 1998. Vol. 27. № 1/2. P. 4–7.

53. A maximal lattice of implicational logics // Bulletin of the Section of Logic. 1998. Vol. 27. № 1/2. P. 29–32.

54. Jaśkowski's criterion and three-valued paraconsistent logics // Logic and Logical Philosophy. 1999. № 7. P. 81–86.

55. Логика в России. Вторая половина XX-го века // Вопросы философии. 1999. № 9. С. 148–158.

56. Импликации следования, строгая, релевантная, интуиционистская и классическая и их взаимоотношения // Логические исследования. Вып. 6. М.: Наука, 1999. С. 76–80.

57. В.А. Смирнов как логик и методолог науки (Обзор) // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Серия 3: Философия. Реферативный журнал. № 1. М.: ИНИОН, 1999. С. 58–78. (В соавторстве с А.М. Анисовым)

58. Характеризация классов натуральных чисел посредством логических матриц // Труды научно-исследовательского семинара Логического центра Института философии РАН. 1998. М., 1999. С. 217–225.

59. Логика на рубеже тысячелетий // Логические исследования. Вып. 7. М.: Наука, 2000. С. 7–60.

60. Рецензия: А.А. Крушинский. Логика «И цзина»: Дедукция в Древнем Китае // Восток. 2000. № 5. С. 196–198.

61. A maximal paraconsistent logic: The combination of two three-valued isomorphs of classical propositional logic // D. Batens et al. (eds.) Frontiers of Paraconsistent Logic. Baldock: RSP, 2000. P. 181–187.

62. Three-valued paraconsistent logics // Multiple-Valued Logic. 2000. Vol. 5. No. 2. P. 117–123.

63. V.A. Smirnov (1931–1996) // Studia Logica. 2000. Vol. 66. № 2. P. 201–204.

64. The Classification of propositional calculi // Studia Logica. 2000. Vol. 66. № 2. P. 253–271.

65. Владимир Александрович Смирнов: вехи биографии и научная деятельность // В.И. Шалак (ред.) Логико-философские труды В.А. Смирнова. М.: УРСС, 2001. С. 9–13.

66. В.А. Смирнов о модально-временных логиках // В.И. Шалак (ред.) Логико-философские труды В.А. Смирнова. М.: УРСС, 2001. С. 211–214.

67. Подструктурные логики: гильбертовский подход // Логические исследования. Вып. 8. М.: Наука, 2001. С. 65–71.

68. Не истинностно-функциональная логика Клини с антибулевой операцией // Труды научно-исследовательского семинара Логического центра Института философии РАН. Вып. XV. М., 2001. С. 41–45.

69. Нерегулярность и «существенная» немонотонность логики Юрьева Y3 // Труды научно-исследовательского семинара логического центра Института философии РАН. Вып. XVI. М., 2002. С. 54–58.

70. Atomic and molecular paraconsistent logics // Logique et Analyse. 2002. Vol. 45. № 177–178. P. 31–37.

71. Современные исследования в философской логике // Вопросы философии. 2003. № 9. С. 54–75. (Переиздано в: Философия. Наука. Культура. «Вопросам философии» 60 лет (под редакцией академика В.А. Лекторского). М.: Вече, 2008. С. 725–742).

72. Рецензия на книгу В.К. Финна «Интеллектуальные системы и общество» // Вопросы философии. 2003. № 10. С. 174–183.

73. Современное состояние исследований в философской логике // Логические исследования. Вып. 10. М.: Наука, 2003. С. 61–93.

74. Современные исследования в философской логике: Мировой уровень, отечественная наука, РФФИ // Вестник РФФИ. 2003. № 3. С. 52–71.

75. Дуал трехзначной логики Гейтинга // Труды научно-исследовательского семинара Логического центра Института философии РАН. Вып. XVII. М., 2004. С. 68–71.

76. Предмет логики в свете основных тенденций ее развития // Логические исследования. Вып. 11. М., 2004. С. 149–171.

77. Предмет логики в свете основных тенденций ее развития // Вестник Санкт-Петербургского Университета. Серия 6. 2004. Вып. 3. С. 4–13.

78. Неклассические логики *versus* классической // Логико-философские штудии. 2005. Вып. 3. С. 48–73.

79. Логика: некоторые аспекты современного развития // Вестник Российской академии естественных наук. 2006. № 1. С. 69–73.

80. Применение логики в компьютерных науках // А.Е. Конверский (ред.) Логика: Перспективы развития. Киев: ИПЦ «Киевский университет», 2006. С. 35–41.

81. К определению логического: Традиция и современный взгляд // Вестник Вятского государственного гуманитарного университета. 2007. № 4. С. 12–19.

82. Трехзначные характеристические матрицы для классической пропозициональной логики // Труды научно-исследовательского семинара Логического центра Института философии РАН. Вып. XVIII. М.: ИФ РАН, 2007. С. 50–62. (В соавторстве с Л.Ю. Девяткиным и В.М. Поповым).

83. Modern study in philosophical logic: worldwide level and Russian science // Studies in Logic, Grammar and Rhetoric. 2008. Vol. 14. № 27. P. 35–71.

84. Эндоморфизмы в качестве операторов истинности и модальная логика **V2** // Труды научно-исследовательского семинара логического центра Института философии РАН. Вып. XIX. М.: ИФРАН, 2009. С. 54–63.

85. Континуальность трехзначных логик: проблемы и гипотезы // Логические исследования. 2010. Вып. 16. С. 127–133.

86. На пути к протологике // Логические исследования. 2011. Вып. 17. С. 152–166.

87. Аристотель и Лукасевич о законе противоречия (К публикации книги Яна Лукасевича «О принципе противоречия у Аристотеля. Критическое исследование») // Вопросы философии. 2012. № 8. С. 154–165.

88. Ян Лукасевич против Яна Лукасевича (вступительная статья) // Лукасевич Ян. О принципе противоречия у Аристотеля. Критический анализ / Пер. с польск. М.-СПб.: ЦГИ, 2012. С. 19–50.

89. Moscow logical schools (Period of ideology 1917–1991) // A. Schumann (ed.) Logic in Central and Eastern Europe: History, Science and Discourse. University Press of America, 2013. P. 317–380.

90. Preface. The History of Finish-Soviet Logic Colloquium // Logical Investigations (Special Issue). 2013. Vol. 19. P. 5–9.

91. Von Wright's truth logic and around // Logical Investigations (Special Issue). 2013. Vol. 19. P. 39–50.

92. Философский принцип полноты. Ч. I, II // Вопросы философии. 2013. № 6. С. 58–70; № 7. С. 95–108.

93. Философский принцип полноты. Ч. I, II // Философия и культура. 2013. № 11. С. 1508–1522; № 12. С. 1660–1679 (доп. и перераб.)

94. Philosophical problems of foundations of logic // Studia Humana. 2014. Vol. 3. № 1. P. 13–26.

95. Many-valued logic in Poland: The golden age // K. Mulligan et al. (eds.) The History and Philosophy of Polish Logic. Essays in Honour of Jan Woleński. Palgrave Macmillan, 2014. P. 13–48.

96. Основной вопрос метафизики // Философский журнал. 2014. № 2(13). С. 51–73.

97. Решетки четырехзначных модальных логик // Логические исследования. 2015. Т. 21. № 1. С. 122–134.

98. Предисловие. Многообразие трехзначности // Девяткин Л.Ю., Преловский Н.Н., Томова Н.Е. В границах трехзначности. М.: ИФ РАН, 2015. С. 9–33.

99. В поисках реальности: Исчезновение // Философия науки и техники. 2015. Т. 20. № 1. С. 36–81.

100. Тоска по философии (памяти А.П. Огурцова). Статья I // Вопросы философии. 2015. № 10. С. 133–149.

101. Тоска по философии (памяти А.П. Огурцова). Статья II // Вопросы философии. 2016. № 2. С. 137–147.

102. Сверхреализм. Часть I: От мыслимого к возможному // Философский журнал. 2016. Том 9. № 2. С. 5–23.

103. Сверхреализм. Часть II: От возможного к реальности // Философский журнал. 2016. Том 9. № 3. С. 5–24.

104. Яакко Хинтикка (1929–2015) // Логические исследования. 2016. Т. 22. № 1. С. 9–12.

105. Модальная пропозициональная логика истины Tr и ее полнота // Логические исследования. 2016. Т. 22. № 1. С. 13–31.

106. Bochvar's three-valued logic and literal paralogics: Their lattice and functional equivalence // Logic and Logical Philosphy. 2017. Vol. 26. № 2. P. 207–235. (в соавторстве с Н.Е. Томовой)

107. Four-valued logics BD and DM4: Expansions // Bulletin of the section of logic. 2017. Vol. 46. № 1/2. P. 33–45.

108. Контрфактуальное мышление // Логические исследования. 2017. Т. 23. № 2. С. 98–122.

Статьи в энциклопедиях

1. Контрфактические высказывания // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 385–386.

2. Логика // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 424–428.

3. Логика высказываний // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 385–386.

4. Логика многозначная // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 432–434.

5. Логика символическая // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 440–441.

6. Логические исследования в России и СССР // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 450–454. (в соавторстве с В.А. Бажановым)

7. Логическое следование // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 457–458.

8. Ложь // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 460.

9. Металогика // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 489.

10. Неклассические логики // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 592–594.

11. Немонотонные логики // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 594.

12. Нечеткая логика // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 603–604.

13. «Принципы математики» («Principia Mathematica») // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 739–740.

14. Символическая логика // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 859–860.

15. Фатализм логический // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 1031.

16. Философская логика // Энциклопедия Эпистемологии и Философии Науки. М., 2009. С. 1084–1086.

17. Алгебра логики // Новая философская энциклопедия. Т. 1. М.: Мысль, 2010 С. 75–76.

18. Логика в России // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 409–414.

19. Логика высказываний // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 415–418.

20. Логика символическа // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 429–431.

21. Логический фатализм // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 443.

22. Ложь // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 447–448.

23. Лукасевич // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 456–457. (В соавторстве с В.Н. Порусом)

24. Многозначные логики // Новая философская энциклопедия. Т. 2. М.: Мысль, 2010. С. 586–588.

25. Неклассические логики // Новая философская энциклопедия. Т. 3. М.: Мысль, 2010. С. 49–51.

26. Непротиворечия закон // Новая философская энциклопедия. Т. 3. М.: Мысль, 2010. С. 77. (В соавторстве с В.А. Смирновым)

27. Паранепротиворечивая логика // Новая философская энциклопедия. Т. 3. М.: Мысль, 2010. С. 198–199.

28. Пост Эмиль Леон // Новая философская энциклопедия. Т. 3. М.: Мысль, 2010. С. 292–293.

29. Смирнов Владимир Александрович // Новая философская энциклопедия. Т. 3. М.: Мысль, 2010. С. 574–575.

30. Философская логика // Новая философская энциклопедия. Т. 4. М.: Мысль, 2010. С. 246–247.

31. Логика // Новая Российская Энциклопедия. Т. IX(2). М.: Издательство «Энциклопедия», «ИНФРА-М», 2012. С. 492–495.

32. Логика в России // Новая Российская Энциклопедия. Т. IX(2). М.: Издательство «Энциклопедия», «ИНФРА-М», 2012. С. 498–500.

33. Неклассические логики // Большая Российская Энциклопедия. Т. 22. М.: Издательство «Большая Российская Энциклопедия», 2013. С. 323.

Публикации в материалах научных мероприятий

1. Каков истинностный статус утверждений о будущих событиях? // Модальные и интенсиональные логики (Тезисы координационного совещания, Москва, июнь 5-7, 1978 г.) М.: ИФАН, 1978. С. 66–70.

2. Фактор-семантика и классы многозначных систем логики // Релевантные логики и теория следования (Материалы II Советско-финского коллоквиума по логике. Москва, декабрь 3–7, 1979 г.) М.: ИФАН, 1979. С. 67–75.

3. The T-F-interpretation of some n-valued logics // 6th International Congress for Logic, Methodology and Philosophy of Science. Abstracts. Section 5. Hannover, 1979. P. 98–102.

4. Матричная логика для простых чисел // Модальные и интенсиональные логики (Материалы VIII Всесоюзной конференции «Логика и методология науки»). М.: ИФАН, 1982. С. 51–54.

5. Характеристическая матрица для простых чисел и ее свойства // Шестая Всесоюзная конференция по математической логике. Тбилиси, 1982. С. 76.

6. Аристотель и «морское сражение» // Симпозиум по логике Аристотеля. Тезисы докладов. Тбилиси, 1983. С. 51–54.

7. Будет или не будет завтра морское сражение? // Логика Аристотеля (Материалы симпозиума). Тбилиси, 1985. С. 98–111.

8. Истинностные значения // Интенсиональные логики и логическая теория теорий (Тезисы докладов IV Советско-финского коллоквиума по логике). Тбилиси, 1985. С. 72–73.

9. Логика как теория истинностных значений // Философские проблемы истории логики и методологии науки. Материалы Всесоюзной конференции. Часть 1. М.: ИФАН, 1986. С. 80–83.

10. Логика как истинностное значение // Логика и системные методы анализа научного знания. Тезисы докладов к IX Всесоюзному совещанию по логике, методологии и философии науки. М., 1986. С. 135–137.

11. Logic as truth-value // 8th International Congress of Logic, Methodology of Philosophy of Science. Moscow. Section 5. 1987. P. 263–265.

12. Связь импликации Лукасевича с импликацией Гейтинга // Семиотические аспекты формализации интеллектуальной деятельности. Тезисы докладов и сообщений. М., 1988. С. 79. (В соавторстве с В.Л. Васюковым)

13. T-F-sequences and their sets as truth-values // Intensional logic, history of philosophy and methodology. Budapest, 1988. P. 109–119.

14. Истинностные значения. Что это такое? // Исследования по неклассическим логикам. VI Советско-финский коллоквиум. М.: Наука, 1989. С. 38–53.

15. Взаимоотношение между n-значными логиками Лукасевича $Ł_n$ и Гёделя G_n // X Всесоюзная конференция по логике, методологии и философии науки. Тезисы докладов и выступлений (секции 1–5). Минск, 1990. С. 68–69.

16. Classification of implicational logics // 9th International Congress on Logic, Methodology and Philosophy of Science. Abstracts. Vol. 1. Section 5. Uppsala, 1991. P. 131.

17. Проблема о расширении $H_{\rightarrow}$ // Одиннадцатая Межреспубликанская конференция по математической логике. Казань, 1992. С. 69.

18. Classical propositional logic as the 7-dimensional cube // Book of abstracts of XIX World Congress of Philosophy. Moscow, 1993. Section 4. C. 274.

19. Классификация пропозициональных логик // Логика, методология, философия науки (XI Международная конференция, Москва–Обнинск, 1995). Т. 2. С. 33–36.

20. Classification of propositional logic // 10th International Congress on Logic, Methodology and Philosophy of Science. Volume of Abstracts. Florence, 1995.

21. Функциональная предполнота и закон порождения простых чисел // Современная логика: Проблемы теории, истории и применения в науке. Тезисы докладов. СПб., 1996. С. 22–26.

22. Два трехзначных изоморфа классической логики и их комбинации // Международная конференция «Развитие логики в России: Итоги и перспективы». М.: МГУ, 1997. С. 31–33.

23. Максимальная решетка импликативных логик // Международная конференция «Смирновские чтения». М., 1997. С. 18–19.

24. Two three-valued isomorphs of classical propositional logic and their combinations // 1st International Congress of Paraconsistent Logic. Brussel, 1997. P. 92–94.

25. The law of prime numbers generation: Logic and computer realization // International Conference on Information and Control. Proceedings. Vol. 2. St. Petersburg, 1997. P. 494–495.

26. Трехзначные импликативно-негативные паранепротиворечивые логики // Современная логика: Проблемы теории, истории и применения в науке. Материалы V Ощероссийской научной конференции. СПб., 1998. С. 137–140.

27. Булевы каскады импликативных логик // Смирновские чтения. 2. Международная конференция. М.: ИФРАН, 1999. С. 41–44.

28. Молекулярная паранепротиворечивая логика со свойством Сетте // Второй Российский Философский Конгресс. Екатеринбург, 1999. С. 246.

29. Two lattices of implicational logics // 11th International Congress of Logic, Methodology of Science and Philosophy. Krakow, 1999.

30. Логика: Феномены XX века // Современная логика: Проблемы теории, истории и применения в науке. Материалы VI Общероссийской научной конференции. СПб., 2000. С. 461–465.

31. Подструктурные логики: гильбертовские исчисления // 3 Международная конференция «Смирновские чтения». М.: ИФРАН, 2001. С. 36–38.

32. Логика и логические системы // Современная логика: Проблемы теории, истории и применения в науке. Материалы VII Общероссийской научной конференции. СПб., 2002. С. 455–458.

33. Современные исследования в философской логике // 4 Международная конференция «Смирновские чтения». М.: Институт философии РАН, 2003. С. 95–100.

34. Опыт преподавания спецкурса по многозначным логикам для студентов философского факультета МГУ // Проблемы преподавания логики и дисциплин логического цикла (Материалы Международной научно-практической конференции). Киев, 2004. С. 42.

35. Логика и логики // Смирновские чтения по логике (Материалы 5-ой конференции). М.: ИФРАН, 2007. С. 143–144.

36. Вклад А.А. Зиновьева в многозначную логику // Зиновьевские чтения (1-я Международная конференция). М.: Изд-во Московского гуманитарного ун-та, 2007. С. 90–93.

37. P-логики // Современная логика: проблемы теории, истории и применения в науке. Материалы X Общероссийской научной конференции, 26–28 июня 2008 г., Санкт-Петербург. СПб., 2008. С. 278–280.

38. Проблема континуальности трехзначных логик // Шестые Смирновские чтения по логике. М.: Современные тетради, 2009. С. 13–18.

39. Логика истинности Tr // Современная логика: Проблемы теории и истории. Материалы XI Международной конференции. СПб., 2010. С. 442–444.

40. Логика, обучение логике и протологика // Проблемы преподавания логики и дисциплин логического цикла. Киев, 2010. С. 18–19.

41. Philosophical problems of foundations of logic // Седьмые Смирновские чтения по логике. М., 2011. С. 76–79.

42. Тезис Сушко и его функциональное опровержение // Восьмые Смирновские чтения: Междунар. науч. конф., Москва, 19–21 июня 2013 г. С. 15–17.

43. Non-classical logics vs. classical logic // Philosophy and Logic – 2013, International Workshop, May 23–25, 2013. URL: `http://logic.univ.kiev.ua/wp-content/uploads/2013/03/Karpenko.docx` (дата обращения — 25.04.20)

44. Решетка фундаментальных четырехзначных модальных логик // 9-е Смирновские чтения: материалы Междунар. науч. конф. (г. Москва, 17–19 июня 2015 г.). М., 2015. С. 24–26.

45. Lattice of Three-Valued Literal Paralogics // Логико-философские штудии. 2016. Том 13. № 2. Материалы XII Международной конференции «Современная логика: проблемы теории и истории, СПб., 2016. С. 176–178. (В соавторстве с Н.Е. Томовой)

Методические пособия

1. Учебно-методические материалы по спецкурсу «Пропозициональная многозначная логика» (для студентов старших курсов). М.: МГУ, 1986.

2. Введение в многозначную пропозициональную логику М.: Институт философии РАН. Государственный университет гуманитарных наук, 2003. 112 с.

Публицистика, поэзия, проза

1. Псевдонимы / Составление А.С. Карпенко. М.–СПб.: ЦГИ, 2011. 224 с.

2. Альтернативная реальность: Гай Юлий Цезарь // Литературная Россия. 30.05.2014. № 22(2657). С. 14.

3. Александр Чаха. Дневник Белого Кардинала. М.–СПб.: ЦГИ, 2017. 280 с.

Информация для авторов

• Журнал «Логические исследования» принимает к публикации рукописи, содержащие изложение оригинальных результатов из различных областей современной логики, ранее не публиковавшиеся и не представленные к публикации в других изданиях. (Рубрики см. на сайте журнала `http://logicalinvestigations.ru`)

• Все статьи, планируемые к публикации в журнале «Логические исследования», проходят процедуру анонимного рецензирования.

• Решение о публикации текста принимается главным редактором с учетом мнения редколлегии и оценки рецензентов. Решение о публикации принимается в течение двух месяцев с момента предоставления рукописи.

• Плата за опубликование рукописей не взимается.

• Рукопись должна быть представлена в электронном виде и оформлена в формате LaTeX 2_ε (по согласованию с редколлегией — в MS Word с обязательным предоставлением pdf-файла).

• При подготовке рукописи в LaTeX 2_ε необходимо использовать стилевой класс `LIarticle.cls` и шаблон `LI_template.tex` (если рукопись на русском языке) или `LI_template_eng.tex` (если рукопись на английском языке), которые размещены в правилах предоставления рукописей на сайте `http://logicalinvestigations.ru`. Здесь же размещены подробные рекомендации по подготовке рукописи.

• Объем рукописи не должен превышать 25 стр. (60 тыс. знаков), включая ссылки, примечания, списоки литературы, аннотации (на русском и английском языках).

Статьи следует направлять в электронном виде в редакцию через сайт
`logicalinvestigations@gmail.com`

Information for authors

• *Logical Investigations* accepts for submission papers containing original results in all areas of logic. The papers should not have been published or simultaneously submitted to another publication. (Sections of the journal: `http://logicalinvestigations.ru`)

• All materials published in *Logical Investigations* undergo peer review process.

• The Editor in Chief makes the decision which of the submitted articles should be published, with due account for opinions of the Editorial Board and the reviewers. The decision is made within two months since the date of submission of the manuscript.

• Authors are not charged for the publication.

• Papers should be submitted electronically in the $\LaTeX\,2_\varepsilon$ format (special permission of the editorial board is needed for submissions to be made in the MS Word format).

• While typesetting a paper, the class file `LIarticle.cls` and the template `LI_template_eng.tex` should be used; both files can be accessed at `http://logicalinvestigations.ru`. Here you can also find detailed guidelines for preparing your paper.

• Papers should not exceed 25 pages in the above mentioned format (including the notes, the bibliography, the abstract).

Submissions should be sent in electronic form through the website:
`logicalinvestigations@gmail.com`

Научно-теоретический журнал

Логические исследования / Logical Investigations
2020. Том 26. Номер 1

Учредитель и издатель: Федеральное государственное бюджетное учреждение науки Институт философии Российской академии наук

Свидетельство о регистрации СМИ: *ПИ № ФС77-61228 от 03.04.2015 г.*

Главный редактор: *В.И. Шалак*

Ответственный секретарь: *Н.Е. Томова*

Технический редактор: *Е.А. Морозова*, *А.В. Конькова*

Художник: *Н.Н. Попов*, *С.Ю. Растегина*

Подписано в печать с оригинал-макета 29.05.2020.
Формат 70x100 1/16. Печать офсетная. Гарнитура Computer Modern.
Для набора греческого текста исользован пакет Teubner.
Усл. печ. л. 13,22. Уч.-изд. л. 10,18. Тираж 1 000 экз. Заказ № 11.

Оригинал-макет изготовлен в Институте философии РАН

Разработка LaTeX-класса стилевого оформления оригинал-макета: *А.Е. Коньков*

Компьютерная верстка: *Н.Е. Томова*

Отпечатано в ЦОП Института философии РАН
109240, г. Москва, ул. Гончарная, д. 12, стр. 1

Информацию о журнале «Логические исследования» см. на сайте:
`http://logicalinvestigations.ru`