

Trivalent Logics and their applications

Proceedings of the ESSLLI 2012 Workshop

Edited by
Paul Égré and David Ripley

Version 08/11/2012

Contents

Foreword and acknowledgements <i>Paul Égré and David Ripley</i>	iii
Programme Committee	v
Workshop Description	vii
Invited papers: <i>Arnon Avron, Grzegorz Malinowski, Katrin Schulz</i>	ix
Contributed papers	1
1 A Multi-Valued Delineation Semantics for Absolute Adjectives <i>Heather Burnett</i>	1
2 Are True and False not Enough? <i>Vincent Degauquier</i>	17
3 Modeling the Suppression Task under Three-Valued Łukasiewicz and Well-Founded Semantics <i>Emmanuelle-Anna Dietz and Steffen Hölldobler</i>	27
4 Reasoning about Rough Sets Using Three Logical Values <i>Beata Konikowska and Arnon Avron</i>	39
5 Doing the right things — trivalence in deontic action logic <i>Piotr Kulicki and Robert Trypuz</i>	53
6 Trivalent logics arising from L-models for the Lambek calculus with constants <i>Stepan Kuznetsov</i>	65
7 A trivalent logic that encapsulates intuitionistic and classical logic <i>Tin Perkov</i>	71
8 TCS for presuppositions <i>Jérémy Zehr and Orin Percus</i>	77

Foreword and acknowledgments

Welcome to this ESSLLI 2012 workshop on Trivalent Logics and their Applications!

Besides our contributors and invited speakers to this workshop, we wish to thank all the people and institutions who helped us organize this event, in particular the members of our programme committee and the additional referees listed below. We also express our gratefulness to the following sponsors:

- Pablo Cobreros' Project 'Borderlineness and Tolerance' (FFI2010-16984) funded by the Ministerio de Ciencia e Innovación, Government of Spain
- François Récanati's CCC project under the European Research Council (FP7/2007-2013) / ERC Advanced Grant agreement n229 441-CCC)
- Philippe Schlenker's EURYI project 'Presupposition: A formal pragmatic approach' hosted by Institut Jean-Nicod
- The CNRS, Institut Jean-Nicod, and the University of Melbourne for additional support.

Special thanks moreover go to Andreas Herzig for encouraging this proposal and to the ESSLLI 2012 organizing committee for their support and for hosting our workshop.

We received 14 submissions for this workshop, 8 of which were accepted for presentation, and 1 as alternative. Three initially scheduled papers could not be included in these proceedings as the speakers had to cancel their participation: a contributed paper by Yasutada Sudo and colleagues, a contributed paper by Mark Jago, and an invited paper by Janneke Huitink. We thank Katrin Schulz in particular for accepting to replace Janneke Huitink as a plenary speaker.

We hope this workshop will be productive and look forward to seeing you there!

Paul Égré and David Ripley, August 2012

Program Committee

Arnon Avron
Pablo Cobreros
Paul Egré (co-chair)
Janneke Huitink
Grzegorz Malinowski
David Ripley (co-chair)
Robert van Rooij

Additional Reviewers

Wojciech Buszkowski
Szymon Frankowski
Hans Smessaert
Anna Wójtowicz

Workshop description

Three-valued logics have been an object of extensive study since at least the work of Lukasiewicz, with applications to a wide range of natural language phenomena, including presupposition, conditionals and vagueness. While many-valued logics can be studied on their own, there has been a regain of interest for three-valued logics in recent years, with the emergence of new perspectives regarding their applicability to natural language.

In the theory of presupposition projection, in particular, the question of whether the projection of presupposition can be dealt with by means of a trivalent truth-functional semantics has been the object of renewed attention, in particular because truth-functional trivalent approaches appear as a main competitor to both dynamic and pragmatic approaches (viz. Beaver and Kraemer 2001, George 2008, Fox 2008, all of them giving special attention to so-called middle-Kleene logic proposed by Peters, and the recent debates with Schlenker). In the area of vagueness, ways have been proposed to combine the canonical paraconsistent and paraconsistent three-valued logics of Kleene and Priest in order to deal with the paradoxes of vagueness, and to account for phenomena such as meaning coarsening and strengthening (viz. Avron et Konikowska 2008, Cobreros et al. 2012). In the literature on conditionals, finally, the question remains largely open of the selection between a wide range of candidates for the definition of a suitable three-valued conditional (viz. Bradley 2002, Cantwell 2008, Huitink 2009, Rothschild 2009). From a more foundational point of view, finally, the meaning attached to the third truth value can vary significantly depending on the problem under consideration and the definition of logical consequence considered to be relevant.

The aim of this workshop is to promote new contributions for the extension of two-valued logic with a third truth-value. Submissions have been encouraged on logical and linguistics aspects of the use of 3-valued logics, with relevance on the following topics:

- applications of trivalent logic to quantification in natural language
- trivalent logics for conditionals / vagueness / presupposition
- are vagueness and presupposition susceptible of a unified treatment in trivalent logic?
- logical consequence and proof-theory for three-valued logic

- unification and classification of 3-valued logics
- connection between 3-valued logics and other non-classical logics
- partial 2-valued logics vs. 3-valued logics
- do we need more than three truth-values? can we dispense with a third truth value?

Invited papers

Arnon Avron (Tel Aviv University) *Using Trivalent Semantics for Paraconsistent Reasoning*

We describe a general method for a systematic and modular generation of cut-free calculi for thousands of paraconsistent logics. The method relies on the use of 3-valued non-deterministic semantics for these logics.

Grzegorz Malinowski (University of Łódź) *Logical three-valuedness and beyond*

The modern history of many-valuedness starts with Łukasiewicz's construction of three-valued logic. This pioneering, philosophically motivated and matrix based construction, first presented in 1918, was in 1922 extended to n-valued cases, including two infinite. Soon several constructions of many-valued logic appeared and the history of the topic became rich and interesting. However, as it is widely known, the problem of interpretation of multiple values is still among vexed questions of contemporary logic. With the talk, which essentially groups my earlier settlements, I intend to put a new thread into discussion on the nature of logical many-valuedness. The topics, touched upon, are: matrices, tautological and non-tautological many-valuedness, Tarski's structural consequence and the Lindenbaum-Wójcicki completeness result, which supports the Suszko's claim on logical two-valuedness of any structural logic. Further to that, two facets of many-valuedness: referential and inferential are unravelled.

The first fits the standard approach and it results in multiplication of semantic correlates of sentences, and not logical values in a proper sense. It is based on the matrix approach and results in a multiple-element referential extensionality. In that paradigm, the central concepts are: the tautological many-valuedness and many-valued consequence.

The second many-valuedness is a metalogical property of quasi-consequence and refers to partition of the matrix universe into more than two disjoint subsets, used in the definition of inference, using the inference rules, which from non-rejected premises lead to the accepted conclusions.

Katrin Schulz (ILLC, Amsterdam) *TBA*

Contributed papers

A Multi-Valued Delineation Semantics for Absolute Adjectives

Heather Burnett

University of California, Los Angeles
hburnett@ucla.edu

1 Introduction

This paper provides a novel semantic analysis of the gradability of adjectives of the *absolute* class within a delineation (i.e. comparison-class-based) semantic framework (first presented in [8]). It has been long observed that the syntactic category of bare adjective phrases can be divided into two principle classes: *scalar* (or *gradable*) vs *non-scalar* (*non-gradable*). The principle test for scalarity of an adjective P is the possibility of P to appear (without coercion) in the explicit comparative construction. Thus, we find a first distinction between adjectives like *tall*, *expensive*, *straight*, *empty*, and *dry* on the one hand (ok: *taller*, *more expensive*, *straighter*, *emptier*, *drier*) and *atomic*, *pregnant*, and *geographical* on the other (?*more atomic*, ?*more pregnant*, ?*more geographical*). It has been argued by many authors that the class of scalar adjectives is further decomposed into two principle subclasses: *relative* adjectives (henceforth *RAs*: ex. *tall*, *short*, *expensive*, *intelligent*) and *absolute* adjectives (henceforth *AAs*: ex. *empty*, *straight*, *dry*, *clean*). Although RAs and AAs behave differently in many syntactic and semantic constructions, the fundamental difference between these two classes of adjectives is generally taken to be that members of the former class have context-sensitive semantic denotations (denotations that vary depending on contextually given comparison classes); whereas, members of the latter class have semantic denotations that are independent of context (cf. [16], [7], [12]). As discussed in [13], this empirical observation raises a puzzle for the delineation approach, since, as will be outlined below, in this framework, the scales associated with adjectival constituents are derived through looking at how their semantic denotations vary across comparison classes. The inability of comparison-class-based frameworks to treat the difference between absolute and relative adjectives has been taken (for example by [6]) to be a major argument against a delineation semantics for scalar adjectives and in favour of a semantics in which degrees and scales are primitives.

In this paper, I present a new solution to the puzzle of the gradability of AAs within the delineation approach, one that takes into account the empirical observation that these constituents can be used imprecisely or vaguely (cf. [10], [11], [6], a.o.). I show that by integrating a simplified version of [8]’s comparison-class-based logical system with the similarity-based multi-valued logical framework proposed by [4] to model the vagueness/imprecision associated with these predi-

cates, we can arrive at new logical framework that can treat the absolute/relative distinction without degrees in the ontology.

The paper is organized as follows: in section 2, I present the delineation framework for the semantic analysis of gradable predicates. Then, in section 3, I present the main ways in which adjectives like *tall* differ from adjectives like *empty*, and I argue that the latter adjectives challenge the comparison-class-based approach. In section 4, I present the empirical observation that absolute adjectives are subject to the phenomenon of vagueness/imprecision, and I introduce the multi-valued logical system that I will be employing to model this phenomenon, [4]’s *Tolerant, Classical, Strict* (TCS). Finally, in section 5, I give my analysis of the gradability of AAs within a delineation extension of TCS. In particular, I propose that the non-trivial scales associated with AAs are derived through looking at comparison-class-based variation in predicate-relative similarity/indifference relations, and I show how these relations can be constructed within this new approach using methods in the same vein as [1] and [14]. The new framework is formally laid out (with the proofs of the main results of the paper) in the appendix.

2 Delineation Semantics

Delineation semantics is a framework for analyzing the semantics of gradable expressions that takes the observation that they are context sensitive to be their key feature. A delineation approach to the semantics of positive and comparative constructions was first proposed by [8], and has been further developed by many authors in the past 30 years. In this framework, scalar adjectives denote sets of individuals and, furthermore, they are evaluated with respect to comparison classes, i.e. subsets of the domain D . The basic idea is that the extension of a gradable predicate can change depending on the set of individuals that it is being compared with. In other words, the semantic denotation of the positive form of the scalar predicate (i.e. *tall*) can be assigned a different set of individuals in different comparison classes.

Definition 1. *CC-relativized interpretation of predicates (informal).*

1. For a scalar adjective P and a contextually given comparison class $X \subseteq D$,

$$(1) \quad \llbracket P \rrbracket_X \subseteq X.$$

2. For an individual a , a scalar adjective P , and a contextually given comparison class $X \subseteq D$,

$$(2) \quad \llbracket a \text{ is } P \rrbracket_X = 1 \text{ iff } \llbracket a \rrbracket \in \llbracket P \rrbracket_X.$$

Unlike degree semantics (cf. [6]), delineation semantics takes the positive form as basic and derives the semantics of the comparative form from quantification over comparison classes. Informally, *John is taller than Mary* is true just in case there is some comparison class with respect to which John counts as tall and Mary counts as not tall.

Definition 2. Semantics for the comparative (informal). For two individuals a, b and a scalar adjective P , $\llbracket a \text{ is } P\text{-er than } b \rrbracket = 1$ iff $a >_P b$, where $>_P$ is defined as:

- (3) $x >_P y$ iff there is some comparison class X such that $x \in \llbracket P \rrbracket_X$ and $y \notin \llbracket P \rrbracket_X$.

As it stands, the analysis of the comparative in definition 2 is very weak and allows some very strange and un-comparative-like relations¹, if we do not say anything about how the extensions of gradable predicates can change in different comparison classes (CCs). A solution to this problem involves imposing some constraints on how predicates like *tall* can be applied in different CCs. In this work, I will adopt the set of constraints on the application of gradable predicates presented in [1] and [2]. Van Benthem proposes three axioms governing the behaviour of individuals across comparison classes. They are the following (presented in my notation):

For $x, y \in D$ and $X \subseteq D$ such that $x \in \llbracket P \rrbracket_X$ and $y \notin \llbracket P \rrbracket_X$,

- (4) **No Reversal (NR:)** There is no $X' \subseteq D$ such that $y \in \llbracket P \rrbracket_{X'}$ and $x \notin \llbracket P \rrbracket_{X'}$.
- (5) **Upward difference (UD):** For all X' , if $X \subseteq X'$, then there is some $z, z' : z \in \llbracket P \rrbracket_{X'}$ and $z' \notin \llbracket P \rrbracket_{X'}$.
- (6) **Downward difference (DD):** For all X' , if $X' \subseteq X$ and $x, y \in X'$, then there is some $z, z' : z \in \llbracket P \rrbracket_{X'}$ and $z' \notin \llbracket P \rrbracket_{X'}$.

No Reversal states that if $x >_P y$, there is no X' such that y is in $\llbracket P \rrbracket_{X'}$, but x is not. **Upward Difference** states that if, in the comparison class X , there is a $P/\text{not } P$ contrast, then a $P/\text{not } P$ contrast is preserved in every larger CC. Finally, **Downward Difference** says that if in some comparison class X , there is a $P/\text{not } P$ contrast involving x and y , then there remains a contrast in every smaller CC that contains both x and y . van Benthem shows that these axioms give rise to *strict weak orders*: irreflexive, transitive and almost connected relations².

¹ For example, suppose in the CC $\{a, b\}$, $a \in \llbracket P \rrbracket_{\{a, b\}}$ and $b \notin \llbracket P \rrbracket_{\{a, b\}}$. So $a >_P b$. And suppose moreover that, in the larger CC $\{a, b, c\}$, $b \in \llbracket P \rrbracket_{\{a, b, c\}}$ and $a \notin \llbracket P \rrbracket_{\{a, b, c\}}$. So $b >_P a$. But clearly, natural language comparatives do not work like this: If John is {taller, fatter, wider...} than Mary, Mary cannot also be {taller, fatter, wider...} than John. In other words, $>_P$ must be asymmetric.

² The definitions of *irreflexivity*, *transitivity* and *almost connectedness* are given below.

Definition 3. Irreflexivity. A relation $>$ is irreflexive iff there is no $x \in D$ such that $x > x$.

Definition 4. Transitivity. A relation $>$ is transitive iff for all $x, y, z \in D$, if $x > y$ and $y > z$, then $x > z$.

Definition 5. Almost Connectedness. A relation $>$ is almost connected iff for all $x, y \in D$, if $x > y$, then for all $z \in D$, either $x > z$ or $z > y$.

Definition 6. *Strict weak order.* A relation $>$ is a strict weak order just in case $>$ is *irreflexive, transitive, and almost connected*.

As discussed in [8], [2] and [14], strict weak orders (also known as *ordinal scales* in measurement theory) intuitively correspond to the types of relations expressed by many kinds of comparative constructions³. Thus, the theorem in 1 is an important result in the semantic analysis of comparatives, and it shows that scales associated with gradable predicates can be constructed from the context-sensitivity of the positive form and certain axioms governing the application of the predicate across different contexts.

Theorem 1. *Strict Weak Order.* For all P , $>_P$ is a strict weak order.

Proof. [1]; [2], p. 116. □

This analysis seems appropriate for relative predicates like *tall* and *short*; however, as we will see in the next section, it does not capture the certain aspects of the meaning of absolute predicates like *empty* and *straight*.

3 The Absolute/Relative Distinction

Following many authors, I take the principle way in which AAs like *empty* and *straight* differ from RAs like *tall* and *fat* is that AAs are not context-sensitive in the same way that RAs are. One test that shows this is the *definite description* test. As observed by [6] and [15] a.o., adjectives like *tall* and *empty* differ in whether they can ‘shift’ their thresholds (i.e. criteria of application) to distinguish between two individuals in a two-element comparison class when they appear in a definite description. For example, suppose there are two containers (A and B), and neither of them are particularly tall; however, A is (noticeably) taller than B. In this situation, if someone asks me (7-a), then it is very clear that I should pass A. Now suppose that container A has less liquid than container B, but neither container is particularly close to being completely empty. In this situation, unlike what we saw with *tall*, (7-b) is infelicitous.

- (7) a. Pass me **the tall one**.
 b. Pass me **the empty one**.

In other words, unlike RAs, AAs cannot change their criteria of application to distinguish between objects that lie in the middle of their associated scale. Using this test, we can now make the argument that adjectives like *full*, *straight*,

³ For example, one cannot be taller than oneself; therefore $>_{tall}$ should be irreflexive. Also, if John is taller than Mary, and Mary is taller than Peter, then we know that John is also taller than Peter. So $>_{tall}$ should be transitive. Finally, suppose John is taller than Mary. Now consider Peter. Either Peter is taller than Mary (same height as John or taller) or he is shorter than John (same height as Mary or shorter). Therefore, $>_{tall}$ should be almost connected.

and *bald* are absolute, since (8-a) is infelicitous if neither object is (close to) completely full/straight/bald. Likewise, we can make the argument that *dirty*, *wet*, and *bent* are also absolute, since (8-b) is infelicitous when comparing two objects that are at the middle of the dirtiness/wetness/curvature scale (i.e. both of them are dirty/wet/bent).

- (8) **Absolute Adjectives**
- a. Pass me the **full/straight/bald** one.
 - b. Pass me the **dirty/wet/bent** one.

Furthermore, we can make the argument that *long*, *expensive*, and even colour adjectives like *blue* are relative, since the (9) is felicitous when comparing two objects when both or neither are particularly long/expensive/blue⁴.

- (9) **Relative Adjectives**
Pass me the **long/expensive/blue** one.

How can we capture this distinction in a delineation framework? An idea that has been present in the literature for a long time, and has recently been incarnated in, for example, [7] and [12], is that unlike *tall* or *long* that have a context sensitive meaning, adjectives like *straight*, *empty* or *bald* are not context sensitive (hence the term *absolute* adjective). That is, in order to know who the bald people are or which rooms are empty, we do not need compare them to a certain group of other individuals, we just need to look at their properties. To incorporate this idea into the delineation approach, I propose (following an idea in [14]) that, in a semantic framework based on comparison classes, what it means to be non-context-sensitive is to have your denotation be invariant across classes. Thus, for an absolute adjective Q and a comparison class X , it suffices to look at what the extension of Q is in the maximal CC, the domain D , in order to know what $\llbracket Q \rrbracket_X$ is. I therefore propose that a different axiom set governs the semantic interpretation of the members of the absolute class that does not apply to the relative class: the singleton set containing the *absolute adjective axiom*.

- (10) **Absolute Adjective Axiom (AAA).**
If $Q \in AA$, then for all $X \subset D$ and $x \in X$, $x \in \llbracket Q \rrbracket_X$ iff $x \in \llbracket Q \rrbracket_D$.

In other words, the semantic denotation of an absolute adjective is set with respect to the total domain, and then, by the AAA, the interpretation of Q in D is replicated in each smaller comparison class. The AAA is very powerful: as shown by theorem 2, the scales that the semantic denotations of absolute constituents give rise to are very small, essentially trivial.

Theorem 2. *If Q satisfies the AAA, $>_Q$ is homomorphic to the two element boolean algebra (sometimes written 2).*

Proof. Consider the function $h : D \rightarrow 2$.

⁴ For an example of the use of a colour adjective like *blue* to distinguish between two not particularly blue objects, see [5].

(11) For all $x \in D$, $h(x) = 1$ iff $x \in \llbracket Q \rrbracket_D$.

Show for all $x, y \in D$, if $x >_Q y$ then $h(x) >_2 h(y)$. Immediately from the definition of h . \square

Absolute adjectives thus raise a puzzle for delineation analyses:

(12) **The Puzzle of Absolute Adjectives:**

If AAs have non-context-sensitive semantic denotations, how can they be gradable?

In the rest of the paper, I will provide a solution to this puzzle.

4 Vagueness/Imprecision with AAs

Of course, saying that adjectives like *empty*, *bald*, and *straight* are not at all context-sensitive is clearly false. As observed by very many authors (ex. [16], [10], [11], [7], [6] a.o.), the criteria for applying an absolute adjective can vary depending on context, as exemplified in (13).

- (13) a. Only two people came to opening night; the theatre was empty.
 b. Two people didn't evacuate; the theatre wasn't empty when they started fumigating.

Rather than being attributed directly to the context-sensitivity of their semantic denotation, the contextual variation in the application of absolute predicates is generally attributed to something that is variably called “imprecision”, “loose talk” or “vagueness”, among other things. I therefore propose that the context-sensitivity that allows for the construction of non-trivial scales is not semantic, as in the case of relative adjectives (as outlined in section 2), but **pragmatic**: although the semantic denotation of an absolute predicate does not vary across comparison classes, its denotation on its imprecise use does.

As mentioned in the introduction, the approach that I will adopt to model the effects of vagueness/imprecision is [4]’s *Tolerant, Classical, Strict* (TCS). This system was developed as a way to preserve the intuition that vague and imprecise predicates⁵ are *tolerant* (i.e. satisfy $\forall x \forall y [P(x) \ \& \ x \sim_P y \rightarrow P(y)]$, where \sim_P is a ‘little by little’ or indifference relation for a predicate P), without running into the Sorites paradox⁶. [4] adopt a non-classical logical framework with three notions of satisfaction: classical satisfaction, tolerant satisfaction, and its dual,

⁵ The system in [4] was proposed to model the puzzling properties of vague language with relative predicates like *tall*; however, I suggest that the results in this paper show that it has a natural application to modelling similar effects with absolute adjectives.

⁶ Note that on their imprecise use, absolute predicates like *bald* and *empty* give rise to Soritical-type reasoning: how many hairs must someone have before they stop being considered bald? How many seats must be filled before a theatre is no longer considered empty?

strict satisfaction. Formulas are tolerantly/strictly satisfied based on classical truth and predicate-relative, possibly non-transitive *indifference relations*. For a given predicate P , an indifference relation, \sim_P , relates those individuals that are viewed as sufficiently similar with respect to P . For example, for the predicate *empty*, \sim_{empty} would be something like the relation “differ by a number of objects that is irrelevant for our purposes/contain roughly the same number of objects”. Since these relations are given by context, we assume that they are part of the model. I give the definition of the indifference relations (within a comparison class-based framework) below.

Definition 7. CC-relativized indifference relations. For all scalar adjectives P and comparison classes $X \subseteq D$,

$$(14) \quad \sim_P^X \text{ is a binary relation on the elements of } X \text{ that is reflexive and symmetric (but not necessarily transitive).}$$

In this framework, we say that *Room A is empty* is tolerantly true just in case Room A contains a number of objects that do not cause us to make a distinction between it and a completely empty room in the context. For the purposes of the analyses in this paper, I will suppose that classical satisfaction and classical denotations correspond to regular semantic satisfaction and semantic denotations, while tolerant and strict satisfaction and denotations correspond to pragmatic notions⁷. The three notions of satisfaction are defined (informally) within a comparison-class-based system⁸ as shown below.

Definition 8. Classical ($\llbracket \cdot \rrbracket^c$), tolerant ($\llbracket \cdot \rrbracket^t$), and strict ($\llbracket \cdot \rrbracket^s$) interpretation of predicates. For all scalar adjectives P and $X \subseteq D$,

1. $\llbracket P \rrbracket_X^c \subseteq X$.
2. $\llbracket P \rrbracket_X^t = \{x : \exists d \sim_P^X x \text{ and } d \in \llbracket P \rrbracket_X^c\}$.
3. $\llbracket P \rrbracket_X^s = \{x : \forall d \sim_P^X x, d \in \llbracket P \rrbracket_X^c\}$.

Definition 9. Classical, tolerant, and strict satisfaction. For all individuals a , scalar predicates P , and comparison classes $X \subseteq D$,

$$(15) \quad \llbracket a \text{ is } P \rrbracket_X^{t/c/s} = 1 \text{ iff } \llbracket a \rrbracket \in \llbracket P \rrbracket_X^{t/c/s}.$$

The definitions of the tolerant and strict comparative relations are parallel to the classical comparative (definition 2).

Definition 10. Classical/tolerant/strict comparative (informal). For two individuals a, b and a scalar adjective P , $\llbracket a \text{ is } P\text{-er than } b \rrbracket^{t/c/s} = 1$ iff $a >_P^{t/c/s} b$, where $>_P^{t/c/s}$ is defined as:

⁷ As such, my interpretation of the framework bares many similarities with [9]’s *Pragmatic Halos* approach to modelling “pragmatic slack” or “loose talk”.

⁸ Note that, in his 1980 paper, Klein adopts a supervaluationist account of the vagueness of scalar adjectives. Thus, the integration of Klein’s basic semantics for the comparative construction with a similarity-based account of vagueness is a departure from the system presented in [8].

- (16) $x >_P^{t/c/s} y$ iff there is some comparison class X such that $x \in \llbracket P \rrbracket_X^{t/c/s}$ and $y \notin \llbracket P \rrbracket_X^{t/c/s}$.

The precise definition of TCS, set within a comparison-class-based approach to the semantics of scalar terms, is given in the appendix.

5 Analysis of Absolute Adjectives

In order to account for how AAs can have, at the same time, a semantic denotation that is constant across CCs, but at the same time be associated with non-trivial scales, I propose that what can vary across CCs are the indifference relations i.e., the \sim_Q^X s. For example, if I compare Homer Simpson, who has exactly two hairs, directly with Yul Brynner (who has zero hairs), the two would not be considered indifferent with respect to baldness (Homer has hair!). However, if I add Marge Simpson into the comparison class (she has a very large hairdo), then Yul and Homer start looking much more similar, when it comes to baldness. Thus, I propose, it should be possible to order individuals with respect to how close to being completely bald (or empty or straight) they are by looking at in which comparison classes they are considered indifferent to completely bald/empty/straight individuals⁹.

In what follows, I present a set of axioms that constrain indifference relations between individuals across comparison classes. Recall that I proposed that, unlike relative adjectives which are only subject to van Benthem's axioms (NR, UD, and DD), absolute adjectives are subject to the AAA. Then, in the spirit of [1] and [13], I will show that these axioms will allow us to construct non-trivial strict weak orders from the tolerant meaning of absolute predicates¹⁰.

5.1 Pragmatic Axiom Set

I propose the following axioms to constrain indifference relations¹¹.

⁹ The idea is conceptually similar in some sense (although extremely different in its execution) to a suggestion made by [12], with respect to how an adjective like *empty* can be both absolute and gradable.

¹⁰ For lack of space, I will only address the analysis of so-called *total* or *universal* AAs like *empty*, *bald*, and *straight*. However, the analysis of partial/existential AAs like *dirty* and *wet* is essentially the dual of the analysis of total AAs, with non-trivial scales being constructed out of strict denotations instead of tolerant ones. See [3] for discussion.

¹¹ One of the axioms (T-NS) makes reference to a 'tolerantly greater than or equal relation' (\geq^t): We first define an equivalence relation \approx_P :

Definition 11. Tolerantly Equivalent. (\approx^t) For a predicate Q and $a, b \in D$,

- (i) $a \approx_Q^t b$ iff $a \not\prec_Q^t b$ and $b \not\prec_Q^t a$.

Now we define \geq^t :

- (17) **Tolerant No Skipping (T-NS):** For an AA Q , $X \in \mathcal{P}(D)$ and $x, y \in X$, if $x \sim_Q^X y$ and there is some $z \in X$ such that $x \geq_Q^t z \geq_Q^t y$, then $x \sim_Q^X z$.

Tolerant No Skipping says that, if person A is indistinguishable from person B, and there's a person C lying in between persons A and B on the relevant tolerant scale, then A and C (the greater two of $\{A, B, C\}$) are also indistinguishable. As discussed in the appendix, T-NS performs a very similar function to van Benthem's *No Reversal*.

We now have two axioms that talk about how indifference relations can change across comparison classes. I call these the *granularity* axioms.

- (18) **Granularity 1 (G1):** For an AA Q , $X \in \mathcal{P}(D)$, and $x, y \in X$, if $x \sim_Q^X y$, then for all $X' \subseteq D : X \subseteq X'$, $x \sim_Q^{X'} y$.

G1 says that if person A and person B are indistinguishable in comparison class X , then they are indistinguishable in all supersets of X . This is meant to reflect the fact that the larger the domain is (i.e. the larger the comparison class is), the more things can cluster together¹².

- (19) **Granularity 2 (G2):** For an AA Q , $X, X' \subseteq D$, and $x, y \in X$, if $X \subset X'$ and $x \not\sim_Q^X y$ and $x \sim_Q^{X'} y$, then $\exists z \in X' - X : x \not\sim_Q^{X'} z$.

G2 says that, if person A and person B are distinguishable in one CC, X , and then there's another CC, X' , in which they are indistinguishable, then there is some person C in $X' - X$ that is distinguishable from person A. This axiom is similar in spirit to van Benthem's *Upward Difference* in that it ensures that, if there is a contrast/distinction in one comparison class, the existence of contrast is maintained in all the larger CCs.

The final axiom that we need is *Minimal Difference*:

- (20) **Minimal Difference (MD):** For an AA Q and $x, y \in D$, if $x >_Q^c y$, then $x \not\sim_Q^{\{x,y\}} y$.

Minimal Difference says that, if, at the finest level of granularity, you would make a classical distinction between two individuals, then they are not indistinguishable at that level of granularity. MD is similar in spirit to van Benthem's *Downward Difference* because it allows us to preserve contrasts down to the smallest comparison classes.

With these axioms, we can prove the main result of the paper (which is proved in the appendix):

Definition 12. Tolerantly greater than or equal. (\geq^t) For $a, b \in D$,

- (ii) $a \geq_Q^t b$ iff $a >_Q^t b$ or $a \approx_Q^t b$.

¹² G1 can be weakened a bit to allow some indifference relations to be undone in larger CCs. In this case, we derive semi-orders instead of strict weak orders (cf. [3]).

- (21) **Theorem 3.** If Q is an absolute adjective, then $>_Q^t$ is a strict weak order.

6 Conclusion

In this paper, I gave a new analysis of the semantics and pragmatics of absolute adjectives, and, in particular, I addressed the question of how AAs can have a non-context-sensitive semantic denotation but still be gradable with a delineation framework. I showed that the scales (i.e. strict weak orders) that are associated with absolute predicates can be derived in within the multi-valued delineation TCS system from certain intuitive statements about how individuals can and cannot be indifferent across comparison classes. Thus, I argue that the puzzles raised by absolute adjectives for the delineation approach can be solved, provided that we have an appropriate framework to treat vagueness and imprecision.

References

1. J. van Benthem. (1982). “Later than late: On the logical origin of the temporal order”. *Pacific Philosophical Quarterly*, 63:193-203.
2. J. van Benthem. (1990). *The Logic of Time*. Dordrecht: Reidel.
3. H. Burnett. (2012). *The Grammar of Tolerance: On Vagueness, Context-Sensitivity, and the Origin of Scale Structure*. PhD Dissertation. UCLA.
4. P. Cobreros, P. Égré, D. Ripley, and R. van Rooij. (2011). “Tolerant, Classical, Strict.” *Journal of Philosophical Logic*. (forthcoming).
5. D. Fara. (2000). “Shifting Sands: An interest-relative theory of vagueness.” *Philosophical Topics*. 20: 45-81.
6. C. Kennedy. (2007). “Vagueness and Grammar: The study of relative and absolute gradable predicates.” *Linguistics and Philosophy*. 30: 1-45.
7. C. Kennedy and L. McNally. (2005). “Scale structure and the semantic typology of gradable predicates”. *Language*. 81:345-381.
8. E. Klein. (1980). “A semantics for positive and comparative adjectives”. *Linguistics and Philosophy*. 4:1-45.
9. P. Lasnik. (1999). “Pragmatic Halos.” *Language*. 75: 522-571.
10. D. Lewis. (1979). “Score-keeping in a language game”. *Journal of Philosophical Logic*, 8: 339-359.
11. M. Pinkal. (1995). *Logic and Lexicon*. Dordrecht: Kluwer Academic Publishers.
12. F. Récanati. (2010). *Truth-Conditional Pragmatics*. Oxford: OUP.
13. R. van Rooij. (2011). “Vagueness and Linguistics”. In G. Ronzitti, editor, *The vagueness handbook*, Dordrecht: Springer. pp. 1-57.
14. R. van Rooij. (2011). “Implicit vs explicit comparatives”. In Paul Égré and Nathan Klinedinst (eds.), *Vagueness and Language Use*, Palgrave Macmillan.
15. K. Syrett et al. (2010). “Meaning and context in children’s understanding of gradable adjectives”. *Journal of Semantics*, 27:1-35.
16. P. Unger. (1975). *Ignorance*. Oxford: Clarendon Press.

7 Appendix: Framework and Proofs

7.1 The Framework: Delineation TCS

Language

Definition 13. Vocabulary. *The vocabulary consists of the following expressions:*

1. A series of individual constants: $a_1, a_2, a_3 \dots$
2. A series of individual variables: $x_1, x_2, x_3 \dots$
3. Two series of unary predicate symbols:
 - Relative scalar adjectives: $P_1, P_2, P_3 \dots$
 - Absolute scalar adjectives: $Q_1, Q_2, Q_3 \dots$
4. For every unary predicate symbol P , there is a binary predicate $>_P$.
5. Quantifiers and connectives \forall, \vee and \neg , plus parentheses.

Definition 14. Syntax.

1. Variables and constants (and nothing else) are terms.
2. If t is a term and P is a predicate symbol, then $P(t)$ is a well-formed formula (wff).
3. If t_1 and t_2 are terms and P is a predicate symbol, then $t_1 >_P t_2$ is a wff.
4. For any variable x , if ϕ and ψ are wffs, then $\neg\phi$, $\phi \vee \psi$, and $\forall x\phi$ are wffs.
5. Nothing else is a wff.

Semantics

Definition 15. C(lassical)-model. *A c-model is a tuple $M = \langle D, m \rangle$ where D is a non-empty domain of individuals, and m is a function from pairs consisting of a member of the non-logical vocabulary and a comparison class (a subset of the domain) satisfying:*

- For each individual constant a_1 , $m(a_1) \in D$.
- For each $X \in \mathcal{P}(D)$ and for each predicate P , $m(P, X) \subseteq X$.

Definition 16. T(olerant)-model. *A t-model is a tuple $M = \langle D, m, \sim \rangle$, where $\langle D, m \rangle$ is a model and \sim is a function from predicate/comparison class pairs such that:*

- For all P and all $X \in \mathcal{P}(D)$, \sim_P^X is a binary relation on X that is reflexive, symmetric, but not necessarily transitive.

Definition 17. Assignment. *An assignment for a c/t-model M is a function $g : \{x_n : n \in \mathbb{N}\} \rightarrow D$ (from the set of variables to the domain D).*

Definition 18. Interpretation. *An interpretation $\llbracket \cdot \rrbracket_{M,g}$ is a pair $\langle M, g \rangle$, where M is a t-model, and g is an assignment.*

Definition 19. Interpretation of terms ($\llbracket \cdot \rrbracket_{M,g}$). For a model M , an assignment g ,

1. If x_1 is a variable, $\llbracket x_1 \rrbracket_{M,g} = g(x_1)$.
2. If a_1 is a constant, $\llbracket a_1 \rrbracket_{M,g} = m(a_1)$.

In what follows, for an interpretation $\llbracket \cdot \rrbracket_{M,g}$, a variable x_1 , and a constant a_1 , let $g[a_1/x_1]$ be the assignment for M which maps x_1 to a_1 , but agrees with g on all variables that are distinct from x_1 .

Definition 20. Classical Satisfaction ($\llbracket \cdot \rrbracket^c$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

1. $\llbracket P(t_1) \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in m(P, X) \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X - m(P, X) \\ i & \text{otherwise} \end{cases}$
2. $\llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^c = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^c = 0 \\ 0 & \text{otherwise} \end{cases}$
3. $\llbracket \neg \phi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \\ i & \text{otherwise} \end{cases}$
4. $\llbracket \phi \vee \psi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = \llbracket \psi \rrbracket_{M,g,X}^c = 0 \\ i & \text{otherwise} \end{cases}$
5. $\llbracket \forall x_1 \phi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 0 \\ i & \text{otherwise} \end{cases}$

Definition 21. Tolerant Satisfaction ($\llbracket \cdot \rrbracket^t$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

1. $\llbracket P(t_1) \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if there is some } a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} : \llbracket P(a_1) \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X, \text{ and there is no } a_1 \in X : a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} \\ i & \text{otherwise} \end{cases}$
2. $\llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^t = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^t = 0 \\ 0 & \text{otherwise} \end{cases}$
3. $\llbracket \neg \phi \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \\ i & \text{otherwise} \end{cases}$
4. $\llbracket \phi \vee \psi \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^t = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = \llbracket \psi \rrbracket_{M,g,X}^t = 0 \\ i & \text{otherwise} \end{cases}$

$$5. \llbracket \forall x_1 \phi \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 0 \\ i & \text{otherwise} \end{cases}$$

Definition 22. Strict Satisfaction($\llbracket \cdot \rrbracket^s$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

$$\begin{aligned} 1. \llbracket P(t_1) \rrbracket_{M,g,X}^s &= \begin{cases} 1 & \text{if for all } a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} : \llbracket P(a_1) \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X, \text{ and there is no } a_1 \in X : a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} \\ i & \text{otherwise} \end{cases} \\ 2. \llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^s &= \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^s = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^s = 0 \\ 0 & \text{otherwise} \end{cases} \\ 3. \llbracket \neg \phi \rrbracket_{M,g,X}^s &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \\ i & \text{otherwise} \end{cases} \\ 4. \llbracket \phi \vee \psi \rrbracket_{M,g,X}^s &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^s = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = \llbracket \psi \rrbracket_{M,g,X}^t = 0 \\ i & \text{otherwise} \end{cases} \\ 5. \llbracket \forall x_1 \phi \rrbracket_{M,g,X}^s &= \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 0 \\ i & \text{otherwise} \end{cases} \end{aligned}$$

Proposed Axioms for AAs ¹³

- (22) **Absolute Adjective Axiom (AAA)**: For all $X \in \mathcal{P}(D)$ and $a_1 \in X$, $\llbracket Q_1(a_1) \rrbracket_{M,g,X}^c = 1$ iff $\llbracket Q_1(a_1) \rrbracket_{M,g,D}^c = 1$.
- (23) **Tolerant No Skipping (T-NS)**: For an AA Q_1 , $X \in \mathcal{P}(D)$ and $a_1, a_2 \in X$, if $a_1 \sim_{Q_1}^X a_2$ and there is some $a_3 \in X$ such that $\llbracket a_1 \geq_{Q_1} a_3 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_3 \geq_{Q_1} a_2 \rrbracket_{M,g,X}^t = 1$, then $a_1 \sim_{Q_1}^X a_3$.
- (24) **Granularity 1 (G1)**: For an AA Q_1 , $X \in \mathcal{P}(D)$, and $a_1, a_2 \in X$, if $a_1 \sim_{Q_1}^X a_2$, then for all $X' \in \mathcal{P}(D) : X \subseteq X'$, $a_1 \sim_{Q_1}^{X'} a_2$.
- (25) **Granularity 2 (G2)**: For an AA Q_1 , $X, X' \in \mathcal{P}(D)$, and $a_1, a_2 \in X$, if $X \subset X'$ and $a_1 \not\sim_{Q_1}^X a_2$ and $a_1 \sim_{Q_1}^{X'} a_2$, then $\exists a_3 \in X' - X : a_1 \not\sim_{Q_1}^{X'} a_3$.
- (26) **Minimal Difference (MD)**: For an AA Q_1 and $a_1, a_2 \in D$, if $\llbracket a_1 >_{Q_1} a_2 \rrbracket_{M,g,X}^c = 1$, then $a_1 \not\sim_{Q_1}^{\{x,y\}} a_2$.

¹³ **Tolerantly greater than or equal**. (\geq^t) For an interpretation $\llbracket \cdot \rrbracket_{M,g,X}$, a predicate P , $a_1, a_2 \in D$, $a_1 \geq_P^t a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^t = 1$ or $a_1 \approx_P^t a_2$.

7.2 Proofs

Firstly, Minimal Difference ensures that classical absolute denotations are subsets of tolerant denotations:

Lemma 1. Tolerant Subset. *If $Q \in AA$, then, for all $X \subseteq D$, $a_1, a_2 \in D$, if $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^c$, then $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t$.*

Proof. Suppose $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^c$. Then, by definition 20, there is some $X' \subseteq D$ such that $\llbracket Q(a_1) \rrbracket_{M,g,X'}^c = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X'}^c = 0$. Now consider $\{a_1, a_2\}$. By downward difference, $\llbracket Q(a_1) \rrbracket_{M,g,\{a_1,a_2\}}^c = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,\{a_1,a_2\}}^c = 0$. By the definition of $\llbracket \cdot \rrbracket^t$, $\llbracket Q(a_1) \rrbracket_{M,g,\{a_1,a_2\}}^t = 1$. Furthermore, by Minimal Difference, $a_1 \not\sim_Q^{\{a_1,a_2\}} a_2$. So $\llbracket Q(a_2) \rrbracket_{M,g,\{a_1,a_2\}}^t = 0$. By definition 21, $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t$. \square

Secondly, with only T-No Skipping, we can prove that a version of van Benham's No Reversal holds at the tolerant level.

Lemma 2. Tolerant No Reversal (T-NR): *For $X \subseteq D$, and $a_1, a_2 \in D$ if $\llbracket Q(a_1) \rrbracket_{M,g,X}^t = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X}^t = 0$, then there is no $X' \subseteq D$ such that $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 1$ and $\llbracket Q(a_1) \rrbracket_{M,g,X'}^t = 0$.*

Proof. Suppose $\llbracket Q(a_1) \rrbracket_{M,g,X}^t = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X}^t = 0$. Suppose, for a contradiction that there is an $X' \subseteq D$ such that $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 1$ and $\llbracket Q(a_1) \rrbracket_{M,g,X'}^t = 0$. Therefore, $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_1 \rrbracket_{M,g,X}^t = 1$. Furthermore, by assumption and definition 21, there is some $a_3 \sim_Q^X a_1$ such that $\llbracket Q(a_3) \rrbracket_{M,g,X}^c = 1$, and $a_3 \not\sim_Q^X a_2$. Since $\llbracket Q(a_1) \rrbracket_{M,g,X'}^t = 0$, by the AAA, $\llbracket Q(a_1) \rrbracket_{M,g,X'}^c = 0$. So $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^c = 1$. By lemma 1, $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 1$, and so $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_1 \rrbracket_{M,g,X}^t = 1$. Since $a_3 \sim_Q^X a_1$, by No Skipping, $a_3 \sim_Q^X a_2$. \perp \square

Using the complete axiom set $\{\text{NS}, \text{G1}, \text{G2}, \text{MD}\}$, we can show that, for all $Q \in AA$, the tolerant comparative $(>_Q^t)$ is a strict weak order.

Lemma 3. Irreflexivity. *For all $X \subseteq D$ and $a_1 \in D$, $\llbracket a_1 >_Q a_1 \rrbracket_{M,g,X}^t = 0$.*

Proof. Since it is impossible, for any $X \subseteq D$, for an element to be both in $\llbracket Q \rrbracket_{M,g,X}^t$ and not in $\llbracket Q \rrbracket_{M,g,X}^t$, by definition 21, $>_Q^t$ is irreflexive. \square

Lemma 4. Transitivity. *For all $X \subseteq D$ and $a_1, a_2, a_3 \in D$, if $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_3 \rrbracket_{M,g,X}^t = 1$, then $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$.*

Proof. Suppose $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_3 \rrbracket_{M,g,X}^t = 1$ to show that $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. Then there is some $X' \subseteq D$ such that $\llbracket Q(a_1) \rrbracket_{M,g,X'}^t = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 0$. Thus, there is some $a_4 \in X' : \llbracket Q(a_4) \rrbracket_{M,g,X'}^c = 1$, and $a_4 \sim_Q^{X'} a_1$. Now consider $X' \cup \{a_3\}$. By the AAA and the assumption that $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_3 \rrbracket_{M,g,X}^t = 1$, $\llbracket Q(a_2) \rrbracket_{M,g,X \cup \{a_3\}}^c = 0$ and

$$\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 0.$$

Case 1: $X' \cup \{a_3\} = X'$. Since $\llbracket Q(a_3) \rrbracket_{M,g,X'}^t = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 0$, by theorem 2, $\llbracket Q(a_3) \rrbracket_{M,g,X'}^t = 0$, and $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. \checkmark **Case 2:** $X' \subset X' \cup \{a_3\}$. Since $X' \subset X' \cup \{a_3\}$ and $a_4 \sim_Q^X a_1$, by G1, $a_4 \sim_Q^{X' \cup \{a_3\}} a_1$. By the AAA, $\llbracket Q(a_4) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 1$. So $\llbracket Q(a_1) \rrbracket_{M,g,X \cup \{a_3\}}^t = 1$. Suppose, for a contradiction that $\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^t = 1$. Then there is some $a_5 \in X' \cup \{a_3\} : \llbracket Q(a_5) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 1$ and $\mathbf{a}_5 \sim_Q^{X' \cup \{a_3\}} \mathbf{a}_3$. By assumption and since $\llbracket Q(a_2) \rrbracket_{M,g,X'}^c$, by MD, $\llbracket a_5 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. So by T- No Skipping, $a_5 \sim_Q^{X' \cup \{a_3\}} a_2$. Since $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 0$, $a_5 \not\sim_Q^{X'} a_2$. So by G2, since $X' \cup \{a_3\} - X' = \{a_3\}$, $\mathbf{a}_5 \not\sim_Q^{X' \cup \{a_3\}} \mathbf{a}_3$. \perp . So $\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^t = 0$, and $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. \checkmark \square

Lemma 5. Almost Connectedness. For all $X \subseteq D$ and $a_1, a_2 \in D$, if $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ then for all $a_3 \in D$, either $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$ or $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 1$.

Proof. Let $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 0$ to show $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$.

Case 1: $\llbracket Q(a_1) \rrbracket_{M,g,D}^c = 1$. Since $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 0$, $\llbracket Q(a_3) \rrbracket_{M,g,D}^c = 0$. So $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^c = 1$, and, by lemma 1, $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. \checkmark **Case 2:** $\llbracket Q(a_1) \rrbracket_{M,g,D}^c = 0$. Since $\llbracket a_1 >_Q a_2 \rrbracket_{M,g,X}^t = 1$, there is some $X' \subseteq D$ such that $\llbracket Q(a_1) \rrbracket_{M,g,X'}^t = 1$ and $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 0$. So there is some $a_4 \in X' : \llbracket Q(a_4) \rrbracket_{M,g,X'}^c = 1$ and $a_4 \sim_Q^{X'} a_1$. Now consider $X' \cup \{a_3\}$. Since $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 0$, $\llbracket Q(a_1) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 0$ and $\llbracket Q(a_2) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 0$ and $\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 0$. Since $a_4 \sim_Q^{X'} a_1$, by G1, $a_4 \sim_Q^{X' \cup \{a_3\}} a_1$ and by the AAA, $\llbracket Q(a_4) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 1$. So, by definition 21, $\llbracket Q(a_1) \rrbracket_{M,g,X' \cup \{a_3\}}^t = 1$. Now suppose for a contradiction that $\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^t = 1$. Then there is some $a_5 \in X' \cup \{a_3\} : \llbracket Q(a_5) \rrbracket_{M,g,X' \cup \{a_3\}}^c = 1$ and $\mathbf{a}_5 \sim_Q^{X' \cup \{a_3\}} \mathbf{a}_3$. Since $\llbracket Q(a_5) \rrbracket_{X' \cup \{a_3\}}^c = 1$ and $\llbracket Q(a_2) \rrbracket_{X' \cup \{a_3\}}^c = 0$, $\llbracket a_5 >_Q a_2 \rrbracket_{M,g,X}^c = 1$; so by lemma 1, $\llbracket a_5 >_Q a_2 \rrbracket_{M,g,X}^t = 1$. Furthermore, since, by assumption, $\llbracket a_3 >_Q a_2 \rrbracket_{M,g,X}^t = 0$, $\llbracket a_2 \geq_Q a_3 \rrbracket_{M,g,X}^t = 1$. Since $\llbracket a_5 \geq_Q a_2 \rrbracket_{M,g,X}^t = 1$ and $\llbracket a_2 \geq_Q a_3 \rrbracket_{M,g,X}^t = 1$, and $a_5 \sim_Q^{X' \cup \{a_3\}} a_3$, by Tolerant No Skipping, $a_5 \sim_Q^{X' \cup \{a_3\}} a_2$. However, since $\llbracket Q(a_2) \rrbracket_{M,g,X'}^t = 0$, and by the AAA, $\llbracket Q(a_5) \rrbracket_{M,g,X'}^c = 1$, $a_5 \not\sim_Q^{X'} a_2$. Since $X' \subset X' \cup \{a_3\}$ and $a_5 \sim_Q^{X' \cup \{a_3\}} a_2$, by G2, there is some $a_6 \in X' \cup \{a_3\} - X'$ such that $a_6 \not\sim_Q^{X' \cup \{a_3\}} a_3$. Since $X' \cup \{a_3\} - X' = \{a_3\}$, $\mathbf{a}_5 \not\sim_Q^{X' \cup \{a_3\}} \mathbf{a}_3$. \perp So $\llbracket Q(a_3) \rrbracket_{M,g,X' \cup \{a_3\}}^t = 0$ and $\llbracket a_1 >_Q a_3 \rrbracket_{M,g,X}^t = 1$. \checkmark \square

We can now prove the main theorem of the paper:

Theorem 3. If Q is an absolute adjective, $<_Q^t$ is a strict weak order.

Proof. Immediate from lemmas 3, 4 and 5. \square

Are True and False Not Enough?*

Vincent Degauquier

Université catholique de Louvain, Institut supérieur de Philosophie
Place du Cardinal Mercier 14, 1348 Louvain-la-Neuve, Belgium
`vincent.degauquier@uclouvain.be`

Abstract. Trivalent logic is usually interpreted as a non-classical logic that violates the principle of bivalence. However, this logic can also be interpreted as a bivalent logic that ignores either the principle of completeness or the principle of consistency. In light of these bivalent interpretations, the three-valued interpretation can be understood either in a paraconsistent or paracomplete perspective. With this in mind, we investigate the semantic and proof-theoretic relationships between the paraconsistent and paracomplete conceptions of trivalent logic. More specifically, our purpose is to provide a unified framework for characterizing the notion of logical consequence specific to these conceptions. To do this, we propose a notion of validity and an associated hypersequent-inspired calculus for each of the two conceptions of trivalent logic.

Keywords: bivalence, completeness, consistency, logical consequence, sequent calculus.

Introduction

Logic is traditionally defined according to underlying principles. Among them, three seem particularly important. The principle of bivalence (understood in its etymological sense) says that there are exactly two truth values, usually called True and False. The principle of completeness states that a sentence has at least one truth value. The principle of consistency states that a sentence has at most one truth value. A logic that satisfies the conjunction of these three principles is called classical. By contrast, a logic is called non-classical if it does not obey at least one of them.

In relation to these principles, trivalent logic is usually interpreted as a non-classical logic that violates the principle of bivalence but requires the addition of the principles of completeness and consistency. However, this logic can also be interpreted as a bivalent logic that ignores exactly one of these last two principles. In light of the bivalent interpretations, the three-valued interpretation can be understood either in a paraconsistent or paracomplete perspective. Depending on which viewpoint is embraced, the meaning of the third truth value is different.

* This is an abridged and slightly modified version of the paper ‘Cuts, Gluts and Gaps’ due to appear in *Logique et Analyse*.

By means of these bivalent interpretations, we investigate the semantic and proof-theoretic relationships between the paraconsistent and paracomplete conceptions of trivalent logic. More specifically, our purpose is to provide a unified framework for characterizing the notion of logical consequence within each of these conceptions. To do this, we identify the core of the notion of logical consequence common to these two conceptions of trivalent logic.

The proof-theoretic approach we have chosen is sequent calculus. We propose a notion of validity and an associated hypersequent-inspired calculus for each of the two aforementioned conceptions of trivalent logic. In order to develop a theoretical framework broad enough to reach the core of the trivalent notion of logical consequence, we make use of a generalized notion of sequent closely related to that of hypersequent. In this way, we will identify the rules common to the paraconsistent and paracomplete trivalent calculi as well as the rules specific to each of them.

1 Language

A first-order predicate language \mathcal{L} is composed of a countable set of symbols consisting of a non-empty set of n -ary relation symbols, a set of n -ary function symbols, a countable set of variables and the usual logical symbols (\neg , \wedge , \vee , \rightarrow , \forall and \exists). The nullary function symbols are called constants and the nullary relation symbols are called propositional symbols.

As for syntax, the notions of term and formula are defined in the usual way. Nevertheless, formulas equivalent up to their bound variables are identified so that their bound variables are supposed to be Bourbaki's squares. In this way, any occurrence of a variable in a formula is free. The substitution operation is defined as follows: If A is a formula, $\alpha_1, \dots, \alpha_n$ are distinct variables and t_1, \dots, t_n are terms, then $A[\alpha_1 := t_1, \dots, \alpha_n := t_n]$ denotes the formula resulting from the simultaneous substitution of t_i for α_i in A , for all i ($1 \leq i \leq n$).

2 Bivalent logics

Three bivalent logics differ from classical logic insofar as they ignore the principle of completeness and/or the principle of consistency: consistent logic (which is closely related to Kleene's strong three-valued logic) satisfies the principle of consistency, complete logic (which is closely related to Priest's logic of paradox) satisfies the principle of completeness and positive logic (which is closely related to Belnap's useful four-valued logic) ignores both of these principles. In addition to classical logic, three bivalent (non-classical) logics can therefore be distinguished.

A *bivalent model* \mathcal{M} (simply called *model* hereafter) for a language \mathcal{L} is composed of a structure for \mathcal{L} and an interpretation of the proper symbols of \mathcal{L} in this structure (see [4]).

A *structure* for \mathcal{L} consists of a universe, a set of relations on this universe and a set of functions defined on this universe and with values in this universe.

For every $n \in \mathbb{N}$, if \mathcal{L} has n -ary relation symbols, the structure must have at least one n -ary relation. For every $n \in \mathbb{N}$, if \mathcal{L} has n -ary function symbols, the structure must have at least one n -ary function.

The universe $|\mathcal{M}|$ of a model \mathcal{M} is a non-empty set. An n -ary relation R is an ordered pair of subsets of $|\mathcal{M}|^n$ such that $R = \langle R^+, R^- \rangle$. The first term of the ordered pair denotes the set of n -tuples of elements of the universe that verify the relation R and the second term of the ordered pair denotes the set of n -tuples of elements of the universe that falsify the relation.

An *interpretation* of \mathcal{L} assigns an object in the universe to every constant, an n -ary function defined on the universe to every n -ary function symbol of \mathcal{L} and an n -ary relation to every n -ary relation symbol of \mathcal{L} . The interpretation of an n -ary relation symbol R of \mathcal{L} in the universe of the model \mathcal{M} is denoted $R_{\mathcal{M}}$ and is equated to the ordered pair $\langle (R^n)_{\mathcal{M}}^+, (R^n)_{\mathcal{M}}^- \rangle$ of subsets of $|\mathcal{M}|^n$.

A *valuation* is an assignment of objects to variables. If v is a valuation, $\vec{\alpha}$ is a sequence of distinct variables $\alpha_1, \dots, \alpha_n$ and \vec{o} is a sequence of elements o_1, \dots, o_n in $|\mathcal{M}|$, then $v[\vec{\alpha} \mapsto \vec{o}]$ is the valuation that differs from v insofar as the variable α_i denotes the element o_i , for all i ($1 \leq i \leq n$). More specifically, the valuation $v[\vec{\alpha} \mapsto \vec{o}]$ is defined as follows:

- $v[\vec{\alpha} \mapsto \vec{o}](\beta) = o_i$, if β is the same variable as α_i ($1 \leq i \leq n$).
- $v[\vec{\alpha} \mapsto \vec{o}](\beta) = v(\beta)$, if β is distinct from every α_i ($1 \leq i \leq n$).

By combining a valuation v and an interpretation of constants and function symbols in \mathcal{M} , all terms of the language are given a value. The joint extension of interpretation and valuation is denoted $v_{\mathcal{M}}$. Moreover, if \vec{t} is a sequence of terms t_1, \dots, t_n , then $v_{\mathcal{M}}(\vec{t})$ denotes the sequence $v_{\mathcal{M}}(t_1), \dots, v_{\mathcal{M}}(t_n)$. It is required that:

- $v_{\mathcal{M}}(\beta) = v(\beta)$, for every variable β .
- $v_{\mathcal{M}}(F\vec{t}) = F_{\mathcal{M}}(v_{\mathcal{M}}(\vec{t}))$, for every n -ary function symbol F .

Truth and *falsity* of a formula A are defined in a model under a valuation. Given a model \mathcal{M} and a valuation v , truth (denoted by $\mathcal{M} \models_v^+ A$) and falsity (denoted by $\mathcal{M} \models_v^- A$) of formulas of the language are defined inductively:

- $\mathcal{M} \models_v^+ R t_1 \dots t_n$ if and only if $\langle v_{\mathcal{M}}(t_1), \dots, v_{\mathcal{M}}(t_n) \rangle \in R_{\mathcal{M}}^+$
- $\mathcal{M} \models_v^- R t_1 \dots t_n$ if and only if $\langle v_{\mathcal{M}}(t_1), \dots, v_{\mathcal{M}}(t_n) \rangle \in R_{\mathcal{M}}^-$
- $\mathcal{M} \models_v^+ \neg A$ if and only if $\mathcal{M} \models_v^- A$
- $\mathcal{M} \models_v^- \neg A$ if and only if $\mathcal{M} \models_v^+ A$
- $\mathcal{M} \models_v^+ (A \wedge B)$ if and only if $\mathcal{M} \models_v^+ A$ and $\mathcal{M} \models_v^+ B$
- $\mathcal{M} \models_v^- (A \wedge B)$ if and only if $\mathcal{M} \models_v^- A$ and/or $\mathcal{M} \models_v^- B$
- $\mathcal{M} \models_v^+ (A \vee B)$ if and only if $\mathcal{M} \models_v^+ A$ and/or $\mathcal{M} \models_v^+ B$
- $\mathcal{M} \models_v^- (A \vee B)$ if and only if $\mathcal{M} \models_v^- A$ and $\mathcal{M} \models_v^- B$
- $\mathcal{M} \models_v^+ (A \rightarrow B)$ if and only if $\mathcal{M} \models_v^- A$ and/or $\mathcal{M} \models_v^+ B$
- $\mathcal{M} \models_v^- (A \rightarrow B)$ if and only if $\mathcal{M} \models_v^+ A$ and $\mathcal{M} \models_v^- B$
- $\mathcal{M} \models_v^+ \forall \alpha A$ if and only if $\mathcal{M} \models_{v[\alpha \mapsto o]}^+ A$, for all $o \in |\mathcal{M}|$
- $\mathcal{M} \models_v^- \forall \alpha A$ if and only if $\mathcal{M} \models_{v[\alpha \mapsto o]}^- A$, for some $o \in |\mathcal{M}|$

$$\begin{aligned} \mathcal{M} \models_v^+ \exists \alpha A \text{ if and only if } \mathcal{M} \models_{v[\alpha \rightarrow o]}^+ A, \text{ for some } o \in |\mathcal{M}| \\ \mathcal{M} \models_v^- \exists \alpha A \text{ if and only if } \mathcal{M} \models_{v[\alpha \rightarrow o]}^- A, \text{ for all } o \in |\mathcal{M}| \end{aligned}$$

Remark 1. Although the notation might suggest it, the definitions of the connectives listed above are not equivalent to the usual definitions. By distinguishing truth from non-falsity and falsity from non-truth, the positive definitions of truth and falsity make the traditional definitions of logical connectives ambiguous. While it does not exist in classical logic, this ambiguity must be removed in a non-classical bivalent logic. It is therefore possible to propose several positive definitions corresponding to the classical definition of a connective but none of them fit perfectly with it. This translation problem is particularly acute in the cases of negation and implication.

A model \mathcal{M} is *consistent* if and only if $(R^n)_{\mathcal{M}}^+ \cap (R^n)_{\mathcal{M}}^- = \emptyset$, for every n -ary relation R on $|\mathcal{M}|$. A model \mathcal{M} is *complete* if and only if $(R^n)_{\mathcal{M}}^+ \cup (R^n)_{\mathcal{M}}^- = |\mathcal{M}|^n$, for every n -ary relation R on $|\mathcal{M}|$. In this sense, a model is called *classical* if and only if it is both consistent and complete.

Fact 1 (Principle of completeness). *Let \mathcal{M} be a complete model.*

1. *if $\mathcal{M} \not\models_v^+ A$, then $\mathcal{M} \models_v^- A$, for all formulas A .*
2. *if $\mathcal{M} \not\models_v^- A$, then $\mathcal{M} \models_v^+ A$, for all formulas A .*

Fact 2 (Principle of consistency). *Let \mathcal{M} be a consistent model.*

1. *if $\mathcal{M} \models_v^+ A$, then $\mathcal{M} \not\models_v^- A$, for all formulas A .*
2. *if $\mathcal{M} \models_v^- A$, then $\mathcal{M} \not\models_v^+ A$, for all formulas A .*

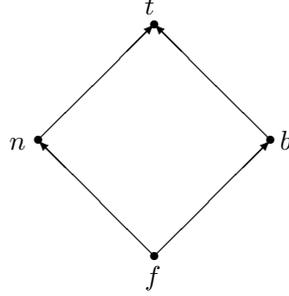
Depending on whether a bivalent logic restricts the class of models to that of consistent, complete or classical models, this logic will be called consistent, complete or classical, respectively. In general, bivalent logic that takes into account the class of models without restriction is called positive.

3 Towards a third truth value

Starting with a bivalent interpretation, the development of trivalent logic does not simply consist in adding a third truth value. Instead, it consists in giving new definitions of truth and falsity. In general, we can say that a formula is true in a many-valued logic if and only if it is both true and not false in bivalent logic. Similarly, we can say that a formula is false in a many-valued logic if and only if it is both false and not true in bivalent logic.

As for the third truth value, two options are possible from the bivalent viewpoint. The consistent interpretation suggests that the third truth value in trivalent logic is assigned to a formula if and only if it is neither true nor false in bivalent logic (see [7]). On the other hand, the complete interpretation suggests that the third truth value in trivalent logic is assigned to a formula if and only if it is both true and false in bivalent logic (see [9]).

This choice concerning the third truth value highlights the distinction between paracomplete and paraconsistent interpretations of trivalent logic. In light of the bivalent interpretation, four new truth values can therefore be defined (see [2]). Let $\langle \{t, f, n, b\}, \leq \rangle$ be a lattice such that:



Using the lattice above and assuming that $\bar{t} = f$, $\bar{f} = t$, $\bar{n} = n$ and $\bar{b} = b$, an inductive definition of a many-valued model can be provided. A *many-valued translation* of a model \mathcal{M} under a valuation v for a language \mathcal{L} , denoted by \mathcal{M}_v^T , is a function from the set of formulas of \mathcal{L} to the set $\{t, f, n, b\}$ such that:

$$\mathcal{M}_v^T(Rt_1 \dots t_n) = t \text{ if and only if } \mathcal{M} \models_v^+ Rt_1 \dots t_n \text{ and } \mathcal{M} \not\models_v^- Rt_1 \dots t_n$$

$$\mathcal{M}_v^T(Rt_1 \dots t_n) = f \text{ if and only if } \mathcal{M} \not\models_v^+ Rt_1 \dots t_n \text{ and } \mathcal{M} \models_v^- Rt_1 \dots t_n$$

$$\mathcal{M}_v^T(Rt_1 \dots t_n) = n \text{ if and only if } \mathcal{M} \not\models_v^+ Rt_1 \dots t_n \text{ and } \mathcal{M} \not\models_v^- Rt_1 \dots t_n$$

$$\mathcal{M}_v^T(Rt_1 \dots t_n) = b \text{ if and only if } \mathcal{M} \models_v^+ Rt_1 \dots t_n \text{ and } \mathcal{M} \models_v^- Rt_1 \dots t_n$$

$$\mathcal{M}_v^T(\neg A) = \overline{\mathcal{M}_v^T(A)}$$

$$\mathcal{M}_v^T(A \wedge B) = \inf(\{\mathcal{M}_v^T(A), \mathcal{M}_v^T(B)\})$$

$$\mathcal{M}_v^T(A \vee B) = \sup(\{\mathcal{M}_v^T(A), \mathcal{M}_v^T(B)\})$$

$$\mathcal{M}_v^T(A \rightarrow B) = \sup(\{\mathcal{M}_v^T(A), \mathcal{M}_v^T(B)\})$$

$$\mathcal{M}_v^T(\forall \alpha A) = \inf(\{\mathcal{M}_{v[\alpha \mapsto o]}^T(A) \mid o \in |\mathcal{M}|\})$$

$$\mathcal{M}_v^T(\exists \alpha A) = \sup(\{\mathcal{M}_{v[\alpha \mapsto o]}^T(A) \mid o \in |\mathcal{M}|\})$$

The following proposition establishes the correspondence between the bivalent interpretation and the many-valued interpretation of truth values.

Proposition 1. *For all formulas A :*

1. $\mathcal{M}_v^T(A) = t$ if and only if $\mathcal{M} \models_v^+ A$ and $\mathcal{M} \not\models_v^- A$
2. $\mathcal{M}_v^T(A) = f$ if and only if $\mathcal{M} \not\models_v^+ A$ and $\mathcal{M} \models_v^- A$
3. $\mathcal{M}_v^T(A) = n$ if and only if $\mathcal{M} \not\models_v^+ A$ and $\mathcal{M} \not\models_v^- A$
4. $\mathcal{M}_v^T(A) = b$ if and only if $\mathcal{M} \models_v^+ A$ and $\mathcal{M} \models_v^- A$

The many-valued translation of a model is called *consistent*, *complete* or *classical* if and only if its image is $\{t, f, n\}$, $\{t, f, b\}$ or $\{t, f\}$, respectively. From proposition 1 and facts 1 and 2, it immediately follows that the many-valued translation of a model \mathcal{M} is consistent, complete or classical if and only if \mathcal{M} is consistent, complete or classical, respectively.

For this reason, we will apply the term ‘positive logic’, without distinction, to the bivalent logic that takes into account the class of models without restriction and to the many-valued logic that takes into account the class of many-valued translations without restriction. Similarly, a many-valued logic that restricts the class of many-valued translations to that of consistent, complete or classical translations will simply be called consistent, complete or classical logic, respectively.

4 Sequent calculi

A sequent is called valid, glut-valid, gap-valid and classic-valid if it is semantically correct in positive logic, complete logic, consistent logic and classical logic, respectively. Similarly, a sequent is called derivable, glut-derivable, gap-derivable and classic-derivable if it is proof-theoretically correct in positive logic, complete logic, consistent logic and classical logic, respectively.

A *sequent* is a quadruple $\langle \Pi, \Gamma, \Delta, \Sigma \rangle$, where Π , Γ , Δ and Σ are finite multisets over the set of formulas of the language (see [1], [6] and [8]). The sequent $\langle \Pi, \Gamma, \Delta, \Sigma \rangle$ is denoted $\Pi; \Gamma \vdash \Delta; \Sigma$. A multiset is a sequence modulo the ordering.

More specifically, a *multiset* M over S is an ordered pair $\langle S, f \rangle$, where S is a set and $f : S \rightarrow \mathbb{N}$ is a function that indicates the multiplicity of each element of S . The *underlying set* of a multiset $M = \langle S, f \rangle$ is the set μ such that $\mu = \{s \in S \mid f(s) \neq 0\}$. M is called *finite*, if μ is finite, and M is called *empty*, if μ is empty. The notation $M \neq \emptyset$ means that M is not empty.

Let M_1 and M_2 be multisets such that $M_1 = \langle S, f_1 \rangle$ and $M_2 = \langle S, f_2 \rangle$. M is the *intersection* of M_1 and M_2 , denoted $M_1 \cap M_2$, if $M = \langle S, f \rangle$ is a multiset, where $f(s) = f_1(s)$, if $f_1(s) \leq f_2(s)$, and $f(s) = f_2(s)$, if $f_2(s) \leq f_1(s)$, for all $s \in S$. M is the *union* of M_1 and M_2 , denoted $M_1 \cup M_2$, if $M = \langle S, f \rangle$ is a multiset, where $f(s) = f_1(s) + f_2(s)$, for all $s \in S$. The multisets $M_1 \cup M_2$ and $\langle S, f \rangle$, where $\{s \in S \mid f(s) \neq 0\} = \{A\}$ and $f(A) = 1$, are denoted M_1 , M_2 and A , respectively.

Let $\Pi; \Gamma \vdash \Delta; \Sigma$ be a sequent such that π , γ , δ and σ are the underlying sets of Π , Γ , Δ and Σ , respectively. Then, $\Pi; \Gamma \vdash \Delta; \Sigma$ is *valid* if and only if for every model \mathcal{M} and valuation v , $\mathcal{M} \not\vdash_v^- A$, for all $A \in \pi$, and $\mathcal{M} \vdash_v^+ A$, for all $A \in \gamma$, implies $\mathcal{M} \vdash_v^+ A$, for some $A \in \delta$, and/or $\mathcal{M} \not\vdash_v^- A$, for some $A \in \sigma$.

The definition of validity can be preserved for consistent and/or complete logics. Depending on whether the notion of valid sequent is restricted to consistent models or to complete models, a sequent is called *gap-valid* or *glut-valid*, respectively. If only the class of models which are both consistent and complete is taken into account, then a sequent is called *classic-valid*.

For each bivalent logic, a sequent calculus and a notion of derivability corresponding to that of validity are now set out. (A completeness proof for these sequent calculi is provided in [3].) The rules for these sequent calculi are as follows. The main feature of these rules is that the weakening and contraction structural rules are absorbed into the rules of inference (see [3]).

$$\begin{array}{c}
\frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma}{\Pi; \Gamma, \neg A \Vdash \Delta; \Sigma} \neg_L^i \\
\frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma}{\Pi, \neg A; \Gamma \Vdash \Delta; \Sigma} \neg_L^e \\
\frac{\Pi; \Gamma, A, B \Vdash \Delta; \Sigma}{\Pi; \Gamma, (A \wedge B) \Vdash \Delta; \Sigma} \wedge_L^i \\
\frac{\Pi, A, B; \Gamma \Vdash \Delta; \Sigma}{\Pi, (A \wedge B); \Gamma \Vdash \Delta; \Sigma} \wedge_L^e \\
\frac{\Pi; \Gamma, A \Vdash \Delta; \Sigma \quad \Pi; \Gamma, B \Vdash \Delta; \Sigma}{\Pi; \Gamma, (A \vee B) \Vdash \Delta; \Sigma} \vee_L^i \\
\frac{\Pi, A; \Gamma \Vdash \Delta; \Sigma \quad \Pi, B; \Gamma \Vdash \Delta; \Sigma}{\Pi, (A \vee B); \Gamma \Vdash \Delta; \Sigma} \vee_L^e \\
\frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma \quad \Pi; \Gamma, B \Vdash \Delta; \Sigma}{\Pi; \Gamma, (A \rightarrow B) \Vdash \Delta; \Sigma} \rightarrow_L^i \\
\frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma \quad \Pi, B; \Gamma \Vdash \Delta; \Sigma}{\Pi, (A \rightarrow B); \Gamma \Vdash \Delta; \Sigma} \rightarrow_L^e \\
\frac{\Pi; \forall \alpha A, \Gamma, A[\alpha := t] \Vdash \Delta; \Sigma}{\Pi; \Gamma, \forall \alpha A \Vdash \Delta; \Sigma} \forall_L^i \\
\frac{\forall \alpha A, \Pi, A[\alpha := t]; \Gamma \Vdash \Delta; \Sigma}{\Pi, \forall \alpha A; \Gamma \Vdash \Delta; \Sigma} \forall_L^e \\
\frac{\Pi; \Gamma, A[\alpha := \beta] \Vdash \Delta; \Sigma}{\Pi; \Gamma, \exists \alpha A \Vdash \Delta; \Sigma} \exists_L^i \\
\frac{\Pi, A[\alpha := \beta]; \Gamma \Vdash \Delta; \Sigma}{\Pi, \exists \alpha A; \Gamma \Vdash \Delta; \Sigma} \exists_L^e \\
\frac{\Pi, A; \Gamma \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \neg A, \Delta; \Sigma} \neg_R^i \\
\frac{\Pi; \Gamma, A \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; \neg A, \Sigma} \neg_R^e \\
\frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma \quad \Pi; \Gamma \Vdash B, \Delta; \Sigma}{\Pi; \Gamma \Vdash (A \wedge B), \Delta; \Sigma} \wedge_R^i \\
\frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma \quad \Pi; \Gamma \Vdash \Delta; B, \Sigma}{\Pi; \Gamma \Vdash \Delta; (A \wedge B), \Sigma} \wedge_R^e \\
\frac{\Pi; \Gamma \Vdash A, B, \Delta; \Sigma}{\Pi; \Gamma \Vdash (A \vee B), \Delta; \Sigma} \vee_R^i \\
\frac{\Pi; \Gamma \Vdash \Delta; A, B, \Sigma}{\Pi; \Gamma \Vdash \Delta; (A \vee B), \Sigma} \vee_R^e \\
\frac{\Pi, A; \Gamma \Vdash B, \Delta; \Sigma}{\Pi; \Gamma \Vdash (A \rightarrow B), \Delta; \Sigma} \rightarrow_R^i \\
\frac{\Pi; \Gamma, A \Vdash \Delta; B, \Sigma}{\Pi; \Gamma \Vdash \Delta; (A \rightarrow B), \Sigma} \rightarrow_R^e \\
\frac{\Pi; \Gamma \Vdash A[\alpha := \beta], \Delta; \Sigma}{\Pi; \Gamma \Vdash \forall \alpha A, \Delta; \Sigma} \forall_R^i \\
\frac{\Pi; \Gamma \Vdash \Delta; A[\alpha := \beta], \Sigma}{\Pi; \Gamma \Vdash \Delta; \forall \alpha A, \Sigma} \forall_R^e \\
\frac{\Pi; \Gamma \Vdash A[\alpha := t], \Delta, \exists \alpha A; \Sigma}{\Pi; \Gamma \Vdash \exists \alpha A, \Delta; \Sigma} \exists_R^i \\
\frac{\Pi; \Gamma \Vdash \Delta; A[\alpha := t], \Sigma, \exists \alpha A}{\Pi; \Gamma \Vdash \Delta; \exists \alpha A, \Sigma} \exists_R^e
\end{array}$$

The usual restrictions for the \forall_R^i , \forall_R^e , \exists_L^i and \exists_L^e rules hold. The *eigenvariable* β must not appear in the conclusion of these rules.

The notion of *derivation* as well as those of *initial sequent* and *endsequent* are defined inductively in the usual way. Roughly speaking, a derivation is a finite rooted tree in which the nodes are sequents. The root of the tree (at the bottom) is called the endsequent and the leaves of the tree (at the top) are called initial sequents. The *length* of a derivation is the number of sequents in that derivation.

A sequent is *derivable* if and only if there exists a derivation in which it is the endsequent and all initial sequents are axiomatic. A sequent $\Pi; \Gamma \Vdash \Delta; \Sigma$ is *axiomatic* if and only if $\Gamma \cap \Delta \neq \emptyset$ and/or $\Pi \cap \Sigma \neq \emptyset$.

The definition of derivable sequent can be preserved for consistent and/or complete logics by changing the definition of axiomatic sequent.

A sequent is *gap-derivable* if and only if there exists a derivation in which it is the endsequent and all initial sequents are gap-axiomatic. A sequent $\Pi; \Gamma \Vdash \Delta; \Sigma$ is *gap-axiomatic* if and only if it is axiomatic and/or $\Gamma \cap \Sigma \neq \emptyset$.

A sequent is *glut-derivable* if and only if there exists a derivation in which it is the endsequent and all initial sequents are glut-axiomatic. A sequent $\Pi; \Gamma \Vdash \Delta; \Sigma$ is *glut-axiomatic* if and only if it is axiomatic and/or $\Pi \cap \Delta \neq \emptyset$.

Finally, a sequent is *classic-derivable* if and only if there exists a derivation in which it is the endsequent and all initial sequents are classic-axiomatic. A sequent $\Pi; \Gamma \Vdash \Delta; \Sigma$ is *classic-axiomatic* if and only if it is gap-axiomatic and/or glut-axiomatic.

5 Unifying logical consequence

Several formulations of the redundancy of cut are possible in the sequent calculi mentioned above. Indeed, four different forms of cut are distinguishable. Only two hold for positive sequent calculus while all of them hold for classical sequent calculus. As for complete and consistent sequent calculi, they only admit one form of cut in addition to the two that hold for positive sequent calculus. These results are contained in theorems 1 and 2. Only a sketch of the proofs is given here (for details, see [3]).

Theorem 1. *For all formulas A :*

1. *if $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ and $\Pi; \Gamma, A \Vdash \Delta; \Sigma$ are derivable, then $\Pi; \Gamma \Vdash \Delta; \Sigma$ is derivable.*
2. *if $\Pi; \Gamma \Vdash \Delta; A, \Sigma$ and $\Pi, A; \Gamma \Vdash \Delta; \Sigma$ are derivable, then $\Pi; \Gamma \Vdash \Delta; \Sigma$ is derivable.*

Proof. The proof of the first assertion proceeds by a main induction on the complexity of A . When A is an atomic formula or when A is a quantified formula of the form $\exists\beta B$ or $\forall\beta B$, the proof uses a subinduction on the sum of the derivation lengths of the sequents $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ and $\Pi; \Gamma, A \Vdash \Delta; \Sigma$. The weakening and contraction structural properties as well as the inversion property of inference rules are presupposed as proved. The second assertion is treated symmetrically. \square

Theorem 2. *For all formulas A :*

1. *if $\Pi; \Gamma \Vdash \Delta; A, \Sigma$ and $\Pi; \Gamma, A \Vdash \Delta; \Sigma$ are glut-derivable, then $\Pi; \Gamma \Vdash \Delta; \Sigma$ is glut-derivable.*
2. *if $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ and $\Pi, A; \Gamma \Vdash \Delta; \Sigma$ are gap-derivable, then $\Pi; \Gamma \Vdash \Delta; \Sigma$ is gap-derivable.*

Proof. The proof is similar to that of theorem 1. \square

According to the position of the cut formula, four different forms of the original cut rule can be distinguished (see [5]).

$$\frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma \quad \Pi, A; \Gamma \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; \Sigma} \text{cut}_{e-e}$$

$$\frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma \quad \Pi; \Gamma, A \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; \Sigma} \text{cut}_{i-i}$$

$$\frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma \quad \Pi; \Gamma, A \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; \Sigma} \text{cut}_{e-i}$$

$$\frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma \quad \Pi, A; \Gamma \Vdash \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; \Sigma} \text{cut}_{i-e}$$

In view of theorems 1 and 2, it is interesting to note that the cut_{e-e} and cut_{i-i} rules preserve derivability, glut-derivability, gap-derivability and classic-derivability. By contrast, the cut_{e-i} and cut_{i-e} rules do not preserve derivability. In addition, the cut_{e-i} rule does not preserve gap-derivability and the cut_{i-e} rule does not preserve glut-derivability. In other words, positive sequent calculus admits only cut_{e-e} and cut_{i-i} , while complete and consistent sequent calculi admit, in addition to these rules, the cut_{e-i} and cut_{i-e} rules, respectively. As for classical sequent calculus, it admits the four cut rules.

Using the definition of axiomatic sequent, the weakening property and theorem 2, the following propositions can be easily proved. Proposition 4 means that $\Pi; \Gamma \Vdash \Delta; \Sigma$ is a classic-derivable sequent if and only if $\Pi, \Gamma \Vdash \Delta, \Sigma$ is deducible in a classical sequent calculus of the usual kind.

Proposition 2. *For all formulas A:*

1. *if $\Pi; \Gamma, A \Vdash \Delta; \Sigma$ is glut-derivable, then $\Pi, A; \Gamma \Vdash \Delta; \Sigma$ is glut-derivable.*
2. *if $\Pi; \Gamma \Vdash \Delta; A, \Sigma$ is glut-derivable, then $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ is glut-derivable.*

Proposition 3. *For all formulas A:*

1. *if $\Pi, A; \Gamma \Vdash \Delta; \Sigma$ is gap-derivable, then $\Pi; \Gamma, A \Vdash \Delta; \Sigma$ is gap-derivable.*
2. *if $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ is gap-derivable, then $\Pi; \Gamma \Vdash \Delta; A, \Sigma$ is gap-derivable.*

Proposition 4. *For all formulas A:*

1. *$\Pi; \Gamma, A \Vdash \Delta; \Sigma$ is classic-derivable if and only if $\Pi, A; \Gamma \Vdash \Delta; \Sigma$ is classic-derivable.*
2. *$\Pi; \Gamma \Vdash \Delta; A, \Sigma$ is classic-derivable if and only if $\Pi; \Gamma \Vdash A, \Delta; \Sigma$ is classic-derivable.*

These properties can be expressed using the following rules.

$$\frac{\Pi, A; \Gamma \Vdash \Delta; \Sigma}{\Pi; \Gamma, A \Vdash \Delta; \Sigma} \rightarrow_L \qquad \frac{\Pi; \Gamma \Vdash A, \Delta; \Sigma}{\Pi; \Gamma \Vdash \Delta; A, \Sigma} \rightarrow_R$$

$$\frac{\Pi; \Gamma, A \Vdash \Delta; \Sigma}{\Pi, A; \Gamma \Vdash \Delta; \Sigma} \leftarrow_L \qquad \frac{\Pi; \Gamma \Vdash \Delta; A, \Sigma}{\Pi; \Gamma \Vdash A, \Delta; \Sigma} \leftarrow_R$$

Propositions 2 and 3 assert, respectively, the admissibility of the $\leftarrow_{L/R}$ rules in complete sequent calculus and the admissibility of the $\rightarrow_{L/R}$ rules in consistent sequent calculus. By proposition 4, each of these rules is admissible in classical sequent calculus. However, the $\rightarrow_{L/R}$ rules do not hold for complete sequent calculus and the $\leftarrow_{L/R}$ rules do not hold for consistent sequent calculus. Moreover, none of them are admissible in positive sequent calculus.

Proposition 5. *Let $\Pi; \Gamma \Vdash \Delta; \Sigma$ be a sequent.*

1. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is glut-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rule cut_{e-i} .
2. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is gap-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rule cut_{i-e} .
3. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is classic-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rules cut_{e-i} and cut_{i-e} .

Proposition 6. *Let $\Pi; \Gamma \Vdash \Delta; \Sigma$ be a sequent.*

1. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is glut-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rules $\leftarrow_{L/R}$.
2. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is gap-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rules $\rightarrow_{L/R}$.
3. $\Pi; \Gamma \Vdash \Delta; \Sigma$ is classic-derivable if and only if $\Pi; \Gamma \Vdash \Delta; \Sigma$ is provable in positive sequent calculus plus the additional rules $\leftarrow_{L/R}$ and $\rightarrow_{L/R}$.

The foregoing propositions provide two different characterizations of the notions of glut-derivability, gap-derivability and classic-derivability obtained from the more general notion of derivability by adding rules. Proposition 5 underlines the crucial part played by the cut properties in characterizing the notion of logical consequence within complete and/or consistent sequent calculi. Proposition 6 makes the underlying principles of completeness and consistency explicit in the definitions of glut-axiomatic, gap-axiomatic and classic-axiomatic sequent. These two characterizations of the consistent and/or complete notions of logical consequence not only allow us to identify the common thread between these notions, but also allow us to better understand the relationships between the notions of cut, completeness and consistency.

References

1. Avron, A.: A constructive analysis of **RM**. *The Journal of Symbolic Logic* **52**(4) (1987) 939–951
2. Belnap, N.D.: A useful four-valued logic. In Dunn, J.M., Epstein, G., eds.: *Modern Uses of Multiple-Valued Logic*. Reidel (1977) 8–37
3. Degauquier, V.: *Recherches sur la bivalence*. PhD thesis, Université catholique de Louvain (2011)
4. Dunn, J.M.: Intuitive semantics for first-degree entailments and ‘coupled trees’. *Philosophical Studies* **29**(3) (1976) 149–168
5. Gentzen, G.: Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift* **39**(1) (1935) 176–210
6. Girard, J.Y.: Three-valued logic and cut-elimination : the actual meaning of Takeuti’s conjecture. *Dissertationes Mathematicae (Rozprawy Matematyczne)* **136** (1976) 1–49
7. Kleene, S.C.: *Introduction to metamathematics*. Van Nostrand, New York (1952)
8. Muskens, R.: On partial and paraconsistent logics. *Notre Dame Journal of Formal Logic* **40**(3) (1999) 352–374
9. Priest, G.: The logic of paradox. *Journal of Philosophical Logic* **8**(1) (1979) 219–241

Modeling the Suppression Task under Three-Valued Lukasiewicz and Well-Founded Semantics

Emmanuelle-Anna Dietz,* Steffen Hölldobler**

International Center for Computational Logic, Technische Universität Dresden
D-01062 Dresden, Germany

Abstract. Two three-valued approaches for modeling human conditional reasoning are presented. The suppression task, a study from psychological reasoning, formalized by Stenning and van Lambalgen, is examined. It is known that the suppression task is adequately solved using logic programs under the three-valued Lukasiewicz semantics. In this paper we examine this approach and compare it with well-founded semantics. We show that, by some modification, both approaches yield to identical results for a specific class of programs.

1 Introduction

Three-valued logics seem to be adequate for modeling some aspects of human reasoning as shown by Stenning and van Lambalgen [1]. They analyze the suppression task, a well-known experiment within cognitive science. In this psychological study, Byrne [2] has shown that graduate students with no previous exposure to formal logic did suppress previously drawn conclusions when additional information became available. Interestingly, in some instances the previously drawn conclusions were valid whereas in other instances the conclusions were invalid with respect to classical two-valued logic. Consider the following example: *If she has an essay to finish then she will study late in the library* and *She has an essay to finish*. Most subjects (96%) conclude: *She will study late in the library*. If subjects, however, receive an additional conditional: *If the library stays open she will study late in the library* then only 38% of them conclude: *She will study late in the library*. This shows, that, although the conclusion is still correct, the conclusion is suppressed by an additional conditional. This is an excellent example for human capability to draw *non-monotonic* inferences. Table 1 shows the abbreviations that will be used in this paper. The participants received conditionals (A, B, C) and facts $E, \neg E, L, \neg L$ from which they had to draw inferences. Table 2 gives an account of the findings of the first part of the suppression task.

Computational approaches modeling human reasoning must be classified regarding cognitive adequacy. We distinguish between conceptual and inferential measures. In our context, a system is *conceptually adequate* if it appropriately rep-

* dietz@iccl.tu-dresden.de

** sh@iccl.tu-dresden.de

Table 1. The suppression task and used abbreviations.

<i>A</i>	<i>If she has an essay to finish then she will study late in the library.</i>
<i>B</i>	<i>If she has a textbook to read then she will study late in the library.</i>
<i>C</i>	<i>If the library stays open she will study late in the library.</i>
<i>E</i>	<i>She has an essay to finish.</i>
$\neg E$	<i>She does not have an essay to finish.</i>
<i>L</i>	<i>She will study late in the library.</i>
$\neg L$	<i>She will not study late in the library.</i>

Table 2. The drawn conclusions of the experiment of Byrne [2].

Conditional(s)	Fact	Exp.	Findings
<i>A</i>	<i>E</i>	96%	conclude <i>L</i> .
<i>A, B</i>	<i>E</i>	96%	conclude <i>L</i> .
<i>A, C</i>	<i>E</i>	38%	conclude <i>L</i> .
<i>A</i>	$\neg E$	46%	conclude $\neg L$.
<i>A, B</i>	$\neg E$	4%	conclude $\neg L$.
<i>A, C</i>	$\neg E$	63%	conclude $\neg L$.

resents human knowledge. *Inferential adequacy* measures whether the computations behave similarly to human reasoning. It is straightforward to see that classical two-valued logic cannot model the suppression task adequately. At least a non-monotonic operator is needed. Stenning and van Lambalgen [1] suggest that human reasoning is modeled by, firstly, reasoning towards an appropriate representation or logical form (conceptual adequacy) and, secondly, reasoning with respect to this representation (inferential adequacy). As appropriate representation to model the suppression task, Stenning and van Lambalgen propose logic programs under completion semantics based on the three-valued logic used by Fitting [3], which itself is based on the three-valued Kleene [4] logic. Unfortunately, some technical claims made by Stenning and van Lambalgen [1] are wrong. Hölldobler and Kencana Ramli [5] have shown that the three-valued logic proposed by Fitting is inadequate for the suppression task. Somewhat surprisingly, the suppression task can be adequately modeled if the three-valued Lukasiewicz [6] logic is used.

The computational logic approach in [5, 7] models the suppression task as logic programs together with their weak completion. They show that the conclusions drawn with respect to least models correspond to the findings in [2] and conclude that the derived logic programs under Lukasiewicz logic are inferentially adequate for the suppression task.

In this paper we examine and compare Lukasiewicz semantics (L semantics) with well-founded semantics. Well-founded semantics is a widely accepted approach in the field of nonmonotonic reasoning [8]. Compared to Clark’s completion [9] or stable model semantics [10], well-founded semantics behaves more intuitive for programs with positive or negative cycles [11]. Under Clark’s completion, a program with a positive cycle might not reflect the intended interpretation of the program. For instance, consider $\mathcal{P} = \{l \leftarrow e \wedge \neg ab_1, ab_1 \leftarrow ab_2, e \leftarrow \top\}$. Under Clark’s completion we conclude *l*. However, if we extend \mathcal{P} with $ab_2 \leftarrow ab_1$, which is implied by the completion of \mathcal{P} , we cannot conclude *l* anymore. This seems to be counterintuitive. Furthermore, programs containing negative cycles might not be consistent under Clark’s completion and might not have a model under stable model semantics like e.g. $p \leftarrow \neg p$.

In the following two sections we provide the necessary definitions and explain the logic programs representing the suppression task. After that, we give the characterization of the semantics in consideration and point out their differences.

Table 3. Truth table for three-valued logics.

\neg								
	\wedge	\vee	\leftarrow_L	\leftrightarrow_L	\leftarrow_K	\leftrightarrow_S	\leftarrow_S	
\top	\top	\top	\top	\top	\top	\top	\top	\top
\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp
\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}	\mathbf{U}
	\perp	\top	\perp	\perp	\perp	\perp	\perp	\perp
	\perp	\perp	\perp	\top	\top	\top	\top	\top
	\perp	\mathbf{U}	\perp	\mathbf{U}	\mathbf{U}	\perp	\perp	\perp
	\mathbf{U}	\top	\mathbf{U}	\mathbf{U}	\mathbf{U}	\perp	\perp	\perp
	\mathbf{U}	\perp	\perp	\mathbf{U}	\top	\perp	\perp	\top
	\mathbf{U}	\mathbf{U}	\top	\top	\mathbf{U}	\top	\top	\top

Section 5 shows that by some modifications both approaches yield the same results. The last section concludes.

2 Three-valued Logics

Three-valued logics were introduced by Lukasiewicz [6] and since then different interpretations about the connectives have been proposed (see e.g. [4]). In this paper, we stick to the original definition by Lukasiewicz as it has been shown in [5, 7] that this logic adequately solves the suppression task. Table 3 gives the truth tables of three-valued conjunction, disjunction and several variants of implication and equivalence. The three-valued logics are the Lukasiewicz logic (\leftarrow_L), Kleene's strong implication [4] (\leftarrow_K), and seq_3 logic [12] (\leftarrow_S). The seq_3 logic corresponds to the logic used under well-founded semantics [13]. The truth values of the equivalences \leftrightarrow_L and \leftrightarrow_S are obtained from those of the respective implications, \leftarrow_L and \leftarrow_S , by conjoining them in both directions. \top , \perp , and \mathbf{U} denote *true*, *false*, and *unknown*, respectively. Stenning and van Lambalgen [1] have taken the logic $\{\neg, \wedge, \vee, \leftarrow_K, \leftrightarrow_S\}$ to model the suppression task. Hölldobler and Kenana Ramli [5] show that their logic is inadequate and propose to use $\{\neg, \wedge, \vee, \leftarrow_L, \leftrightarrow_L\}$ which corrects their mistakes and adequately models the suppression task.

3 Preliminaries

We define the necessary notations we will use throughout this paper and restrict ourselves to propositional logic as this is sufficient to solve the suppression task.

3.1 Logic Programs

A *logic program* \mathcal{P} is a finite set of clauses of the form

$$A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m \quad (1)$$

where A is an atom and called *head* and $A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ is called *body* of the clause. The body can be either a conjunction of atoms A_i with $0 \leq i \leq n$

and negated atoms B_j with $0 \leq j \leq m$, where i or j is ≥ 1 . \top and \perp are special cases of atoms where a clause of the form $A \leftarrow \top$ is called *positive fact* and a clause of the form $A \leftarrow \perp$ is called *negative fact*. Let $\mathcal{A}_{\mathcal{P}}$ be the set of all atoms occurring in \mathcal{P} . An atom is *not defined* in \mathcal{P} if it is not the head of any clause in \mathcal{P} ; $nd(\mathcal{P}) = \{A \mid \text{there is no clause } C \text{ in } \mathcal{P} \text{ such that } A \text{ is the head of } C\}$. Without loss of generality, we will restrict the use of \top and \perp in the body of the clauses on facts. A program \mathcal{P}^- is the program \mathcal{P} without negative facts.

3.2 Interpretations and Models

An *interpretation* I is a mapping from formulas to the set of truth values $\{\top, \perp, \text{U}\}$. With respect to a particular logic, an interpretation is represented by a pair $I = \langle I^\top, I^\perp \rangle$ of disjoint sets of atoms with the understanding that all atoms mapped to \top are in I^\top and all atoms mapped to \perp are in I^\perp and vice versa. If atoms are mapped to U , they are neither in I^\top nor in I^\perp . A *total interpretation* is an interpretation $I = \langle I^\top, I^\perp \rangle$ such that $\mathcal{A}_{\mathcal{P}} = I^\top \cup I^\perp$.

An interpretation I is a *model* for \mathcal{P} if I maps each clause occurring in \mathcal{P} to \top . A *basic model* I is a model determined according to the seq_3 logic. As apparent from table 3 whenever a formula is true under the S implication it is true under the L implication and vice versa. Thus, every basic model is also a model under Lukasiewicz logic. The *greatest model* of \mathcal{P} is the union of all models of \mathcal{P} .

A *well-founded model* is defined based on *Well-founded semantics* [14], which is only defined for logic programs without negative facts. The definition of a well-founded model requires the definition of *unfounded sets*: $U \subseteq \mathcal{A}$ is an unfounded set of \mathcal{P} with respect to I if each atom $A \in U$ satisfies the following condition: For each clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$, (at least) one of the following holds:

1. there exists $A_i \in I^\perp$, $1 \leq i \leq n$, or there exists $B_j \in I^\top$, $1 \leq j \leq m$.
2. there exists $A_i \in U$, $1 \leq i \leq n$.

Let I be an interpretation and \mathcal{P} be a program without negative facts, i.e. programs \mathcal{P}^- . The operators $T_{\mathcal{P}}$, $\mathcal{U}_{\mathcal{P}}$, and $W_{\mathcal{P}}$ which determine the well-founded semantics of \mathcal{P} , are defined as follows:

- $T_{\mathcal{P}}(I)$ is the set of all literals A such that there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m \in \mathcal{P}$, where each A_i is in I^\top and each B_j is in I^\perp .
- $\mathcal{U}_{\mathcal{P}}(I)$ is the greatest unfounded set of \mathcal{P} with respect to I .
- $W_{\mathcal{P}}(I) = T_{\mathcal{P}}(I) \cup \neg \mathcal{U}_{\mathcal{P}}(I)$.

where the greatest unfounded set $\mathcal{U}_{\mathcal{P}}(I)$ of \mathcal{P} with respect to I is the union of all unfounded sets of \mathcal{P} with respect to I . The negation of $\mathcal{U}_{\mathcal{P}}(I)$, $\neg \mathcal{U}_{\mathcal{P}}(I)$ contains all negated atoms of $\mathcal{U}_{\mathcal{P}}(I)$. The well-founded model of \mathcal{P} is the least fixed point of $W_{\mathcal{P}}(I)$.

A *L model* of a logic program \mathcal{P} is the least fixed point of the $\Phi_{\mathcal{P}}$ operator under Lukasiewicz logic. Following the definition of [1], let I be an interpretation in

Table 4. Representational Form of the Suppression Task

\mathcal{P}	Clauses			Facts	Byrne
\mathcal{P}_{AE}	$l \leftarrow e \wedge \neg ab_1$		$ab_1 \leftarrow \perp$	$e \leftarrow \top$	96% L
\mathcal{P}_{ABE}	$l \leftarrow e \wedge \neg ab_1$	$l \leftarrow t \wedge \neg ab_2$	$ab_1 \leftarrow \perp$	$ab_2 \leftarrow \perp$	$e \leftarrow \top$ 96% L
\mathcal{P}_{ACE}	$l \leftarrow e \wedge \neg ab_1$	$l \leftarrow o \wedge \neg ab_3$	$ab_1 \leftarrow \neg o$	$ab_3 \leftarrow \neg e$	$e \leftarrow \top$ 38% L
$\mathcal{P}_{A\bar{E}}$	$l \leftarrow e \wedge \neg ab_1$		$ab_1 \leftarrow \perp$	$e \leftarrow \perp$	46% $\neg L$
$\mathcal{P}_{AB\bar{E}}$	$l \leftarrow e \wedge \neg ab_1$	$l \leftarrow t \wedge \neg ab_2$	$ab_1 \leftarrow \perp$	$ab_2 \leftarrow \perp$	$e \leftarrow \perp$ 4% $\neg L$
$\mathcal{P}_{AC\bar{E}}$	$l \leftarrow e \wedge \neg ab_1$	$l \leftarrow o \wedge \neg ab_3$	$ab_1 \leftarrow \neg o$	$ab_3 \leftarrow e$	$e \leftarrow \perp$ 63% $\neg L$

$\Phi_{\mathcal{P}}(I) = \langle J^{\top}, J^{\perp} \rangle$, where

$$\begin{aligned}
 J^{\top} &= \{A \mid \text{there exists } A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m \in \mathcal{P} \text{ with} \\
 &\quad A_1, \dots, A_n \subseteq I^{\top} \text{ and } B_1, \dots, B_m \subseteq I^{\perp}\}, \\
 J^{\perp} &= \{A \mid \text{there exists } A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m \in \mathcal{P} \text{ and} \\
 &\quad \text{for all } A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m \in \mathcal{P} \\
 &\quad \text{there exists } A_i \in I^{\perp} \text{ or there exists } B_j \in I^{\top}\}.
 \end{aligned}$$

As shown in [5] for any \mathcal{P} , the least fixed point of $\Phi_{\mathcal{P}}$ can be computed by iterating $\Phi_{\mathcal{P}}$ starting with the empty interpretation.

3.3 Reasoning Towards an Appropriate Logical Form

Stenning and van Lambalgen [1] have argued that the first step in modeling human reasoning is reasoning towards an appropriate logical form. In particular, they argue that conditionals shall not be encoded by implications straight away but rather by licenses for implications. For example, the conditional A in Table 1 should be encoded by the clause $l \leftarrow e \wedge \neg ab_1$, where ab_1 is an *abnormality* predicate which expresses that something abnormal is known. In other words, l holds if e is true and nothing abnormal is known.

Table 4 shows the representational form of the first six examples of the suppression task as modeled by Stenning and van Lambalgen.¹ The last column of Table 4 shows the findings of Byrne [2].

The predicates ab_1 , ab_2 and ab_3 represent different kinds of abnormality. For instance, \mathcal{P}_{ABE} and \mathcal{P}_{ACE} contain two clauses which yield to the same conclusion l . Both programs differ in the way that for \mathcal{P}_{ABE} , the premise of the second clause is an *alternative* to the first clause and in \mathcal{P}_{ACE} the premise of the second clause is an *additional* to the first clause. The difference is represented by the additional clause in \mathcal{P}_{ACE} where ab_1 is true when *The library does not stay open* and ab_3 is true when *She does not have an essay to finish*. We assume that Stenning and van Lambalgen adequately model this part of the suppression task and adopt this reasoning step.

4 L Semantics and Well-Founded Semantics

In this section we give the characterizations of L and well-founded semantics by level mappings and compare them.

¹ The other 6 cases require abduction (see [7]) which is beyond the scope of this paper.

4.1 Level Mapping

A *level mapping* for a program \mathcal{P} is a function $L : \mathcal{A}_{\mathcal{P}} \rightarrow \mathbb{N}$. A *partial level mapping* for \mathcal{P} is a level mapping that may be undefined for some atoms. An *I-partial level mapping* for \mathcal{P} is a partial mapping for an interpretation where the domain of L is $dom(L) = I^{\top} \cup I^{\perp}$. Consider the following three programs

$$\begin{aligned}\mathcal{P}_1 &= \{p \leftarrow q, q \leftarrow \top, q \leftarrow \neg r\} \\ \mathcal{P}_2 &= \{p \leftarrow p\} \\ \mathcal{P}_3 &= \{p \leftarrow \neg q, q \leftarrow \neg p\}\end{aligned}$$

For instance, in \mathcal{P}_1 we have that p and q are in the domain $dom(L)$ but r is not, because r is neither in I^{\top} nor in I^{\perp} . If we want to identify some property for this program, we can define some level mapping, e.g. $L(\top) = 1$, $L(r) = 2$, $L(q) = 3$, and $L(p) = 4$. This is a level mapping where the head of each clause in \mathcal{P}_1 is strictly higher than the atoms in its body. It is not possible to define such a level mapping for \mathcal{P}_2 and \mathcal{P}_3 because they contain a *cycle*. We have a cycle in a program if at least one atom *depends on* itself. We say that p depends on q if and only if $p \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ such that $q = A_i$ or $q = B_j$ where $1 \leq i \leq n$ and $1 \leq j \leq m$. We say that p *depends positively* on q , if $q = A_i$ and p *depends negatively* on q if $q = B_j$. Dependency is transitive, thus if p depends on q and q depends on r , then p depends on r where one negative dependency is enough to define the whole dependency as negative.

Accordingly, if an atom depends positively on itself, such as in program \mathcal{P}_2 , then we call this a *positive cycle* and if an atom depends negatively on itself, such as in program \mathcal{P}_3 then we call this a *negative cycle*. By restricting the level mapping conditions we can specify these properties: \mathcal{P} is *acyclic* with respect to L if for every clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} we find that

$$L(A) > L(A_i) \text{ for all } i, 1 \leq i \leq n \text{ and } L(A) > L(B_j) \text{ for all } j, 1 \leq j \leq m.$$

\mathcal{P} is acyclic if it is acyclic with respect to some level mapping. \mathcal{P}_1 is acyclic, whereas \mathcal{P}_2 and \mathcal{P}_3 are not. We distinguish two cases: \mathcal{P} is *stratified* with respect to L if for every clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} we find that

$$L(A) \geq L(A_i) \text{ for all } i, 1 \leq i \leq n \text{ and } L(A) > L(B_j) \text{ for all } j, 1 \leq j \leq m.$$

\mathcal{P} is stratified if it is stratified with respect to some level mapping [15, 16]. Programs which only have positive cycles, are stratified. \mathcal{P}_2 is stratified but \mathcal{P}_3 is not. \mathcal{P} is *tight* with respect to L if for every clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} we find that

$$L(A) > L(A_i) \text{ for all } i, 1 \leq i \leq n \text{ and } L(A) \geq L(B_j) \text{ for all } j, 1 \leq j \leq m.$$

\mathcal{P} is tight if it is tight with respect to some level mapping [17, 18]. Programs which only have negative cycles, are tight programs. \mathcal{P}_3 is tight but \mathcal{P}_2 is not.

4.2 L Semantics

Kenana Ramli [19] gives the following characterization of the L semantics for a logic program.

Let \mathcal{P} be a logic program, $I = \langle I^\top, I^\perp \rangle$ be a basic model of \mathcal{P} and \mathbb{L} be an I -partial level mapping for \mathcal{P} . \mathcal{P} is said to satisfy (\mathbb{L}) wrt I and \mathbb{L} if for every $A \in \text{dom}(\mathbb{L})$ one of the following conditions is satisfied:

- (Li) $A \in I^\top$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} such that the following conditions hold:
 - (Lia) for all i with $1 \leq i \leq n$, $A_i \in I^\top$ and $\mathbb{L}(A) > \mathbb{L}(A_i)$
 - (Lib) for all j with $1 \leq j \leq m$, $B_j \in I^\perp$ and $\mathbb{L}(A) > \mathbb{L}(B_j)$.
- (Lii) $A \in I^\perp$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} and for each such clause one of the following conditions holds:
 - (Liiia) there exists i with $1 \leq i \leq n$, $A_i \in I^\perp$ and $\mathbb{L}(A) > \mathbb{L}(A_i)$,
 - (Liiib) there exists j with $1 \leq j \leq m$, $B_j \in I^\top$ and $\mathbb{L}(A) > \mathbb{L}(B_j)$.

If $A \in \text{dom}(\mathbb{L})$ satisfies (Li), then we say that A satisfies (Li) wrt I and \mathbb{L} , and similarly if $A \in \text{dom}(\mathbb{L})$ satisfies (Lii).

Hölldobler and Kenana Ramli [5] show that the model intersection property holds for programs under \mathbb{L} semantics. This property guarantees the existence of a \mathbb{L} model for any logic program.

Theorem 1. *Let \mathcal{P} be a logic program with the \mathbb{L} model M . Then M is the greatest model among all models I for which there exists an I -partial level mapping \mathbb{L} for \mathcal{P} such that \mathcal{P} satisfies (\mathbb{L}) wrt I and \mathbb{L} .*

Intuitively, the level mapping that satisfies (\mathbb{L}) wrt to all $A \in \text{dom}(\mathbb{L})$ is acyclic wrt $\langle I^\top, \emptyset \rangle$ and wrt $\langle \emptyset, I^\perp \rangle$.

4.3 Well-Founded Semantics

Hitzler and Wendt [20] give a characterization for the well-founded semantics.

Let \mathcal{P} be a logic program, $I = \langle I^\top, I^\perp \rangle$ be a basic model of \mathcal{P} and \mathbb{L} be an I -partial level mapping for \mathcal{P} . \mathcal{P} is said to satisfy (WF) wrt I and \mathbb{L} if for every $A \in \text{dom}(\mathbb{L})$ one of the following conditions is satisfied:

- (WFi) $A \in I^\top$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} such that the following conditions hold:
 - (WFia) for all i with $1 \leq i \leq n$, $A_i \in I^\top$ and $\mathbb{L}(A) > \mathbb{L}(A_i)$
 - (WFib) for all j with $1 \leq j \leq m$, $B_j \in I^\perp$ and $\mathbb{L}(A) > \mathbb{L}(B_j)$.
- (WFii) $A \in I^\perp$ and for each clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} , one of the following conditions holds:
 - (WFiia) there exists i with $1 \leq i \leq n$, $A_i \in I^\perp$ and $\mathbb{L}(A) \geq \mathbb{L}(A_i)$,
 - (WFiib) there exists j with $1 \leq j \leq m$, $B_j \in I^\top$ and $\mathbb{L}(A) > \mathbb{L}(B_j)$.

If $A \in \text{dom}(\mathbb{L})$ satisfies (WFi), then we say that A satisfies (WFi) with respect to I and \mathbb{L} , and similarly if $A \in \text{dom}(\mathbb{L})$ satisfies (WFii). Every logic program has a well-founded model [8].

Theorem 2. *Let \mathcal{P} be a logic program with the well-founded model M . Then M is the greatest model among all models I for which there exists an I -partial level mapping \mathbb{L} for \mathcal{P} such that \mathcal{P} satisfies (WF) wrt I and \mathbb{L} .*

Intuitively, the level mapping that satisfies (WF) wrt to all $A \in \text{dom}(\mathbb{L})$ is acyclic wrt $\langle I^\top, \emptyset \rangle$ and stratified wrt $\langle \emptyset, I^\perp \rangle$.

Table 5. Programs with corresponding \mathbb{L} and well-founded models

\mathcal{P}	Property	Possible \mathbb{L} -mapping order	\mathbb{L} model	well-founded model
\mathcal{P}_1	acyclic	$\mathbb{L}(\top) < \mathbb{L}(r) < \mathbb{L}(q) < \mathbb{L}(p)$	$\langle \{p, q\}, \emptyset \rangle$	$\langle \{p, q\}, \{r\} \rangle$
\mathcal{P}_2	stratified		$\langle \emptyset, \emptyset \rangle$	$\langle \emptyset, \{p, q\} \rangle$
\mathcal{P}_3	tight	$\mathbb{L}(p) = \mathbb{L}(q)$	$\langle \emptyset, \emptyset \rangle$	$\langle \emptyset, \emptyset \rangle$

4.4 Properties of \mathbb{L} Semantics and Well-Founded Semantics

The characterizations of \mathbb{L} semantics and well-founded semantics are different in two aspects. First, consider conditions (Lii) and (WFii):

(Lii) $A \in I^\perp$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} and for each such clause one of the following conditions holds: [...]

(WFii) $A \in I^\perp$ and for each clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P} , one of the following conditions holds: [...]

By condition (WFii) all undefined predicates are in I^\perp whereas in \mathbb{L} semantics, they stay undefined. Furthermore, they differ in conditions (Liia) and (WFiia):

(Liia) there exists i with $1 \leq i \leq n$, $A_i \in I^\perp$ and $\mathbb{L}(A) > \mathbb{L}(A_i)$,

(WFiia) there exists i with $1 \leq i \leq n$, $A_i \in I^\perp$ and $\mathbb{L}(A) \geq \mathbb{L}(A_i)$,

In a well-founded model, all predicates which are part of a positive cycle are in I^\perp , whereas in \mathbb{L} semantics these predicates stay undefined.

Table 5 shows the \mathbb{L} and the well-founded models for \mathcal{P}_1 , \mathcal{P}_2 and \mathcal{P}_3 . In the well-founded model of \mathcal{P}_1 , r is in I^\perp because it is an undefined predicate whereas in the \mathbb{L} model r stays undefined. Consider programs \mathcal{P}_2 and \mathcal{P}_3 under \mathbb{L} semantics.

In both cases the I -partial level mapping is undefined for p and q because they are both part of a cycle. Consider an extension of \mathcal{P}_3 , $\mathcal{P}_3^{ext} = \mathcal{P}_3 \cup \{p \leftarrow \top\}$. The \mathbb{L} and the well-founded models are $\langle \{p\}, \{q\} \rangle$ because \mathcal{P}_3^{ext} satisfies (L) and (WF) wrt (the greatest model) $\langle \{p\}, \{q\} \rangle$ and the level mapping $\mathbb{L}(\top) = 0, \mathbb{L}(p) = 1, \mathbb{L}(q)$.

5 Correspondence \mathbb{L} Semantics and Well-Founded Semantics

In the following we show how \mathbb{L} semantics and well-founded semantics yield the same models and recall the examples from the suppression task.

5.1 Claim

There is a relation between \mathbb{L} semantics and well-founded semantics for tight programs. In the following, we look at programs where negative facts are only formulated when p is not the head of any other clause in \mathcal{P} . Under \mathbb{L} semantics this does not restrict the expressions of our programs as by condition (Lii), we can only conclude that p is in I^\perp if *for all* clauses where p is the head of, the body is in I^\perp . Thus $p \leftarrow \perp$ would not add any more information when

Table 6. Suppression Task with the corresponding L and Well-founded models

\mathcal{P}	L Model	Well-founded Model
\mathcal{P}_{AE}	$\langle \{e, l\}, \{ab_1\} \rangle$	$\langle \{e, l\}, \{ab_1\} \rangle$
\mathcal{P}_{ABE}	$\langle \{e, l\}, \{ab_1, ab_2\} \rangle$	$\langle \{e, l\}, \{ab_1, ab_2\} \rangle$
\mathcal{P}_{ACE}	$\langle \{e\}, \{ab_3\} \rangle$	$\langle \{e, ab_1\}, \{o, ab_3, l\} \rangle$
\mathcal{P}_{ACE}^{mod}		$\langle \{e\}, \{ab_3\} \rangle$
$\mathcal{P}_{A\bar{E}}$	$\langle \emptyset, \{ab_1, e, l\} \rangle$	$\langle \emptyset, \{ab_1, e, l\} \rangle$
$\mathcal{P}_{AB\bar{E}}$	$\langle \emptyset, \{ab_1, ab_2, e\} \rangle$	$\langle \emptyset, \{ab_1, ab_2, e, t, l\} \rangle$
$\mathcal{P}_{AB\bar{E}}^{mod}$		$\langle \emptyset, \{ab_1, ab_2, e\} \rangle$
$\mathcal{P}_{AC\bar{E}}$	$\langle \{ab_3\}, \{e, l\} \rangle$	$\langle \{ab_3\}, \{e, l\} \rangle$

there is another clause with p in the head for which the body is not in I^\perp . For well-founded models, we can omit negative facts because they are implicit when predicates are not defined.

Theorem 3. *Let \mathcal{P} be a tight logic program. We claim that the L model of \mathcal{P} is the well-founded model of \mathcal{P}^{mod} , where the modified program $\mathcal{P}^{mod} = \mathcal{P}^- \cup \{A \leftarrow \neg A \mid A \in nd(\mathcal{P}^-)\}$.*

The proof is in the appendix.

5.2 The Suppression Task

Table 6 shows how the L semantics and well-founded semantics are applied for the six cases of the suppression task. Programs \mathcal{P}_{ACE} and $\mathcal{P}_{AB\bar{E}}$ reflect the actual suppression of additional and alternative information and show that the corresponding L model and well-founded model are different. \mathcal{P}_{ACE}^{mod} and $\mathcal{P}_{AB\bar{E}}^{mod}$ in the fourth and seventh row give the well-founded model of the modified programs as proposed in our claim and shows that the well-founded models are indeed the same as the corresponding L models of the program.

6 Conclusion

We investigated L semantics and well-founded semantics for the suppression task. Both approaches have different underlying three-valued interpretations. L semantics adequately models the suppression task, whereas well-founded semantics does not. However, by modifying the programs under consideration for the well-founded semantics, both compute the same adequate models. We show that both approaches yield the same results for the class of tight programs.

7 Acknowledgments

The paper formalized an idea that was brought to our attention by Alexandre Miguel Pinto. Many thanks to Christoph Wernhard and Bertram Fronhöfer.

References

1. Stenning, K., van Lambalgen, M.: Human reasoning and cognitive science. Bradford Books. MIT Press (2008)
2. Byrne, R.M.J.: Suppressing valid inferences with conditionals. *Cognition* **31** (1989) 61–83
3. Fitting, M.: A Kripke-Kleene semantics for logic programs. *J. Log. Program.* **2** (1985) 295–312
4. Kleene, S.C.: Introduction to metamathematics. *Bibl. Matematica*. North-Holland, Amsterdam (1952)
5. Hölldobler, S., Kencana Ramli, C.D.: Logic programs under three-valued lukasiewicz semantics. In: Proceedings of the 25th International Conference on Logic Programming. ICLP '09, Berlin, Heidelberg, Springer-Verlag (2009) 464–478
6. Łukasiewicz: O logice trójwartościowej. *Ruch Filozoficzny* **5** (1920) 169–171 English translation: On Three-Valued Logic. In: *Jan Łukasiewicz Selected Works*. (L. Borkowski, ed.), North Holland, 87–88, 1990.
7. Dietz, E.A., Hölldobler, S., Ragni, M.: A Computational Approach to the Suppression Task. to appear in Proceedings of the 34th Cognitive Science Conference (2012)
8. Van Gelder, A., Ross, K.A., Schlipf, J.S.: The well-founded semantics for general logic programs. *J. ACM* **38** (1991) 619–649
9. Clark, K.L.: Negation as failure. In Minker, J., ed.: *Logic and Data Bases*. Volume 1. Plenum Press, New York, London (1978) 293–322
10. Gelfond, M., Lifschitz, V.: The stable model semantics for logic programming. In Kowalski, R., Bowen, Kenneth, eds.: *Proceedings of International Logic Programming Conference and Symposium*, MIT Press (1988) 1070–1080
11. Przymusiński, T.: Well founded and stationary models of logic programs. *Annals of Mathematics and Artificial Intelligence* **12** (1994) 141–187
12. Gottwald, S.: *A Treatise on Many-Valued Logics*. Volume 9 of *Studies in Logic and Computation*. Research Studies Press, Baldock, UK (2001)
13. Przymusiński, T.C.: Every logic program has a natural stratification and an iterated least fixed point model. In: *Proceedings of the eighth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*. PODS '89, New York, NY, USA, ACM (1989) 11–21
14. Van Gelder, A., Ross, K., Schlipf, J.S.: Unfounded sets and well-founded semantics for general logic programs. In: *Proceedings of the seventh ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*. PODS '88, New York, NY, USA, ACM (1988) 221–230
15. Apt, K.R., Blair, H.A., Walker, A.: *Foundations of deductive databases and logic programming*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1988) 89–148
16. Przymusiński, T.C.: *Foundations of deductive databases and logic programming*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1988) 193–216
17. Fages, F.: Consistency of clark's completion and existence of stable models. *Meth. of Logic in CS* **1** (1994) 51–60
18. Erdem, E., Lifschitz, V.: Tight logic programs. *CoRR* (2003)
19. Kencana Ramli, C.D.: Logic programs and three-valued consequence operators. Master's thesis, Institute for Artificial Intelligence, Department of Computer Science, Technische Universität Dresden (2009)
20. Hitzler, P., Wendt, M.: A uniform approach to logic programming semantics. *Theory and Practice of Logic Programming* **5** (2005) 123–159

8 Appendix

Proof of Theorem 3

$$A \in \mathbf{L}(\mathcal{P}) \text{ if and only if } A \in \mathbf{WF}(\mathcal{P}^{mod})$$

where $\mathbf{L}(\mathcal{P})$ and $\mathbf{WF}(\mathcal{P}^{mod})$ are the \mathbf{L} model of \mathcal{P} and the well-founded model of \mathcal{P}^{mod} , respectively. Both are pairs of interpretations $\langle I^\top, I^\perp \rangle$:

- (1) $A \in \mathbf{L}(\mathcal{P})^\top$ if and only if $A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and
- (2) $A \in \mathbf{L}(\mathcal{P})^\perp$ if and only if $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$

where $\mathbf{L}(\mathcal{P})^\top$ ($\mathbf{L}(\mathcal{P})^\perp$) and $\mathbf{WF}(\mathcal{P}^{mod})^\top$ ($\mathbf{WF}(\mathcal{P}^{mod})^\perp$) are the sets of atoms which are in I^\top (I^\perp) in the \mathbf{L} model of \mathcal{P} and in the well-founded model of \mathcal{P}^{mod} , respectively.

Proof (1)

→ Assume $A \in \mathbf{L}(\mathcal{P})^\top$. This is only the case if condition (Li) holds. Condition (Li) and (WFi) are the same and $\mathcal{P} \subseteq \mathcal{P}^{mod}$, thus $A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$.

← If $A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$, then never because of some clause in the extension of \mathcal{P} : $\{A \leftarrow \neg A\}$, but because of some clause in \mathcal{P} itself. Since conditions (Li) and (WFi) are the same and $A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ then $A \in \mathbf{L}(\mathcal{P})^\top$.

Proof (2)

→ Assume $A \in \mathbf{L}(\mathcal{P})^\perp$. This is only the case if condition (Lii) holds. Condition (Lii) is stricter than condition (WFii) and $\mathcal{P} \subseteq \mathcal{P}^{mod}$, thus $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$ as well.

← (By Contraposition) Assume $A \notin \mathbf{L}(\mathcal{P})^\perp$ (to prove: $A \notin \mathbf{WF}(\mathcal{P}^{mod})^\perp$). Then for condition (Lii) there are two possible cases:

i there is not such a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, B_m$ in \mathcal{P} . That means A is not the head of any clause: $A \in \mathbf{nd}(\mathcal{P})$. Thus $A \leftarrow \neg A \in \mathcal{P}^{mod}$, which is also the only clause in \mathcal{P}^{mod} with A in the head.

Assume $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$. This is only the case if condition (WFii) holds (as the only atom in the body of the clause $A \leftarrow \neg A$ is a negated atom) which is only the case if for each clause with A in the head there is some negated atom B_j with $1 \leq B_j \leq B_m$ occurring in the body we have that $B_j \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and $\mathbf{L}(A) > \mathbf{L}(B_j)$. We have exactly one clause with A in the head, $A \leftarrow \neg A$, but then $\mathbf{L}(A) > \mathbf{L}(A)$. Contradiction.

ii for all clauses in \mathcal{P} with A in the head neither condition Liia nor Liib hold for some clause with A in the head.

Assume $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$. Then for each clause in \mathcal{P} with A in the head, either (WFiia) or (WFiib) holds.

(a) Assume (WFiia) holds: then there exists i with $1 \leq i \leq n$, $A_i \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and $\mathbf{L}(A) \geq \mathbf{L}(A_i)$. Since we restrict ourselves to tight programs, it can only be that $1 \leq i \leq n$, $A_i \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and $\mathbf{L}(A) > \mathbf{L}(A_i)$. Since we assume this for each clause in \mathcal{P}^{mod} and $\mathcal{P} \subseteq \mathcal{P}^{mod}$ (Liia) holds for \mathcal{P} as well and we have that $A \in \mathbf{L}(\mathcal{P})^\perp$. Contradiction.

- (b) Assume (WFiiib) holds: [...]. Then (Liib) holds for \mathcal{P} as well and we have that $A \in \mathbf{L}(\mathcal{P})^\perp$. Contradiction.

If we replace the clause $A \leftarrow \neg A$ by the following two clauses and an additional auxiliary atom e.g. $A \leftarrow \neg \text{not_}A, \text{not_}A \leftarrow \neg A$, then $\mathcal{P}^{mod} = \mathcal{P} \cup \{A \leftarrow \neg \text{not_}A, \text{not_}A \leftarrow \neg A \mid A \in \text{nd}(\mathcal{P})\}$. The proof is the same except for part [i]:

- i either there is not such a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, B_m$ in \mathcal{P} . So we have that A is not the head of any clause and thus $A \in \text{nd}(\mathcal{P})$. We introduce an additional auxiliary variable $\text{not_}A$ and two additional clauses $A \leftarrow \neg \text{not_}A$ and $\text{not_}A \leftarrow \neg A$ for \mathcal{P}^{mod} . These are the only two clauses in \mathcal{P}^{mod} with A and $\text{not_}A$ in the head, respectively.

Assume $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$. This is only the case if condition (WFiiib) holds (as the only atom in the body of the clause $A \leftarrow \neg \text{not_}A$ is a negated atom). That is $\text{not_}A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and $\mathbf{L}(A) > \mathbf{L}(\text{not_}A)$. $\text{not_}A \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ if (WFi) holds. That is, there exists a clause $\text{not_}A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in \mathcal{P}^{mod} such that for all i with $1 \leq i \leq n$, $A_i \in \mathbf{WF}(\mathcal{P}^{mod})^\top$ and $\mathbf{L}(\text{not_}A) > \mathbf{L}(A_i)$ and for all j with $1 \leq j \leq m$, $B_j \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$ and $\mathbf{L}(\text{not_}A) > \mathbf{L}(B_j)$. As we know, $\text{not_}A \leftarrow \neg A$ is the only clause with $\text{not_}A$ in the head in \mathcal{P}^{mod} , thus $A \in \mathbf{WF}(\mathcal{P}^{mod})^\perp$ and $\mathbf{L}(\text{not_}A) > \mathbf{L}(A)$ have to hold. Contradiction.

Reasoning about Rough Sets Using Three Logical Values

Beata Konikowska¹ and Arnon Avron²

¹ Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5,
01-248 Warsaw, Poland

`Beata.Konikowska@ipipan.waw.pl`

² School of Computer Science, Tel Aviv University, Tel Aviv, Israel
`aa@cs.tau.ac.il`

Abstract. The paper presents a logic for reasoning about covering-based rough sets using three logical values: the value **t** corresponding to the positive region of a set, the value **f** — to the negative region, and the undefined value **u** — to the boundary region of that set. Atomic formulas of the logic represent membership of objects of the universe in rough sets, and complex formulas are built out of the atomic ones using three-valued Kleene connectives. In the paper we provide a strongly sound and complete Gentzen-style sequent calculus for the logic.

1 Introduction

Rough sets, developed by Pawlak in the early 1980s [18, 19], represent a simple and yet very powerful notion designed to model vague or imprecise information. Unlike Zadeh’s fuzzy sets, they are not based on any numerical measure of the degree of membership of an object in an imprecisely defined set. Instead, they employ a much more universal and versatile idea of an indiscernibility relation, which groups together objects having the same properties from the viewpoint of a certain application into disjoint equivalence classes.

This concept has proved to be extremely useful in practice. Since their introduction in the early 1980s, rough sets have found numerous applications in areas like control of manufacturing processes [14], development of decisions tables [20], data mining [14], data analysis [21], knowledge discovery [15], and so on. They have also been the subject of an impressive body of research. Though it focused mainly on algebraic properties of rough sets, a number of logicians have also explored this area, presenting and studying various brands of logics connected with rough sets [6, 5, 7, 8, 12, 16, 17, 10, 9, 23, 24].

Over the years, the original notion of rough sets has been generalized by replacing the indiscernibility relation (representing a partition of the universe of objects) underlying Pawlak’s approach with other, less restrictive constructs. The broadest generalization are covering-based rough sets [27, 22], defined based on an arbitrary covering of the universe of objects. By now, this notion has also been examined in many papers (see e.g. [25, 26, 29]), with the main focus again on the algebraic properties of such generalized rough sets.

In this paper we explore the logical aspects of covering-based rough sets, employing for that purpose a three-valued logic. The motivation for using three logical values stems from the fact that, like in case of Pawlak's rough sets, a covering-based approximation space defines three regions of any set X of objects:

- positive region of X , containing all objects of the universe which *certainly belong to X* in the light of the information provided by the covering;
- negative region of X , containing all objects which *certainly do not belong to X* ;
- boundary of X , containing all objects which *cannot be said for sure to either belong or not to belong to X* .

Hence the most natural idea for reasoning about membership of objects in covering-based rough sets is to use a three-valued logic, with the following values:

- **t** — meaning “certainly belongs”, and assigned to objects in the positive region of a given set;
- **f** — meaning “certainly does not belong” and assigned to objects in the negative region of the set; and
- **u** — meaning “not known to either belong or not”, and assigned to the boundary of the set.

Such an idea was first exploited in [3] for the original rough sets based on an equivalence relation on the universe of objects. However, the logic developed there was just a simple propositional logic generated by the three-valued non-deterministic matrix (see [4, 2]), shortly: Nmatrix, which did reflected only some properties of set-theoretic operations on rough sets.

Then next attempt was made in [13] for covering-based rough sets. There the semantics of the logics was based on the natural frameworks for such sets, i.e. covering-based approximations spaces. Atomic formulas of the logic represented either membership of objects of the universe in rough sets or the *subordination relation*³ between objects generated by the covering underlying the approximation space, and complex formulas were formed out of the atomic ones using three-valued Kleene connectives. A Gentzen sequent calculus for the logic was presented, and its strong soundness was proved. However, strong completeness was only proved for the subset of the language containing formulas without set-theoretic operations on rough sets.

In this paper we improve on the results of [13] by providing a strongly sound and complete Gentzen calculus for the logic of rough sets defined as in [13], but without the subordination relation.

The paper is organized as follows. Section 2 presents the fundamentals of covering-based rough sets. Section 3 defines the syntax and semantics of the logic \mathcal{L}_{RS} examined in the paper, including satisfaction and consequence relations

³ Given a covering \mathcal{C} of a universe U , the subordination relation generated by \mathcal{C} is the binary relation $\prec_{\mathcal{C}}$ on U such that $x \prec_{\mathcal{C}} y \Leftrightarrow (\forall C \in \mathcal{C})(y \in C \rightarrow x \in C)$. The relation $\prec_{\mathcal{C}}$ is reflexive and symmetric, and it is an equivalence relation iff \mathcal{C} is a partition, i.e. in case of the original Pawlak's rough sets.

for formulas and sequents, Section 4 presents a strongly sound and complete Gentzen sequent calculus for \mathcal{L}_{RS} , and finally Section 5 presents the conclusions and outlines future work.

2 Covering-based rough sets

In what follows, for any set X , by $\mathcal{P}(X)$ we denote the powerset of X , i.e. the set of all subsets of X , and by $\mathcal{P}^+(X)$ — the set of all nonempty subsets of X .

Definition 1. By a *covering-based approximation space*, or shortly *approximation space*, we mean any ordered pair $\mathcal{A} = (U, \mathcal{C})$, where U is a non-empty universe of *objects*, and $\mathcal{C} \subseteq \mathcal{P}^+(U)$ is a covering of U , i.e. $\bigcup\{C \mid C \in \mathcal{C}\} = U$.

Definition 2. For any approximation space $\mathcal{A} = (U, \mathcal{C})$:

- The *lower approximation* of a set $X \subseteq U$ with respect to the covering \mathcal{C} is

$$L_{\mathcal{C}}(X) = \{x \in U \mid \forall C \in \mathcal{C}(x \in C \Rightarrow C \subseteq X)\}$$

- The *upper approximation* of a set $X \subseteq U$ with respect to the covering \mathcal{C} is

$$H_{\mathcal{C}}(X) = \bigcup\{C \in \mathcal{C} \mid C \cap X \neq \emptyset\}$$

In view of the above definition, one can say that, given the approximate knowledge about objects available in the approximation space \mathcal{A} :

- $L_{\mathcal{C}}(X)$ is the set of all the objects in U which *certainly* belong to X ;
- $H_{\mathcal{C}}(X)$ is the set of all the objects in U which *might* belong to X ;

The above operations have the same basic properties as in case of Pawlak's rough sets based on a partition of the universe, i.e. for any $X, Y \subseteq U$, we have:

$$L_{\mathcal{C}}(X) \subseteq X \subseteq H_{\mathcal{C}}(X)$$

$$\begin{aligned} H_{\mathcal{C}}(X \cup Y) &= H_{\mathcal{C}}X \cup H_{\mathcal{C}}Y & L_{\mathcal{C}}(X \cup Y) &\supseteq L_{\mathcal{C}}X \cup L_{\mathcal{C}}Y \\ L_{\mathcal{C}}(X \cap Y) &= L_{\mathcal{C}}X \cap L_{\mathcal{C}}Y & H_{\mathcal{C}}(X \cap Y) &\subseteq H_{\mathcal{C}}X \cap H_{\mathcal{C}}Y \\ L_{\mathcal{C}}(-X) &= -H_{\mathcal{C}}X & H_{\mathcal{C}}(-X) &= -L_{\mathcal{C}}X \end{aligned} \quad (1)$$

where none of the inequalities in (1) can be replaced by the equality.

Following the example of Pawlak's rough sets, with any subset of a universe U of an approximation space we can associate three regions of that universe: positive, negative and boundary, representing three basic statuses of membership of an object of the universe U in X :

Definition 3. Let $\mathcal{A} = (U, \mathcal{C})$ be an approximation space, and let $X \subseteq U$. Then:

- The *positive region* of X in the space \mathcal{A} with respect to the covering \mathcal{C} is

$$POS_{\mathcal{C}}(X) = L_{\mathcal{C}}(X)$$

- The *negative region* of X in the space \mathcal{A} with respect to the covering \mathcal{C} is

$$NEG_{\mathcal{C}}(X) = L_{\mathcal{C}}(U - X)$$

- The *boundary region* of X in the space \mathcal{A} with respect to the covering \mathcal{C} is

$$BND_{\mathcal{C}}(X) = U - (POS_{\mathcal{C}}(X) \cup NEG_{\mathcal{C}}(X))$$

Corollary 1. For any approximation space $\mathcal{A} = (U, \mathcal{C})$ and any $X \subseteq U$:

$$\begin{aligned} POS_{\mathcal{C}}(X) &= \{x \in U \mid \forall C \in \mathcal{C}(x \in C \Rightarrow C \subseteq X)\} \\ NEG_{\mathcal{C}}(X) &= \{x \in U \mid \forall C \in \mathcal{C}(x \in C \Rightarrow C \subseteq U - X)\} \\ BND_{\mathcal{C}}(X) &= \{x \in U \mid \exists C \in \mathcal{C}(x \in C \wedge C \cap X \neq \emptyset \wedge C \cap (U - X) \neq \emptyset)\} \end{aligned} \quad (2)$$

The regions defined as above are obviously disjoint. Moreover, we can say that, according to the approximate knowledge of the properties of objects in U provided by the covering \mathcal{C} :

- The elements of $POS_{\mathcal{C}}(X)$ *certainly belong* to X ;
- The elements of $NEG_{\mathcal{C}}(X)$ *certainly do not belong* to X ;
- We *cannot tell* if the elements of $BND_{\mathcal{C}}(X)$ *belong* to X or not.

As a result, the most natural solution choice of a logic for reasoning about covering-based rough set is — exactly like in case of Pawlak’s rough sets — to base it on three logical values: **t** — true, **f** – false, **u** — unknown, corresponding to the positive, negative and the boundary region of a set, respectively.

3 Syntax and semantics of the language L_{RS}

Now we shall define the language L_{RS} of the three-valued logic for reasoning about covering-based rough sets described in the introduction. Formulas of L_{RS} will contain expressions representing sets of objects (built from set variables and set constants using set-theoretic operators), variables representing objects, the symbol $\hat{\in}$ of a three-valued binary predicate representing membership of an object in a rough set, and the logical connectives \neg, \wedge, \vee which will be interpreted as 3-valued Kleene connectives.

3.1 Syntax of L_{RS}

Definition 4. Assume that:

- OV is a non-empty denumerable set of *object variables*;
- SV is a non-empty denumerable set of *set variables*,
- **0** and **1** are *set constants*

The syntax of the language L_{RS} is defined as follows:

1. The set SE of *set expressions* of L_{RS} is the least set containing $SV \cup \{\mathbf{0}, \mathbf{1}\}$, and closed under the set-theoretic operators $-, \cup, \cap$;
2. The set of *atomic formulas* of L_{RS} is $\mathcal{A}_{RS} = \{x \hat{\in} A \mid x \in OV, A \in SE\}$;
3. The set \mathcal{F}_{RS} of *formulas* of L_{RS} is the least set \mathcal{F} containing \mathcal{A}_{RS} and closed under the connectives \neg, \vee, \wedge .

3.2 Semantic frameworks for L_{RS} and interpretation of formulas

The semantics of L_{RS} is based on interpreting the formulas of L_{RS} in semantic frameworks for that language, built on covering-based approximation spaces and including valuations of set variables and object variables.

Definition 5. A *semantic framework*, or shortly *framework*, for L_{RS} is an ordered triple $\mathcal{R} = (\mathcal{A}, v, w)$, where:

- $\mathcal{A} = (U, \mathcal{C})$ is a covering-based approximation space;
- $v : OV \rightarrow U$ is a valuation of object variables;
- $w : SV \rightarrow \mathcal{P}(U)$ is a valuation of set variables and constants such that $w(\mathbf{0}) = \emptyset, w(\mathbf{1}) = U$.

For any valuation $w : SV \rightarrow \mathcal{P}(U)$, by w^* we shall denote the extension of w to SE obtained by interpreting $-, \cup, \cap$ as the set-theoretic operations of complement, union and intersection. In other words:

$$w^*(X) = w(X) \text{ for any } X \in SV, w^*(-A) = U - w(A) \text{ for any } A \in SE, \text{ and}$$

$$w^*(A \cup B) = w^*(A) \cup w^*(B), w^*(A \cap B) = w^*(A) \cap w^*(B) \text{ for any } A, B \in SE$$

Definition 6. An *interpretation* of L_{RS} in a framework $\mathcal{R} = (\mathcal{A}, v, w)$, where $\mathcal{A} = (U, \mathcal{C})$, is a mapping $I_{\mathcal{R}} : \mathcal{F}_{RS} \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ defined as follows:

1. For any $x, y \in OV$ and any $A \in SE$,
$$I_{\mathcal{R}}(x \hat{\in} A) = \begin{cases} \mathbf{t} & \text{if } v(x) \in Pos_{\mathcal{C}}(w^*(A)) \\ \mathbf{f} & \text{if } v(x) \in Neg_{\mathcal{C}}(w^*(A)) \\ \mathbf{u} & \text{if } v(x) \in Bnd_{\mathcal{C}}(w^*(A)) \end{cases} .$$
2. For any $\varphi, \psi \in \mathcal{F}$,

$$- I_{\mathcal{R}}(\neg\varphi) = \begin{cases} \mathbf{t} & \text{if } I_{\mathcal{R}}(\varphi) = \mathbf{f} \\ \mathbf{f} & \text{if } I_{\mathcal{R}}(\varphi) = \mathbf{t} \\ \mathbf{u} & \text{if } I_{\mathcal{R}}(\varphi) = \mathbf{u} \end{cases}$$

$$- I_{\mathcal{R}}(\varphi \vee \psi) = \begin{cases} \mathbf{t} & \text{if either } I_{\mathcal{R}}(\varphi) = \mathbf{t} \text{ or } I_{\mathcal{R}}(\psi) = \mathbf{t} \\ \mathbf{f} & \text{if } I_{\mathcal{R}}(\varphi) = \mathbf{f} \text{ and } I_{\mathcal{R}}(\psi) = \mathbf{f} \\ \mathbf{u} & \text{otherwise} \end{cases}$$

$$- I_{\mathcal{R}}(\varphi \wedge \psi) = \begin{cases} \mathbf{t} & \text{if } I_{\mathcal{R}}(\varphi) = \mathbf{t} \text{ and } I_{\mathcal{R}}(\psi) = \mathbf{t} \\ \mathbf{f} & \text{if either } I_{\mathcal{R}}(\varphi) = \mathbf{f} \text{ or } I_{\mathcal{R}}(\psi) = \mathbf{f} \\ \mathbf{u} & \text{otherwise} \end{cases}$$

It can be easily seen that the interpretation $I_{\mathcal{R}}$ is a well-defined mapping of the set of formulas into $\{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$. Indeed, as the regions of a rough set are disjoint, Point 1 provides a well-defined interpretation of atomic formulas, and the clauses for \neg, \vee, \wedge in Point 2 extend it uniquely to complex formulas. In the sequel we will drop the subscript \mathcal{R} at $I_{\mathcal{R}}$ if \mathcal{R} is arbitrary or understood.

3.3 Satisfaction and consequence relations for formulas and sequents

To complete the definition of the semantics of L_{RS} , we need to define the notions of satisfaction and the consequence relation. Since the proof system we are going to develop for L_{RS} will be a sequent calculus, we will define both the notions for formulas as well as for sequents.

Definition 7.

- By a sequent we mean a structure of the form $\Gamma \Rightarrow \Delta$, where Γ and Δ are finite sets of formulas. The set of all sequents over the language L_{RS} is denoted by Seq_{RS} .
- A sequent $\Sigma \in Seq_{RS}$ is called atomic if each formula in Σ is atomic.

Depending on the specific application of rough sets, we can choose either the strong version of the three-valued semantics of L_{RS} — with \mathbf{t} as the only designated value, or its weak version — with two designated values: \mathbf{t}, \mathbf{u} , which leads to a paraconsistent logic. In this paper, we choose the strong semantics like in [3], leaving the weak version for future work. Consequently, we adopt the following definitions of satisfaction and consequence:

Definition 8. 1. A formula $\varphi \in \mathcal{F}_{RS}$ is satisfied by an interpretation I of L_{RS} , in symbols $I \models \varphi$, if $I(\varphi) = \mathbf{t}$.

2. A formula $\varphi \in \mathcal{F}_{RS}$ is valid, in symbols $\models_{RS} \varphi$, if $I \models \varphi$ for any interpretation I of L_{RS} .

3. A set of formulas $T \subseteq \mathcal{F}_{RS}$ is satisfied by an interpretation I , in symbols $I \models T$, if $I \models \varphi$ for all $\varphi \in T$.

4. A sequent $\Sigma = (\Gamma \Rightarrow \Delta)$ is satisfied by an interpretation I , in symbols $I \models \Sigma$, iff either $I \models \varphi$ for some $\varphi \in \Delta$, or $I \not\models \psi$ for some $\psi \in \Gamma$.

5. A sequent $\Sigma = (\Gamma \Rightarrow \Delta)$ is valid, in symbols $\models_{RS} \Sigma$, if $I \models \Sigma$ for any interpretation I of L_{RS} .

6. The formula consequence relation in L_{RS} is the relation \vdash_{RS} on $\mathcal{P}(\mathcal{F}_{RS}) \times \mathcal{F}_{RS}$ such that, for every $T \subseteq \mathcal{F}_{RS}$ and every $\varphi \in \mathcal{F}_{RS}$, $T \vdash_{RS} \varphi$ if each interpretation I of L_{RS} which satisfies T satisfies φ too.

7. The sequent consequence relation in L_{RS} is the relation \vdash_{RS} on $\mathcal{P}(Seq_{RS}) \times Seq_{RS}$ such that, for every $Q \subseteq Seq_{RS}$, and every $\Sigma \in Seq_{RS}$, $Q \vdash_{RS} \Sigma$ iff, for any interpretation I of L_{RS} , $I \models_{RS} Q$ implies $I \models_{RS} \Sigma$.

Note that the use of the same symbol for the formula and sequent consequence relations will not lead to misunderstanding, for the meaning of the symbol will always be clear from the context.

4 Proof system for the logic \mathcal{L}_{RS}

Now we shall present a proof system for the logic \mathcal{L}_{RS} with the language L_{RS} , corresponding to the consequence relation \vdash_{RS} defined in the preceding section. The deduction formalism we use for \mathcal{L}_{RS} is a Gentzen-style sequent calculus.

Sequent calculus CRS**Axioms:** (A1) $\varphi \Rightarrow \varphi$ (A2) $x \hat{\in} \mathbf{0} \Rightarrow$ (A3) $\Rightarrow x \hat{\in} \mathbf{1}$ **Structural rules: weakening, cut****Boolean tautology rules:** for any $A, B \in SE$ such that $A \equiv B$

$$(taut - l) \frac{\Gamma, x \hat{\in} A \Rightarrow \Delta}{\Gamma, x \hat{\in} B \Rightarrow \Delta} \quad (taut - r) \frac{\Gamma \Rightarrow \Delta, x \hat{\in} A}{\Gamma \Rightarrow \Delta, x \hat{\in} B}$$

Intersection rules:

$$(\cap \Rightarrow) \frac{\Gamma, x \hat{\in} B, x \hat{\in} C \Rightarrow \Delta}{\Gamma, x \hat{\in} B \cap C \Rightarrow \Delta} \quad (\Rightarrow \cap) \frac{\Gamma \Rightarrow \Delta, x \hat{\in} B \quad \Gamma \Rightarrow \Delta, x \hat{\in} C}{\Gamma \Rightarrow \Delta, x \hat{\in} B \cap C}$$

Inference rules for Kleene connectives:

$$(\neg \hat{\in} \Rightarrow) \frac{\Gamma, x \hat{\in} -A \Rightarrow \Delta}{\Gamma, \neg(x \hat{\in} A) \Rightarrow \Delta} \quad (\Rightarrow \neg \hat{\in}) \frac{\Gamma \Rightarrow \Delta, x \hat{\in} -A}{\Gamma \Rightarrow \Delta, \neg(x \hat{\in} A)}$$

$$(\neg\neg \Rightarrow) \frac{\Gamma, \varphi \Rightarrow \Delta}{\Gamma, \neg\neg\varphi \Rightarrow \Delta} \quad (\Rightarrow \neg\neg) \frac{\Gamma \Rightarrow \Delta, \varphi}{\Gamma \Rightarrow \Delta, \neg\neg\varphi}$$

$$(\vee \Rightarrow) \frac{\Gamma, \varphi \Rightarrow \Delta \quad \Gamma, \psi \Rightarrow \Delta}{\Gamma, \varphi \vee \psi \Rightarrow \Delta} \quad (\Rightarrow \vee) \frac{\Gamma, \Rightarrow \Delta, \varphi, \psi}{\Gamma \Rightarrow \Delta, \varphi \vee \psi}$$

$$(\neg\vee \Rightarrow) \frac{\Gamma, \neg\varphi, \neg\psi \Rightarrow \Delta}{\Gamma, \neg(\varphi \vee \psi) \Rightarrow \Delta} \quad (\Rightarrow \neg\vee) \frac{\Gamma \Rightarrow \Delta, \neg\varphi \quad \Gamma \Rightarrow \Delta, \neg\psi}{\Gamma \Rightarrow \Delta, \neg(\varphi \vee \psi)}$$

$$(\wedge \Rightarrow) \frac{\Gamma, \varphi, \psi \Rightarrow \Delta}{\Gamma, \varphi \wedge \psi \Rightarrow \Delta} \quad (\Rightarrow \wedge) \frac{\Gamma \Rightarrow \Delta, \varphi \quad \Gamma \Rightarrow \Delta, \psi}{\Gamma \Rightarrow \Delta, \varphi \wedge \psi}$$

$$(\neg\wedge \Rightarrow) \frac{\Gamma, \neg\varphi \Rightarrow \Delta \quad \Gamma, \neg\psi \Rightarrow \Delta}{\Gamma, \neg(\varphi \wedge \psi) \Rightarrow \Delta} \quad (\Rightarrow \neg\wedge) \frac{\Gamma \Rightarrow \Delta, \neg\varphi, \neg\psi}{\Gamma \Rightarrow \Delta, \neg(\varphi \wedge \psi)}$$

In all axioms and inference rules, we assume that $x, y, z \in OV$, $A, B \in SE$.

Note that though the axiom $\varphi, \neg\varphi$ is missing in the formulation of CRS, it is in fact derivable in CRS. Indeed, at the atomic level, from A1 and rule $(\neg \hat{\in} \Rightarrow)$ we can deduce that (1) $\vdash_{CRS} x \hat{\in} A, \neg(x \hat{\in} A) \Rightarrow x \hat{\in} A, x \hat{\in} -A$. In turn, from A1 and rule $(\Rightarrow \cap)$ we can obtain (2) $\vdash_{CRS} x \hat{\in} A, x \hat{\in} -A \Rightarrow x \in (A \cap -A)$. Considering that $A \cap -A \equiv \mathbf{0}$, from rule $(taut - l)$ we can deduce that (3) $\vdash_{CRS} x \in (A \cap -A) \Rightarrow x \hat{\in} \mathbf{0}$. Applying cut, from 1), (2), (3) and Axiom A2 we conclude that $\vdash_{CRS} x \hat{\in} A, \neg(x \hat{\in} A) \Rightarrow$, so $\varphi, \neg\varphi \Rightarrow$ holds for atomic φ . The fact that it holds for complex φ too can be shown by structural induction using the inference rules for Kleene connectives .

Lemma 1.

1. The axioms of the system CRS are valid.
2. For any inference rule ρ of CRS and any framework \mathcal{R} for L_{RS} , if the interpretation I of L_{RS} in \mathcal{R} satisfies all the premises of ρ , then I satisfies the conclusion of ρ as well.

Both parts can be easily verified based on the individual clauses of the definition of I given in Definition 6.

Clearly, from the above Lemma we can immediately conclude that:

Corollary 2. *The inference rules of CRS are strongly sound, i.e. they preserve the validity of sequents.*

5 Strong soundness and completeness of the proof system

To prove strong completeness of CRS, we start with simple characterization of valid single-variable atomic sequents.

Definition 9. *For any $A, B \in SE$, we say that:*

1. A is Boolean-equivalent to B , and write $A \equiv B$, iff $A = B$ is a Boolean tautology;
2. A is Boolean-included in B , and write $A \sqsubseteq B$, iff $A \cap B = A$ is a Boolean tautology, i.e. iff $A \cap B \equiv A$.

Proposition 1. *A sequent $\Sigma = x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow x \hat{\in} B_1, \dots, x \hat{\in} B_l$ is valid iff one of the following conditions is satisfied:*

- (1) $A_1 \cap A_2 \cap \dots \cap A_k \equiv \mathbf{0}$
- (2) $B_i \equiv \mathbf{1}$ for some $i \leq l$
- (3) $A_1 \cap A_2 \cap \dots \cap A_k \sqsubseteq B_i$ for some $i \leq l$

Proof. The backward implication follows easily from the semantics of \mathcal{L}_{RS} . To prove the forward implication, we argue by contradiction. Assume that a sequent Σ of the form given above is such that:

- (I) $A_1 \cap A_2 \cap \dots \cap A_k \not\equiv \mathbf{0}$
- (II) $B_i \not\equiv \mathbf{1}$ for each $i \leq l$
- (III) $A_1 \cap A_2 \cap \dots \cap A_k \not\sqsubseteq B_i$ for each $i \leq l$

Define

$$SV_0 = \{X \in SV \mid X \text{ occurs in } \Sigma\}$$

$$SE_0 = \{A \in SE \mid A \text{ contains only set variables in } SV_0\}$$

As SV_0 is finite, $SV_0 = \{X_1, X_2, \dots, X_n\}$ for some n . The counter-model construction is based on the use of the full disjunctive normal form (DNF) of an expression in SE_0 with respect to SV_0 . Such a DNF is of the form

$$\mathbf{X}^{\bar{\epsilon}} = X_1^{\epsilon_1} \cap X_2^{\epsilon_2} \cap \dots \cap X_n^{\epsilon_n}$$

where $\bar{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_n) \in \{-1, 1\}^n$, and $X_j^1 = X_j, X_j^{-1} = \bar{X}_j$.

Let $A = A_1 \cap A_2 \cap \dots \cap A_k$. Then $A \neq \mathbf{0}$ by (I), so we have

$$\text{DNF}(A) = \mathbf{X}^{\bar{\epsilon}_1} \cup \mathbf{X}^{\bar{\epsilon}_2} \cup \dots \cup \mathbf{X}^{\bar{\epsilon}_p}$$

for some $p \geq 1, \bar{\epsilon}_1, \dots, \bar{\epsilon}_p \in \{-1, 1\}^n$. Since $\text{DNF}(E) \equiv E$ for any $E \in SE_0$, then by (III) we get $\text{DNF}(A) \not\sqsubseteq \text{DNF}(B_i)$ for each $i \leq l$. Hence for each $i \leq l$ there is a $j_i \leq p$ such that $\mathbf{X}^{\bar{\epsilon}_{j_i}}$ does not occur in $\text{DNF}(B_i)$.

Let us assign a unique symbol $a^{\bar{\epsilon}} \notin OV \cup SV$ to any $\bar{\epsilon} \in \{-1, 1\}^n$. As the universe of our counter-model \mathcal{R} we take $U = \{x\} \cup \{a^{\bar{\epsilon}} \mid \bar{\epsilon} \in \{-1, 1\}^n\}$, and as the covering underlying the approximation space — $\mathcal{C} = \{C(u) \mid u \in U\}$, where $C(u) = \{u\}$ for $u \neq x$, and $C(x) = \{x, a^{\bar{\epsilon}_1}, a^{\bar{\epsilon}_2}, \dots, a^{\bar{\epsilon}_p}\}$. The valuation of variables v is given by $v(y) = x$ for any $x \in OV$. Finally, to define the valuation of set variables, we first define $\xi(\mathbf{X}^{\bar{\epsilon}_1}) = \{x, a^{\bar{\epsilon}_1}\}$ and $\xi(\mathbf{X}^{\bar{\epsilon}}) = \{a^{\bar{\epsilon}}\}$ for any $\bar{\epsilon} \in \{-1, 1\}^n \setminus \{\bar{\epsilon}_1\}$. Then we put $w(X) = \emptyset$ for $X \in SV \setminus SV_0$, and define w on SV_0 by taking

$$w(X_j) = \bigcup \{\xi(\mathbf{X}^{\bar{\epsilon}}) \mid \bar{\epsilon} \in \{-1, 1\}^n, \epsilon_j = 1\}$$

for $j = 1, 2, \dots, n$ (recall that $SV_0 = \{X_1, X_2, \dots, X_n\}$). It is easy to check that $\mathcal{R} = ((U, \mathcal{C}), v, w)$ is a well-defined semantic framework for L_{RS} , and

$$w^*(A_1 \cap A_2 \cap \dots \cap A_k) = w^*(X) = \{x, a^{\bar{\epsilon}_1}, a^{\bar{\epsilon}_2}, \dots, a^{\bar{\epsilon}_p}\} = C(x) \quad (3)$$

However, as $w^*(A_1 \cap A_2 \cap \dots \cap A_k) = \bigcap_{r=1}^k w^*(A_r) \subseteq w^*(A_j)$ for each $j \leq k$, (3) implies that $C(x) \subseteq w^*(A_j)$ for any $j \leq k$. Since $C(x)$ is the only set $C \in \mathcal{C}$ such that $x \in C$, then from Corollary 1 we obtain $x \in POS(w^*(A_j))$ and $I_{\mathcal{R}} \models x \hat{\in} A_j$ for $j = 1, 2, \dots, k$. On the other hand, as $\mathbf{X}^{\bar{\epsilon}_{j_i}}$ does not occur in $\text{DNF}(B_i)$ for any $i \leq l$, then $a^{\bar{\epsilon}_{j_i}} \notin w^*(\text{DNF}(B_i)) = w^*(B_i)$ for each $i \leq l$, which in view of $a^{\bar{\epsilon}_{j_i}} \in C(x)$ implies $C(x) \not\subseteq w^*(B_i)$ for each $i \leq l$. As a result, $I_{\mathcal{R}} \not\models x \hat{\in} B_i$ for $i = 1, 2, \dots, l$. Thus $I_{\mathcal{R}} \not\models \Sigma$, which ends the proof.

Since L_{RS} has no means for expressing relationships between object variables, Proposition 1 implies a similar result for multi-variable atomic sequents:

Corollary 3. *An atomic sequent $\Sigma \in Seq_{RS}$ is valid if and only if, for some object variable x occurring in Σ , the sequent Σ_x obtained from Σ by deleting all formulas with variables different from x satisfies one of the conditions of Proposition 1.*

The proof is analogous to that of Proposition 1, with the counter-model for a sequent Σ which does not satisfy any of conditions (1),(2),(3) of that Proposition constructed by combining the individual countermodels for all single-variable subsequents of Σ , constructed exactly like in the proof of Proposition 1.

From the results of the preceding section, we can easily conclude that CRS is complete for atomic sequents:

Proposition 2. *If an atomic sequent $\Sigma \in Seq_{RS}$ is valid, then it is derivable in CRS, i.e. $\vdash_{CRS} \Sigma$.*

Proof. For any variable x occurring in Σ , denote by Σ_x the atomic sequent obtained out of Σ by deleting all formulas with variables different from x . Since Σ is valid, then, by Corollary 3, there exists an x such that Σ_x satisfies one of the conditions (1), (2), (3) of Proposition 1. Hence, assuming that $\Sigma_x = x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow x \hat{\in} B_1, \dots, x \hat{\in} B_l$, we have

either (1) $A_1 \cap A_2 \cap \dots \cap A_k \equiv \mathbf{0}$ or (2) $B_i \equiv \mathbf{1}$ for some $i \leq l$
or (3) $A_1 \cap A_2 \cap \dots \cap A_k \sqsubseteq B_i$ for some $i \leq l$

If (1) holds, then from A1 and rule $(\Rightarrow \cap)$ applied $k-1$ times we can obtain (i) $\vdash_{CRS} x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow x \hat{\in} (A_1 \cap \dots \cap A_k)$. Considering that $A_1 \cap \dots \cap A_k \equiv \mathbf{0}$, from rule $(\text{taut} - l)$ and Axioms A1, A2 we get (ii) $\vdash_{CRS} x \hat{\in} (A_1 \cap \dots \cap A_k) \Rightarrow$. Applying cut to (i) and (ii), we obtain $\vdash_{CRS} x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow$, whence $\vdash_{CRS} \Sigma_x$ by weakening.

If (2) holds, then from Axiom A3 and rule $(\text{taut} - r)$ we get $\vdash_{CRS} \Rightarrow B_i$, whence $\vdash_{CRS} \Sigma_x$ by weakening.

Finally, assume that (3) holds. By what was shown for (1), we have (i) $\vdash_{CRS} x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow x \hat{\in} (A_1 \cap \dots \cap A_k)$. For simplicity, denote $A = A_1 \cap \dots \cap A_k$. Then $A \sqsubseteq B_i$, implying $A \cap B_i \equiv A$ by Definition 9, whence from Axiom A1 and rule $(\text{taut} - l)$ we get (ii) $\vdash_{CRS} x \hat{\in} A \Rightarrow x \hat{\in} A \cap B_i$. In turn, by A1 and rule $(\cap \Rightarrow)$ we have (iii) $\vdash_{CRS} x \hat{\in} A \cap B_i \Rightarrow x \hat{\in} B_i$. Applying cut twice to (i), (ii) and (iii), we obtain $\vdash_{CRS} x \hat{\in} A_1, \dots, x \hat{\in} A_k \Rightarrow x \hat{\in} B_i$, which yields $\vdash_{CRS} \Sigma_x$ by weakening.

Thus $\vdash_{CRS} \Sigma_x$ in all three cases. As $\Sigma_x \subseteq \Sigma$ in the standard sense of sequent inclusion, this implies $\vdash_{CRS} \Sigma$ by weakening.

Proposition 2 is the cornerstone for proving the strong completeness theorem for the logic \mathcal{L}_{RS} :

Theorem 1. *The calculus CRS is finitely strongly sound and complete for \vdash_{RS} , i.e., for any finite set of sequents $S \subseteq \text{Seq}_{RS}$ and any sequent $\Sigma \in \text{Seq}_{RS}$, $S \vdash_{RS} \Sigma$ iff $S \vdash_{CRS} \Sigma$.*

Proof. (Sketch) As the backward implication (soundness) follows from Lemma 1 and Corollary 2, it suffices to prove the forward implication (completeness). The proof is by counter-model construction based on Proposition 2 and the maximum saturated sequent construction used e.g. in [1].

We argue by contradiction. Suppose that for a finite set of sequents S and a sequent $\Sigma = \Gamma \Rightarrow \Delta$ we have $S \vdash_{RS} \Sigma$, but Σ is not derivable from S in CRS. We shall construct a counter-model I such that $I \models S$ but $I \not\models \Sigma$.

Denote by $F(S)$ the set of all formulae belonging to at least one of the sides in some sequent in S , and let SV^* be the set of all set variables which occur either in some $\varphi \in F(S)$ or in Σ . Since S is finite, so are $F(S)$ and SV^* . Using the method shown in in [1], we can construct a sequent $\Gamma' \subseteq \Delta'$ such that

- (i) $\Gamma \subseteq \Gamma', \Delta \subseteq \Delta'$
- (ii) $F(S) \subseteq \Gamma' \cup \Delta'$.

(iii) $\Gamma' \Rightarrow \Delta'$ is not derivable from S in CRS.

The construction is carried out by starting with Σ , and then adding consecutively linearly ordered formulas in $F(S)$ to either the left- or the right-hand side of the sequent constructed up to that time, depending on which option results in a sequent still not derivable from S in CRS. Such a construction is possible because if $S \not\vdash_{CRS} (\Gamma_i \Rightarrow \Delta_i)$, then, for any $\varphi \in F(S)$, we cannot have both $S \vdash_{CRS} (\Gamma_i, \varphi \Rightarrow \Delta_i)$ and $S \vdash_{CRS} (\Gamma_i \Rightarrow \Delta_i, \varphi)$, since this would imply $S \vdash_{CRS} (\Gamma_i \Rightarrow \Delta_i)$ by cut.

Call a sequent *saturated* if it is closed under the inference rules in CRS applied backwards, whereby we assume that closure under the Boolean tautology rules (*taut* – *l*), (*taut* – *r*) is limited only to premises with the set expression A in a full disjunctive normal form with respect to the set SV^* . By way of example, a sequent $\Gamma'' \Rightarrow \Delta''$ is closed under rule ($\vee \Rightarrow$) applied backwards iff $\varphi \vee \psi \subseteq \Gamma''$ implies either $\varphi \in \Gamma''$ or $\psi \in \Gamma''$.

Let $\Gamma^* \Rightarrow \Delta^*$ be the extension of $\Gamma' \Rightarrow \Delta'$ to a saturated sequent which is not derivable from $F(S)$ in CRS (it is easy to see that such a sequent exists; note that the restriction on the closure under tautology rules ensures that the closure adds only a finite number of formulas to $\Gamma' \Rightarrow \Delta'$).

Then we can easily see that:

- (1) $\Gamma \subseteq \Gamma^*, \Delta \subseteq \Delta^*$;
- (2) $F(S) \subseteq \Gamma^* \cup \Delta^*$;
- (3) $\Gamma^* \Rightarrow \Delta^*$ is saturated and it is not derivable from S in CRS.

Now let $\Sigma_a = \Gamma_a \Rightarrow \Delta_a$ be a subsequent of $\Gamma^* \Rightarrow \Delta^*$ consisting of all atomic formulas in $\Gamma^* \Rightarrow \Delta^*$. Then by (3) $S \not\vdash_{CRS} \Sigma_a$, and hence also $\not\vdash_{CRS} \Sigma_a$. As Σ_a is atomic, by Proposition 2, this implies that Σ_a is not valid. Accordingly, there exists a framework \mathcal{R} for L_{RS} and an interpretation I of L_{RS} in \mathcal{R} such that $I \not\models \Sigma_a$. We shall prove that I is the desired counter-model for the original sequent Σ too, i.e. that:

$$\text{(A)} \quad I \not\models (\Gamma \Rightarrow \Delta) \qquad \text{(B)} \quad I \models \Sigma \text{ for each } \Sigma \in S$$

Let us start with (A). As $\Gamma \subseteq \Gamma^*, \Delta \subseteq \Delta^*$, then in order to prove (A) it suffices to show that $I \not\models (\Gamma^* \Rightarrow \Delta^*)$. Since the only designated value in the semantics of L_{RS} is \mathbf{t} and $I(\varphi) \in \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ for any formula $\varphi \in F_{RS}$, this means we have to prove that:

$$I(\gamma) = \mathbf{t} \text{ for any } \gamma \in \Gamma^* \qquad I(\delta) \in \{\mathbf{f}, \mathbf{u}\} \text{ for any } \delta \in \Delta^* \qquad (4)$$

As $I \not\models \Sigma_a$, then (4) holds for all atomic formulas $\gamma \in \Gamma^*, \delta \in \Delta^*$. To show that it holds for complex formulas too, we prove that, for any complex formula φ , the following is true:

$$\text{(A1)} \quad I(\varphi) = \begin{cases} \mathbf{t} & \text{if } \varphi \in \Gamma^* \\ \mathbf{f} & \text{if } \neg\varphi \in \Gamma^* \end{cases} \qquad \text{(A2)} \quad I(\varphi) \in \begin{cases} \{\mathbf{f}, \mathbf{u}\} & \text{if } \varphi \in \Delta^* \\ \{\mathbf{t}, \mathbf{u}\} & \text{if } \neg\varphi \in \Delta^* \end{cases}$$

The proof is by induction on the complexity of φ , and (A1) and (A2) are proved simultaneously, making use of the fact that Σ^* as a saturated sequent is closed under all rules in CRS applied backwards.

To illustrate the method used, consider first the case of $\xi = \neg(x \hat{\in} A)$.

If $\xi \in \Gamma^*$, then $x \hat{\in} - A$ is also in Γ^* , since Σ^* is closed under rule $(\neg \hat{\in} \Rightarrow)$ applied backwards. As (4) holds for all atomic formulas and $x \hat{\in} - A$ is atomic, this yields $I(x \hat{\in} - A) = \mathbf{t}$. However, from Definition 6 and Corollary 1 we can easily conclude that

$$I(x \hat{\in} A) = \begin{cases} \mathbf{t} & \text{iff } I(x \hat{\in} - A) = \mathbf{f} \\ \mathbf{f} & \text{iff } I(x \hat{\in} - A) = \mathbf{t} \\ \mathbf{u} & \text{iff } I(x \hat{\in} - A) = \mathbf{u} \end{cases} \quad (5)$$

which implies $I(\xi) = I(\neg(x \hat{\in} A)) = \mathbf{t}$ by Definition 6.

In turn, if $\xi \in \Delta^*$, then $x \hat{\in} - A$ is also in Δ^* , since Σ^* is closed under rule $(\Rightarrow \neg \hat{\in})$ applied backward. As (4) holds for $x \hat{\in} - A$, then $I(x \hat{\in} - A) \in \{\mathbf{f}, \mathbf{u}\}$, whence in view of (5) we get $I(x \hat{\in} A) \in \{\mathbf{t}, \mathbf{u}\}$. In consequence, $I(\xi) = I(\neg(x \hat{\in} A)) \in \{\mathbf{f}, \mathbf{u}\}$ by Definition 6. Thus (A1) and (A2) hold for ξ .

As another example, assume that (A1), (A2) hold for φ, ψ , and that $\xi = \varphi \vee \psi$. If $\xi \in \Gamma^*$, then either $\varphi \in \Gamma^*$ or $\psi \in \Gamma^*$, since Σ^* is closed under rule $(\vee \Rightarrow)$ applied backwards. As a result, by the inductive assumption on φ, ψ we have either $I(\varphi) = \mathbf{t}$ or $I(\psi) = \mathbf{t}$, and consequently $I(\xi) = \mathbf{t}$ by Definition 6. In turn, if $\xi \in \Delta^*$, then $\varphi, \psi \in \Delta^*$, and $I(\varphi), I(\psi) \in \{\mathbf{f}, \mathbf{u}\}$ by the inductive assumption, whence $I(\xi) \in \{\mathbf{f}, \mathbf{u}\}$ by Definition 6, too. As a result, (A1) and (A2) hold for ξ too.

The proof of other cases is similar, and is left to the reader.

It remains to prove (B), i.e., to show that $I \models \Sigma_0$ for each $\Sigma_0 \in S$. So let $\Sigma_0 \in S$. Then $\Sigma_0 = \varphi_1, \dots, \varphi_k \Rightarrow \psi_1, \dots, \psi_l$ for some integers k, l and formulas $\varphi_i, \psi_j, i = 1, \dots, k, j = 1, \dots, l$. Clearly, we cannot have both $\{\varphi_1, \dots, \varphi_k\} \subseteq \Gamma^*$ and $\{\psi_1, \dots, \psi_l\} \subseteq \Delta^*$, for then $\Gamma^* \Rightarrow \Delta^*$ would be derivable from Σ_0 , and hence from S , by weakening. Since $F(S) \subseteq \Gamma^* \cup \Delta^*$, this implies that either $\varphi_i \in \Delta^*$ for some i , or $\psi_j \in \Gamma^*$ for some j . Hence by (A1) and (A2), which we have already proved, we have either $I \not\models \varphi_i$ for some i , or $I \models \psi_j$ for some j , which implies that $I \models \Sigma$.

6 Conclusions and future work

The crucial feature of three-valued logic presented in the paper is a complete mechanism for reasoning about atomic formulas representing three-valued, rough membership of the objects of the universe in rough sets. The three values taken by the rough membership relation correspond to “crisp” membership of objects in the three basic regions of a rough set: the positive, negative and boundary one. The strong version of semantics with the single designated value \mathbf{t} adopted in the paper amounts to identifying membership of an object x in a rough set A with its belonging to the positive region of A . However, in many applications it is advisable to identify the above membership with x belonging to either the

positive region or the boundary region of A — which corresponds to taking also \mathbf{u} as a designated value. The latter option, which leads to paraconsistent logic, will be the subject of further work.

The use of connectives to form complex formulas enhances the expressive power of the language, but the Kleene 3-valued connectives used here are just one possible choice. Other interesting option, to be explored in the future, are the Lukasiewicz 3-valued connectives (including implication), and the non-deterministic connectives observing the rough set Nmatrix considered in [3]. Exploring these choices is another direction for future work. A still another is to consider a richer language which allows for expressing relationships between objects — and here the most immediate future task will be extending the results of this paper to a language featuring the subordination relation of [13].

The authors would like to thank the anonymous referees for very helpful comments on the paper, including suggestions for directions of future work — which we have included above.

References

1. Avron, A., Konikowska., B. and Ben-Naim, J.: Processing Information from a Set of Sources. In: Towards Mathematical Philosophy, Series: Trends in Logic , Vol. 28, Makinson, David; Malinowski, Jacek; Wansing, Heinrich (Eds.), pp.165-186, Springer Verlag (2008)
2. Avron, A.: Logical Non-determinism as a Tool for Logical Modularity: An Introduction. In: We Will Show Them: Essays in Honor of Dov Gabbay, Vol 1 (S. Artemov, H. Barringer, A. S. d'Avila Garcez, L. C. Lamb, and J. Woods, eds.), pp. 105–124. College Publications (2005).
3. Avron, A. and Konikowska., B.: Rough Sets and 3-valued Logics. *Studia Logica*, vol. **90** (1), pp. 69–92 (2008).
4. Avron, A. and Lev, I.: Non-deterministic Multiple-valued Structures. *Journal of Logic and Computation* 15, pp. 241–261 (2005).
5. Balbiani, P. and Vakarelov, D.: A modal Logic for Indiscernibility and Complementarity in Information Systems. *Fundamenta Informaticae* 45, pp. 173–194 (2001).
6. Banjeeri, M.: Rough sets and 3-valued Lukasiewicz logic. *Fundamenta Informaticae* 32, pp. 213–220 (1997).
7. Demri, S., Orłowska, E., Vakarelov, D.: Indiscernibility and complementarity relations in information systems. In: Gerbrandy, J., Marx, M., de Rijke, M., Venema, Y. (eds.) JFAK: Esays dedicated to Johan van Benthem on the occasion of his 50-th Birthday. Amsterdam University Press (1999).
8. Deneva, A. and Vakarelov, D.: Modal Logics for Local and Global Similarity Relations. *Fundamenta Informaticae*, vol 31, No 3,4, pp. 295–304 (1997).
9. Duentzsch, I. Konikowska, B.: A multimodal logic for reasoning about complementarity. *Journal for Applied Non-Classical Logics*, Vol. 10, No 3–4, pp. 273-302 (2000).
10. Iturrioz, L.: Rough sets and three-valued structures. In: Orłowska, E. (editor), *Logic at Work: Essays Dedicated to the Memory of Helena Rasiowa*. Studies in Fuzziness and Soft Computing, vol. 24, pp. 596–603, Physica-Verlag (1999).
11. Kleene, S.C.: Introduction to metamathematics, D. van Nostrad Co. (1952).

12. Konikowska, B.: A logic for reasoning about relative similarity. Special Issue of *Studia Logica*, E. Orłowska, H. Rasiowa eds., Reasoning with incomplete information. *Studia Logica* 58, pp. 185–226 (1997).
13. Konikowska, B.: Three-Valued Logic for Reasoning about Covering-based Rough Sets. In: Special Volume Dedicated to the Memory of Z. Pawlak, *Intelligent Systems Reference Library*, Springer [to appear] (2012).
14. Lin, T.Y. and Cercone, N. (eds.): *Rough sets and Data Mining. Analysis of Imprecise Data*, Kluwer, Dordrecht (1997).
15. Øhrn, A., Komorowski, J., Skowron, A. and Synak, P.: The design and implementation of a knowledge discovery toolkit based on rough sets — The ROSETTA system. In: Polkowski, L. and Skowron, A. (eds.), *Rough Sets in Knowledge Discovery 1. Methodology and Applications*. Physica Verlag, Heidelberg, pp. 376–399 (1998).
16. Orłowska, E.: Reasoning with Incomplete Information: Rough Set Based Information Logics. In: *Proceedings of SOFTEKS Workshop on Incompleteness and Uncertainty in Information Systems*, pp.16-33, (1993).
17. Pagliani, P. Rough set theory and logic-algebraic structures. In: Orłowska, E. (editor), *Incomplete Information: Rough Set Analysis*. Studies in Fuzziness and Soft Computing, vol. 13, pp. 109–190, Physica-Verlag (1998).
18. Pawlak, Z.: Rough Sets, *Intern. J. Comp. Inform. Sci.*, 11, 341–356 (1982).
19. Pawlak, Z.: *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer, Dordrecht (1991).
20. Pawlak, Z.: Rough set approach to knowledge-based decision support, *European Journal of Operational Research* 29(3), pp. 1–10 (1997).
21. Pawlak, Z.: Rough sets theory and its applications to data analysis. *Cybernetics and Systems* 29, pp. 661–688 (1998).
22. Pomykała, J.A.: Approximation operations in approximation space. *Bull. Pol. Acad. Sci.* 35(9-10), pp. 653–662 (1987).
23. Sen J., Chakraborty, M.K.: A study of interconnections between rough and 3-valued Lukasiewicz logics. *Fundamenta Informaticae* 51, 311–324, 2002.
24. Vakarelov, D.: Information Systems, Similarity Relations and Modal Logics. In: E. Orłowska (ed.) *Incomplete Information: Rough Set Analysis*, pp. 492-550. Studies in Fuzziness and Soft Computing, Physica-Verlag Heidelberg New York (1998).
25. Yao, Y.Y.: Relational interpretations of neighborhood operators and rough set approximation operators. *Information Sciences* 111 (1–4), pp. 239–259 (1998).
26. Yao, Y.Y.: On generalizing rough set theory. In: *The 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrc) 2003*. LNCS vol. 2639, pp. 44-51 (2003).
27. Zakowski, W.: On a concept of rough sets. *Demonstratio Mathematica* XV, 1129-1133 (1982).
28. Zhang, Y.-L. Li, J.J. and Wu, W.-Z.: On axiomatic characterizations of three types of covering-based approximation operators. *Information Sciences* 180, pp. 174–187 (2010).
29. Zhu, W., Wang, F.-Y.: On three types of covering-based rough sets. *IEEE Transactions on Knowledge and Data Engineering* 19(8), pp. 1131–1144 (2007).

Doing the right things – trivalence in deontic action logic

Piotr Kulicki and Robert Trypuz*

John Paul II Catholic University of Lublin

Abstract. Trivalence is quite natural for deontic action logic, where actions are treated as good, neutral or bad. We present the ideas of trivalent deontic logic after J. Kalinowski and its realisation in a 3-valued logic of M. Fisher and two systems designed by the authors of the paper: a 4-valued logic inspired by N. Belnap’s logic of truth and information and a 3-valued logic based on nondeterministic matrices. Moreover, we combine Kalinowski’s idea of trivalence with deontic action logic based on boolean algebra.

Keywords: deontic action logic, many-valued logic, Kalinowski’s deontic logic, Dunn-Belnap’s four-valued matrix

Introduction

Deontic logic can be seen as a formal tool for analysing rational agent’s behaviour in the context of systems of norms. Two main approaches within it can be pointed out: in one of them deontic notions, such as permission, forbiddance or obligation are attributes of situations (in the language – propositions), in the other they are attributes of actions (in the language – names). We are interested in the latter one – deontic action logic (DAL), introduced in the work of G.H. von Wright [13], which is usually treated as the beginning of modern deontic logic.

The first attempt at defining the formal semantics of DAL made by J. Kalinowski [5] has already taken the form of three-valued tables defining truth values of propositions built with deontic operators (permission, forbiddance and obligation) for different types of actions and negations of actions. M. Fisher [4] used a similar methodology but introduced and analysed more operations on actions, namely conjunction and disjunction. More recently a multi-valued approach to DAL appeared in the paper of A. Kouznetsov [7].

Trivalence is quite natural for deontic action logic. The three values in that context refer to actions that are respectively good, neutral or bad. This general idea needs, of course, more detailed specification to be used as a basis for a formal system. In the present paper we just point out the possible ways of using the techniques of multi-valued logic in the area of norms. A more serious study of their application would require much longer and detailed investigations.

* This research was supported by the National Science Center of Poland (DEC-2011/01/D/HS1/04445).

Although other semantic tools have also been successfully applied to DAL we believe that multi-valued semantics for that logic is worth further research. There are two main advantages of that approach (the same as of the use of many-valued techniques in other branches of logic): it is intuitively clear and produces systems with nice computational properties.

We start the paper from a brief recall of Kalinowski's ideas and Fisher's tri-valued DAL. Then we adapt to DAL the ideas introduced into multi-valued logic in a different context: the bilattice known as *FOUR* (bilattice of truth and information) with the respective *FOUR*-valued matrices introduced by N. Belnap [2] and non-deterministic matrices introduced by A. Avron and I. Lev [1]. Finally, we study the relation between DAL, based on boolean algebra introduced by K. Segerberg in [9], and Kalinowski's ideas.

1 Language of deontic action logic

The language of DAL is defined in Backus-Naur notation in the following way:

$$\begin{aligned} \varphi ::= & \alpha = \alpha \mid \mathbf{O}(\alpha) \mid \mathbf{P}(\alpha) \mid \mathbf{P}_w(\alpha) \mid \mathbf{F}(\alpha) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \rightarrow \varphi \mid \varphi \equiv \varphi \\ & \hspace{15em} (1) \\ \alpha ::= & a_i \mid \mathbf{0} \mid \mathbf{1} \mid \bar{\alpha} \mid \alpha \sqcup \alpha \mid \alpha \sqcap \alpha \hspace{10em} (2) \end{aligned}$$

where a_i belongs to a finite set of basic actions Act_0 , “ $\mathbf{0}$ ” is the impossible action and “ $\mathbf{1}$ ” is the universal action; “ $\alpha = \beta$ ” means that α is identical with β (the last three elements of the language will be used only in section 4); “ $\mathbf{O}(\alpha)$ ” – α is obligatory; “ $\mathbf{P}(\alpha)$ ” – α is strongly permitted (i.e. its performance is permitted in combination with any action); “ $\mathbf{P}_w(\alpha)$ ” – α is weakly permitted (i.e. its performance is not forbidden); “ $\mathbf{F}(\alpha)$ ” – α is forbidden, “ $\alpha \sqcup \beta$ ” – α or β (a free choice between α and β); “ $\alpha \sqcap \beta$ ” – α and β (parallel execution of α and β); “ $\bar{\alpha}$ ” – not α (complement of α). Further, for fixed Act_0 , by Act we shall understand the set of formulae defined by (2).

In our considerations we understand actions as types or descriptions rather than individual events in time and space. Thus e.g. “ $\alpha \sqcup \beta$ ” is a description of all actions that are covered by description “ α ” or by description “ β ” (see also [8]).

2 3-valued deontic logics of Kalinowski and Fisher

2.1 Kalinowski's deontic logic

Kalinowski believed that norms like propositions are true or false. In his approach norms describe relations of permission, obligation and prohibition between agents and actions. Thus, if, for instance, permission holds between agent i and action α , then the norm expressing that state of affairs is true. Logical value of norms depends on moral value of actions – *to fulfil norms means to do the right things*. Kalinowski assumed that every action *in genere* is either good (g), or bad (b) or neutral (n), although actions *in concreto* are always

good or bad. Good (bad) actions are such by nature and remain good (bad) in all circumstances. On the other hand neutral actions are those which in some circumstances are good and in other circumstances bad. Kalinowski expressed his philosophical intuitions concerning the meaning of deontic concepts of weak permission, obligation and prohibition by the following matrix:

α	$P_w(\alpha)$	$O(\alpha)$	$F(\alpha)$
b	0	0	1
n	1	0	0
g	1	1	0

One can see that obligatory actions are those which are always good, prohibited actions are those which are always bad and weakly permitted actions are (always) good or neutral.

There is only one internal operator in Kalinowski's logic – action complement. Each action α has its complement (or negations) $\bar{\alpha}$. Action negation is defined by the following matrix:

α	$\bar{\alpha}$
b	g
n	n
g	b

A complement of a good action is bad, a complement of a bad action is good and finally a complement of a neutral action is also neutral. Kalinowski states in [6] that his theory can be enriched by other action operators such as parallel execution or indeterministic choice, but at the same time he is very sceptical about the applicability of theories more expressive than his own deontic theory.

2.2 Fisher's trivalent matrices

One of the extensions of Kalinowski's logic is Fisher's deontic logic. Fisher introduced two operators: parallel execution and indeterministic choice (see two matrixes below), which were missing in Kalinowski's theory.

\sqcap	b	n	g
b	b	b	b
n	b	n	g
g	b	g	g

\sqcup	b	n	g
b	b	b	g
n	b	n	g
g	g	g	g

The two operators are De Morgan duals, i.e.

$$a \sqcup b = \overline{\bar{a} \sqcap \bar{b}} \tag{3}$$

Basic actions are *per se* good, bad or neutral, while the deontic value of other actions, that can be described as combinations of basic actions made with the use of action operators, can be computed using the matrices.

3 Alternative matrix systems

3.1 4-valued deontic semantics

Kalinowski stated in [6] that Fisher’s matrixes are “natural”. We find them, at least at certain points, problematic. In Fisher’s approach a combination (parallel execution) of good and bad actions is bad and free choice between such actions is good. Execution of an action that is composed of two parts of which one is good and the other is bad can be understood as a conflict of norms or values. We want to study formally different possibilities of solving such conflicts.

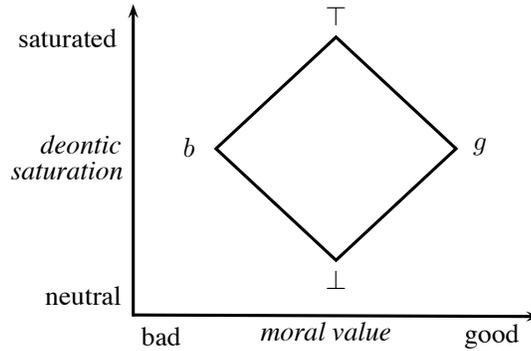
As a first alternative for Fisher’s approach let us investigate the solution in which in such a situation a statement about what is good to do is to be made by an agent. In such a situation the action itself can be seen as neutral. That cannot be achieved with trivalent matrixes with the preservation of associativity, since we would have:

$$(g \sqcap g) \sqcap b = g \sqcap b = n$$

and

$$g \sqcap (g \sqcap b) = g \sqcap n = g.$$

For that reason we take a structure inspired by Belnap’s construction concerning truth information, replacing them respectively by moral value and deontic saturation depicted in the following diagram:



Operator \sqcap is interpreted as supremum in the structure and operator \sqcup as infimum. Negation of \top is \perp and negation of \perp is \top .

The same definitions of the operators can be expressed by the following matrixes:

\sqcap	b	\perp	\top	g
b	b	b	\top	\top
\perp	b	\perp	\top	g
\top	\top	\top	\top	\top
g	\top	g	\top	g

\sqcup	b	\perp	\top	g
b	b	\perp	b	\perp
\perp	\perp	\perp	\perp	\perp
\top	b	\perp	\top	g
g	\perp	\perp	g	g

a	\bar{a}
b	g
\perp	\perp
\top	\top
g	b

The interpretation of deontic atoms is defined by the following matrix:

a	$P_w(a)$	$O(a)$	$F(a)$
b	0	0	1
\perp	1	0	0
\top	1	0	0
g	1	1	0

The matrix shows that intuitively values \perp and \top are treated both as *neutral*. Thus, in a sense, the system remains trivalent, though formally there are *FOUR* values that can be attached to actions.

3.2 A system based on non-deterministic matrices

Another possible approach solves the problem of the conflict of norms by stating that what should be done depends on situation and cannot be predicted deterministically. We formalise it by using nondeterministic matrices. The idea of nondeterministic attachment of deontic value to actions was present in [7] but we use different content of matrix and a more elaborated formalism introduced in [1]. Definition of indeterministic matrices after Avron and Lev is as follows.

- N-matrix for language L is a triple $M = \langle V, D, O \rangle$, where:
 - V is a non-empty set (of truth values);
 - D is a non-empty proper subset of V (designated values);
 - O includes an n -ary function $\underline{x} : V^n \rightarrow 2^V - \emptyset$ for every n -ary operator x of L
- Let M be N-matrix for L . An M -valuation v is a function $v : F_L \rightarrow V$ such that for every n -ary operator x of L and every $\alpha_1, \dots, \alpha_n$,

$$v(x(\alpha_1, \dots, \alpha_n)) \in \underline{x}(v(\alpha_1), \dots, v(\alpha_n)).$$

- A model in M is defined in a usual way.

The following nondeterministic matrices define the deontic value of combined actions:

a	\bar{a}	\sqcap	b	n	g	\sqcup	b	n	g
b	g	b	$\{b\}$	$\{b\}$	$\{b, n, g\}$	b	$\{b\}$	$\{b, n\}$	$\{b, g\}$
n	n	n	$\{b\}$	$\{n\}$	$\{g\}$	n	$\{b, n\}$	$\{n\}$	$\{n, g\}$
g	b	g	$\{b, n, g\}$	$\{g\}$	$\{g\}$	g	$\{b, g\}$	$\{n, g\}$	$\{g\}$

The interpretation of deontic atoms is as in Kalinowski's formalisation.

The most interesting cases are those in which good and bad actions are combined with the operations of parallel execution and free choice. The result in the former case is treated as fully indeterministic. Depending on the estimation of how good is a good component of an action and how bad is its bad component one can treat such a compound action as good, neutral and bad. For the latter case we assume, as a general rule for creating the matrix, that a collection of values of two actions is a value of free choice between them. Thus a free choice between a good and a bad action cannot be neutral, as it was defined in the 4-valued system from the previous section. However, we may also use an alternative non-deterministic matrix in which value n is also possible in this case.

3.3 Some formulas

The following formulas are tautologies in all of the systems defined in the previous sections.

$$F(\alpha) \equiv \neg P_w(\alpha) \quad (4)$$

$$O(\alpha) \rightarrow P_w(\alpha) \quad (5)$$

$$O(\alpha) \equiv \neg P_w(\bar{\alpha}) \quad (6)$$

$$P_w(\alpha) \wedge P_w(\beta) \rightarrow P_w(\alpha \sqcap \beta) \quad (7)$$

$$P_w(\alpha) \wedge P_w(\beta) \rightarrow P_w(\alpha \sqcup \beta) \quad (8)$$

$$P_w(\alpha \sqcup \beta) \rightarrow P_w(\alpha) \vee P_w(\beta) \quad (9)$$

In contrast the following formulas are valid respectively only in Fisher's logic and only in the 4-valued logic:

$$F(\alpha) \rightarrow F(\alpha \sqcap \beta) \quad (10)$$

$$O(\alpha) \rightarrow P_w(\alpha \sqcup \beta) \quad (11)$$

4 Trivalence and deontic action logics based on boolean algebra

4.1 Fundamentals of DAL based on boolean algebra

Deontic action logic based on boolean algebra was introduced by K. Segerberg in [9] and recently studied in [3, 10–12]. Fundamental axioms of \mathcal{DAL} systems are those of Segerberg for $\mathcal{B.O.D.}$, i.e.:

$$P(\alpha \sqcup \beta) \equiv P(\alpha) \wedge P(\beta) \quad (12)$$

$$F(\alpha \sqcup \beta) \equiv F(\alpha) \wedge F(\beta) \quad (13)$$

$$\alpha = \mathbf{0} \equiv F(\alpha) \wedge P(\alpha) \quad (14)$$

It is important to note that permission (prohibition) axiomatised above characterise permitted (prohibited) actions as always permitted (prohibited), i.e. permitted (prohibited) in combination with any action:

$$P(\alpha) \rightarrow P(\alpha \sqcap \beta) \quad (15)$$

$$F(\alpha) \rightarrow F(\alpha \sqcap \beta) \quad (16)$$

Deontic action model for \mathcal{DAL} is a structure $\mathcal{M} = \langle \mathcal{DAF}, \mathcal{I} \rangle$, where $\mathcal{DAF} = \langle E, Leg, Ill \rangle$ is a *deontic action frame* in which $E = \{e_1, e_2, \dots, e_n\}$ is a *nonempty*

set of possible outcomes (events), Leg and Ill are subsets of E and should be understood as sets of legal and illegal outcomes respectively. The basic assumption is that there is no outcome which is legal and illegal:

$$Ill \cap Leg = \emptyset \quad (17)$$

$\mathcal{I} : Act \rightarrow 2^E$ is an interpretation function for \mathcal{DAF} defined as follows:

$$\mathcal{I}(a_i) \subseteq E, \text{ for } a_i \in Act_0 \quad (18)$$

$$\mathcal{I}(\mathbf{0}) = \emptyset \quad (19)$$

$$\mathcal{I}(\mathbf{1}) = E \quad (20)$$

$$\mathcal{I}(\alpha \sqcup \beta) = \mathcal{I}(\alpha) \cup \mathcal{I}(\beta) \quad (21)$$

$$\mathcal{I}(\alpha \sqcap \beta) = \mathcal{I}(\alpha) \cap \mathcal{I}(\beta) \quad (22)$$

$$\mathcal{I}(\bar{\alpha}) = E \setminus \mathcal{I}(\alpha) \quad (23)$$

Additionally we accept that the interpretation of every atom is a singleton:

$$\overline{\overline{\mathcal{I}(\delta)}} = 1 \quad (24)$$

where δ is an atom of Act . A basic intuition is such that an atomic action corresponding to (a set with) one event/outcome is indeterministic. It is important to note two things in this place. The first one is that $\mathcal{I}(\delta)$ is a subset of either Leg , or Ill or $-Leg \cap -Ill$ and the second one is that in every situation an agent's action has only one outcome, which means in practice that what agents really do is to carry out atomic actions.

The definition of an interpretation function makes it clear that every action generator is interpreted as a set of (its) possible outcomes, the impossible action has no outcomes, the universal action brings about all possible outcomes, operations “ \sqcup ”, “ \sqcap ” between actions and “ $-$ ” on a single action are interpreted as set-theoretical operations on interpretations of actions. A class of models defined as above will be represented by \mathbf{C} . Satisfaction conditions for the primitive formulas of \mathcal{DAL} in any model $\mathcal{M} \in \mathbf{C}$ are defined as follows:

$$\begin{aligned} \mathcal{M} \models \mathbf{P}(\alpha) &\iff \mathcal{I}(\alpha) \subseteq Leg \\ \mathcal{M} \models \mathbf{F}(\alpha) &\iff \mathcal{I}(\alpha) \subseteq Ill \\ \mathcal{M} \models \alpha = \beta &\iff \mathcal{I}(\alpha) = \mathcal{I}(\beta) \\ \mathcal{M} \models \neg\varphi &\iff \mathcal{M} \not\models \varphi \\ \mathcal{M} \models \varphi \wedge \psi &\iff \mathcal{M} \models \varphi \text{ and } \mathcal{M} \models \psi \end{aligned}$$

Action α is strongly permitted iff all of its possible outcomes are legal. It means in practice that if α is permitted, then it is permitted in combination with any action (cf. thesis 15). The same is true for forbiddance.

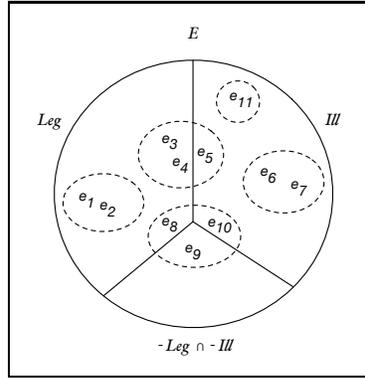


Fig. 1. Five dashed line ovals illustrate some interpretations of \mathcal{DAL} actions. This model includes events which are neither legal nor illegal.

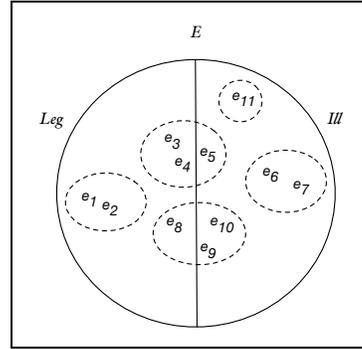


Fig. 2. This model of \mathcal{DAL} is closed in the sense that every event is either legal or illegal.

4.2 Embedding Kalinowski's deontic logic into \mathcal{DAL}

To embed Kalinowski's deontic logic into \mathcal{DAL} we need to make a few additional assumptions. First we assume that Leg and Ill are sets of good and bad events respectively. Then we need to exclude from the models the events which are neither good nor bad, since in Kalinowski's approach each action event is either good or bad. As a result we'd like to obtain models similar to the illustrated below in figure 2.

To obtain the intended result we add a new axiom to \mathcal{DAL} which states that an action is either good or bad or has only good or bad components:

$$P(\delta) \vee F(\delta), \text{ for } \delta \text{ being an atom} \quad (25)$$

Axiom 25 explicitly says that action atoms are good or bad. Additionally we assume that there are actions which are neither good nor bad, to make room for neutral ones:

$$\neg F(\mathbf{1}) \wedge \neg P(\mathbf{1}) \quad (26)$$

Finally, we express Kalinowski's assumption that the complement of a good action is bad:

$$P(\alpha) \equiv F(\bar{\alpha}) \quad (27)$$

The last axiom restricts models of \mathcal{DAL} to the ones illustrated in figure 3. It also shows that "P" and "F" refer to Kalinowski's obligation and forbiddance respectively. They also satisfy semantical conditions restricting obligatory actions only to the good ones and the forbidden actions only to the bad ones. It is also worth noting that each generator (and atom) in \mathcal{DAL} satisfying all the axioms introduced above is interpreted as Leg or Ill (see figure 3).

Finally, we obtain a structure similar to the one resembling Belnap's bilattice from section 3.1. To preserve the intuitions of boolean algebra in figure 4 we reversed the order of saturation.

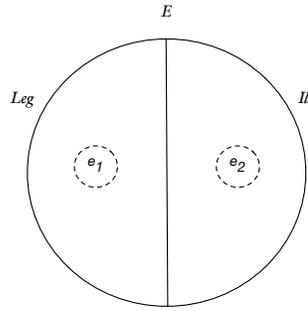


Fig. 3. Model for Kalinowski’s deontic logic

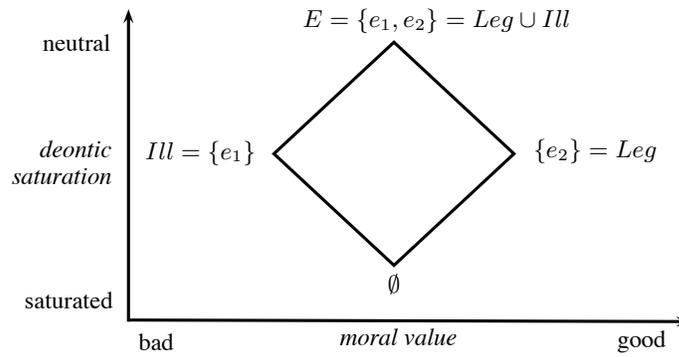


Fig. 4.

It turned out that only the impossible action \emptyset is saturated. Both actions \emptyset and E are neutral.

Since there are only four elements in the algebra we can represent it by the following matrices.

\sqcap	<i>Ill</i>	<i>E</i>	\emptyset	<i>Leg</i>
<i>Ill</i>	<i>Ill</i>	<i>Ill</i>	\emptyset	\emptyset
<i>E</i>	<i>Ill</i>	<i>E</i>	\emptyset	<i>Leg</i>
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
<i>Leg</i>	\emptyset	<i>Leg</i>	\emptyset	<i>Leg</i>

\sqcup	<i>Ill</i>	<i>E</i>	\emptyset	<i>Leg</i>
<i>Ill</i>	<i>Ill</i>	<i>E</i>	<i>Ill</i>	<i>E</i>
<i>E</i>	<i>E</i>	<i>E</i>	<i>E</i>	<i>E</i>
\emptyset	<i>Ill</i>	<i>E</i>	\emptyset	<i>Leg</i>
<i>Leg</i>	<i>E</i>	<i>E</i>	<i>Leg</i>	<i>Leg</i>

<i>a</i>	\bar{a}
<i>Ill</i>	<i>Leg</i>
<i>E</i>	\emptyset
\emptyset	<i>E</i>
<i>Leg</i>	<i>Ill</i>

The only difference between these matrices and 4-valued matrices from section 3.1 lies in the definition of negation. There negation of \perp is \perp and negation of \top is \top , here respective values $-E$ and \emptyset interchange.

Kalinowski’s weak permission can be defined in the following way:

$$P_K(\alpha) \triangleq P(\alpha) \vee N(\alpha) \tag{28}$$

where

$$\mathbf{N}(\alpha) \triangleq \alpha = 0 \vee \neg(\mathbf{P}(\alpha) \vee \mathbf{F}(\alpha)) \quad (29)$$

Action is (syntactically) neutral (**N**) iff it is impossible or is neither good nor bad. Neutrality of the impossible action is assumed for the homogeneity reason. The impossible action is the only one which is at the same time good, bad and neutral. The operator defined in such a way satisfies the desired property that α is neutral iff its complement is neutral too:

$$\mathbf{N}(\alpha) \equiv \mathbf{N}(\bar{\alpha}) \quad (30)$$

A weakly permitted action (in Kalinowski's sense) is thus defined as the one that is either good or neutral.

a	$\mathbf{P}_K(a)$	$\mathbf{P}(a)$	$\mathbf{F}(a)$
<i>Ill</i>	0	0	1
<i>E</i>	1	0	0
\emptyset	1	0	0
<i>Leg</i>	1	1	0

For \mathbf{P}_K the only axiom of Kalinowski's deontic logic K_1 can be proved:

$$\neg\mathbf{P}_K(\bar{\alpha}) \rightarrow \mathbf{P}_K(\alpha) \quad (31)$$

One can also check that the following formula is a theorem of extended \mathcal{DAL} :

$$\mathbf{P}(\alpha) \equiv \neg\mathbf{P}_K(\bar{\alpha}) \quad (32)$$

Conclusions and further work

We have discussed systems of deontic action logic based on three values of actions: good, bad and neutral. In some of them the neutral value was divided into two. The proposed formalism was useful for discussing different approaches to conflicts of norms or values. Future works will be directed to the application of such logics for modelling agents in the context of norms.

References

1. Arnon Avron and Iddo Lev. Non-deterministic multiple-valued structures. *Journal of Logic and Computation*, 15(3):241–261, June 2005.
2. Nuel Belnap. A useful four-valued logic. In J.M. Dunn and G. Epstein, editors, *Modern uses of multiple-valued logic*, pages 8–37. 1977.
3. Pablo F. Castro and T.S.E. Maibaum. Deontic action logic, atomic boolean algebra and fault-tolerance. *Journal of Applied Logic*, 7(4):441–466, 2009.
4. M. Fisher. A three-valued calculus for deontic logic. *Theoria*, 27:107–118, 1961.
5. J. Kalinowski. Theorie des propositions normativess. *Studia Logica*, 1:147–182, 1953.

6. Jerzy Kalinowski. *La logique des normes*. Presses Universitaires de France, 1972.
7. Andrei Kouznetsov. Quasi-matrix deontic logic. In Alessio Lomuscio and Donald Nute, editors, *Deontic Logic in Computer Science*, volume 3065 of *Lecture Notes in Computer Science*, pages 191–208. Springer Berlin / Heidelberg, 2004.
8. Piotr Kulicki and Robert Trypuz. How to build a deontic action logic. In *Logica 2011*. To appear.
9. Krister Segerberg. A deontic logic of action. *Studia Logica*, 41:269–282, 1982.
10. Robert Trypuz and Piotr Kulicki. A systematics of deontic action logics based on boolean algebra. *Logic and Logical Philosophy*, 18:263–279, 2009.
11. Robert Trypuz and Piotr Kulicki. Towards metalogical systematisation of deontic action logics based on boolean algebra. In *Proc. 10th International Conference Deontic Logic in Computer Science*, volume 6181 of *Lecture Notes in Computer Science*. Springer, 2010.
12. Robert Trypuz and Piotr Kulicki. A norm-giver meets deontic action logic. *Logic and Logical Philosophy*, 20:59–72, 2011.
13. G. H. von Wright. Deontic logic. *Mind*, LX(237):1–15, 1951.

Trivalent Logics Arising from L-models for the Lambek Calculus with Constants

Stepan Kuznetsov

Moscow State University
skuzn@inbox.ru

Abstract. We consider language models for the Lambek calculus that allows empty antecedents and enrich them with constants for the empty language and for the language containing only the empty word. No complete calculi are known with respect to these semantics, and in this paper we consider several trivalent systems that arise as fragments of these models' logics.

1 The Lambek Calculus and L-models

We consider the calculus L^* (the Lambek calculus allowing empty antecedents). This calculus is a variant of the original Lambek calculus [4] that allows empty antecedents.

The set $\text{Pr} = \{p_1, p_2, p_3, \dots\}$ is called the set of *primitive types*. Types of L^* are built from primitive types using three binary connectives: \backslash (*left division*), $/$ (*right division*), and \cdot (*multiplication*); we shall denote the set of all types by Tp . Capital letters (A, B, \dots) range over types. Capital Greek letters (except Σ) range over finite (possibly empty) sequences of types. Expressions of the form $\Gamma \rightarrow C$ are called *sequents* of L^* .

Axioms: $A \rightarrow A$.

Rules:

$$\begin{array}{c} \frac{A\Pi \rightarrow B}{\Pi \rightarrow A \backslash B} (\rightarrow \backslash) \qquad \frac{\Pi \rightarrow A \quad \Gamma B \Delta \rightarrow C}{\Gamma \Pi(A \backslash B) \Delta \rightarrow C} (\backslash \rightarrow) \\ \frac{\Pi A \rightarrow B}{\Pi \rightarrow B / A} (\rightarrow /) \qquad \frac{\Pi \rightarrow A \quad \Gamma B \Delta \rightarrow C}{\Gamma(B / A) \Pi \Delta \rightarrow C} (/ \rightarrow) \\ \frac{\Pi \rightarrow A \quad \Delta \rightarrow B}{\Pi \Delta \rightarrow A \cdot B} (\rightarrow \cdot) \qquad \frac{\Gamma A B \Delta \rightarrow C}{\Gamma(A \cdot B) \Delta \rightarrow C} (\cdot \rightarrow) \\ \frac{\Pi \rightarrow A \quad \Gamma A \Delta \rightarrow C}{\Gamma \Pi \Delta \rightarrow C} (\text{cut}) \end{array}$$

Now let Σ be an alphabet (an arbitrary nonempty set, finite or countable). By Σ^* we denote the set of all words over Σ . The set Σ^* with the operation of word concatenation is the *free monoid* generated by Σ ; the empty word ϵ is the unit. Subsets of Σ^* are called *languages* over Σ . The three connectives of L^* correspond to three natural operations on languages ($M, N \subseteq \Sigma^*$): $M \cdot$

$N \Leftrightarrow \{uv \mid u \in M, v \in N\}$, $M \setminus N \Leftrightarrow \{u \in \Sigma^* \mid (\forall v \in M) vu \in N\}$, and $N / M \Leftrightarrow \{u \in \Sigma^* \mid (\forall v \in M) uv \in N\}$ (“ \Leftrightarrow ” here and further means “equals by definition”).

An *L-model* is a pair $\mathcal{M} = \langle \Sigma, w \rangle$, where Σ is an alphabet and w is a function that maps Lambek calculus types to languages over Σ , such that $w(A \cdot B) = w(A) \cdot w(B)$, $w(A \setminus B) = w(A) \setminus w(B)$, and $w(B / A) = w(B) / w(A)$ for all $A, B \in \text{Tp}$. Obviously, w can be defined on primitive types in an arbitrary way, and then it is uniquely propagated to all types.

A sequent of the form $F \rightarrow G$ is considered *true* in a model \mathcal{M} ($\mathcal{M} \models F \rightarrow G$) if $w(F) \subseteq w(G)$. L-models give sound and complete semantics for L, due to the following theorem:

Theorem 1. *A sequent $F \rightarrow G$ is provable in L^* if and only if it is true in all L-models.*

This theorem is proved in [7]; its special case for the product-free fragment (where we keep only types without multiplication) is much easier and appears in [2]. (The notion of truth in an L-model and this theorem can be easily generalized to sequents with zero or more than one type on the left, since $L^* \vdash F_1 F_2 \dots F_n \rightarrow G$ if and only if $L^* \vdash F_1 \cdot F_2 \cdot \dots \cdot F_n \rightarrow G$; for a sequent with an empty antecedent $\rightarrow G$ its derivability is equivalent to the fact that $\epsilon \in w(G)$ for any L-model $\mathcal{M} = \langle \Sigma, w \rangle$).

2 The Lambek Calculus with the Unit

Lambek [5] has also introduced an extension of L with the unit constant $\mathbf{1}$. We denote the new set of types by Tp_1 , and the new calculus L_1 is obtained from L* by adding a new axiom $\rightarrow \mathbf{1}$ and a new rule

$$\frac{\Gamma \Delta \rightarrow C}{\Gamma \mathbf{1} \Delta \rightarrow C} \quad (\mathbf{1} \rightarrow).$$

The set of all languages over an alphabet Σ forms a monoid with respect to the multiplication operation, and the unit of this monoid is $\{\epsilon\}$. It would be natural to define $w(\mathbf{1}) = \{\epsilon\}$, but (see, for example, [6]) this leads to incompleteness: for example, the sequent $(\mathbf{1} / p)(\mathbf{1} / p) \rightarrow \mathbf{1} / p$ is true in any L-model, but is not derivable in L_1 .

There are two possible ways to deal with this situation. The first way is to modify the notion of L-model (change the interpretation of the unit and restrict the interpretations of primitive types) — this is done in [10]. The second option, which we consider more interesting, is to extend the logic to gain completeness with respect to the existing class of L-models, with $w(\mathbf{1}) = \{\epsilon\}$. This remains an open question; here we shall present a fragment of this logic.

For any type B consider the type \tilde{B} , which is B where all occurrences of any primitive type p_i are replaced by $p_i \setminus \mathbf{1}$ (one can symmetrically consider $\mathbf{1} / p_i$ or even mix them). Let $\tilde{\text{Tp}}_1 \Leftrightarrow \{\tilde{B} \mid B \in \text{Tp}_1\}$.

Consider a three-element set, say, $\mathbb{T} \Leftrightarrow \{0, 1, \bar{1}\}$ and two functions $\&, \Rightarrow: \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$ (written in the infix form), defined by the following table:

x	y	$x \& y$	$x \Rightarrow y$
0	0	0	1
0	1	0	1
1	0	0	0
1	1	1	1
0	$\bar{1}$	0	1
$\bar{1}$	0	0	0
1	$\bar{1}$	1	0
$\bar{1}$	1	1	1
$\bar{1}$	$\bar{1}$	$\bar{1}$	$\bar{1}$

Table 1

The first four rows here just represent truth tables for classical conjunction and implication; the other rows extend these operations to their three-valued versions: \Rightarrow is the implication of Sobociński's [9] logic RM_3 (a Gentzen-style calculus for it can be found in [1]), $\&$ is the contentability (fusion) operation. We shall say that a value $x \in \mathbb{T}$ is "true" if $x = 1$ or $x = \bar{1}$, and "false" if $x = 0$. We also introduce an ordering on \mathbb{T} : $0 < \bar{1} < 1$.

Now we define the notion of *trivalent interpretation* for types from Tp_1 , $\tau: \text{Tp}_1 \rightarrow \mathbb{T}$. The function τ is defined in an arbitrary way on primitive types and satisfies the following conditions: $\tau(\mathbf{1}) = \bar{1}$, $\tau(A \cdot B) = \tau(A) \& \tau(B)$, and $\tau(A \setminus B) = \tau(B / A) = (\tau(A) \Rightarrow \tau(B))$.

Theorem 2. *If $F, G \in \widetilde{\text{Tp}}_1$, then the following statements hold:*

1. $\mathcal{M} \models F \rightarrow G$ for any L-model \mathcal{M} if and only if $\tau(F) \leq \tau(G)$ for any trivalent interpretation τ ;
2. $\mathcal{M} \models \rightarrow G$ for any L-model \mathcal{M} if and only if the value $\tau(G)$ is true for any trivalent interpretation τ .

As a corollary we get the result that the set $\mathcal{T} = \{F_1 \dots F_m \rightarrow G \mid F_1, \dots, F_m, G \in \widetilde{\text{Tp}}_1, \text{ and the sequent } F_1 \dots F_m \rightarrow G \text{ is true in every L-model}\}$ is decidable. More precisely, it is co-NP-complete (trivalent interpretations include Boolean conjunction and implication — a complete system — therefore \mathcal{T} is co-NP-hard due to NP-hardness of SAT; trivially, \mathcal{T} itself is in the co-NP class). Note that another fragment, the Lambek calculus itself without the unit, is, vice versa, NP-complete [8].

Proof (of Theorem 2). It easily follows from the definitions that for any $M \subseteq \Sigma^*$

$$M \setminus \{\epsilon\} = \{\epsilon\} / M = \begin{cases} \Sigma^*, & \text{if } M = \emptyset; \\ \{\epsilon\}, & \text{if } M = \{\epsilon\}; \\ \emptyset & \text{in other cases.} \end{cases}$$

We also consider the results of operations \cdot , \setminus and $/$ on languages \emptyset , $\{\epsilon\}$, and Σ^* :

M	N	$M \cdot N$	$M \setminus N$	N / M
\emptyset	\emptyset	\emptyset	Σ^*	Σ^*
\emptyset	Σ^*	\emptyset	Σ^*	Σ^*
Σ^*	\emptyset	\emptyset	\emptyset	\emptyset
Σ^*	Σ^*	Σ^*	Σ^*	Σ^*
\emptyset	$\{\epsilon\}$	\emptyset	Σ^*	Σ^*
$\{\epsilon\}$	\emptyset	\emptyset	\emptyset	\emptyset
Σ^*	$\{\epsilon\}$	Σ^*	\emptyset	\emptyset
$\{\epsilon\}$	Σ^*	Σ^*	Σ^*	Σ^*
$\{\epsilon\}$	$\{\epsilon\}$	$\{\epsilon\}$	$\{\epsilon\}$	$\{\epsilon\}$

Table 2

Statement 1 of the theorem follows from statement 2, because $\mathcal{M} \models F \rightarrow G$ iff $\mathcal{M} \models \rightarrow F \setminus G$ and, on the other hand, $\tau(F) \leq \tau(G)$ iff the value $\tau(F \setminus G) = (\tau(F) \Rightarrow \tau(G))$ is true (see the table for \Rightarrow above).

Let us prove statement 2. First we prove the implication from right to left. Let \mathcal{M} be an L-model and let G be a type from $\widehat{\text{Tp}}_1$ such that the value $\tau(G)$ is true for any trivalent interpretation τ . We shall prove that $\mathcal{M} \models \rightarrow G$. Let $\mathcal{M} = \langle \Sigma, w \rangle$, and let

$$\tau(p_i) = \begin{cases} 0, & \text{if } w(p_i) = \emptyset; \\ \bar{1}, & \text{if } w(p_i) = \{\epsilon\}; \\ 1 & \text{in other cases.} \end{cases}$$

Now, for any non-primitive subtype G' of G , since all occurrences of p_i are in subtypes of the form $p_i \setminus \mathbf{1}$, we have the following correspondence between $w(G')$ and $\tau(G')$: $w(G') = \emptyset$ iff $\tau(G') = 0$; $w(G') = \{\epsilon\}$ iff $\tau(G') = \bar{1}$; $w(G') = \Sigma^*$ iff $\tau(G') = 1$ (other values of $w(G')$ are impossible). For primitive types this is true by definition, and then we proceed by induction on G' (compare Table 1 and Table 2). Now, for G we have $\tau(G) = 1$ or $\tau(G) = \bar{1}$, therefore $w(G) = \Sigma^*$ or $w(G) = \{\epsilon\}$, and in both cases $\epsilon \in w(G)$, so $\mathcal{M} \models \rightarrow G$.

The other direction is similar. Let τ be a trivalent interpretation. Consider, say, $\Sigma = \{a, b\}$, and let

$$w(p_i) = \begin{cases} \emptyset, & \text{if } \tau(p_i) = 0; \\ \{\epsilon\}, & \text{if } \tau(p_i) = \bar{1}; \\ \Sigma^*, & \text{if } \tau(p_i) = 1. \end{cases}$$

Now, since $\epsilon \in w(G)$, $\tau(G)$ is either 1 or $\bar{1}$.

We see that the considered fragment of the theory of L-models with the unit is actually a commutative trivalent logic.

3 The Lambek Calculus with the Empty Set Constant

In this section we consider L-models with a special constant $\mathbf{0}$, interpreted as \emptyset . The set of types here is denoted by $\text{Tp}_{\mathbf{0}}$. We introduce a new set $\mathbb{T} = \{0, 1, \mathcal{L}\}$, and two new operations $\&, \Rightarrow: \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$, defined as follows (this is the implication-conjunction fragment of the strong Kleene [3, § 64] 3-valued logic):

x	y	$x \& y$	$x \Rightarrow y$
0	0	0	1
0	1	0	1
1	0	0	0
1	1	1	1
0	\mathcal{L}	0	1
\mathcal{L}	0	0	\mathcal{L}
1	\mathcal{L}	\mathcal{L}	\mathcal{L}
\mathcal{L}	1	\mathcal{L}	1
\mathcal{L}	\mathcal{L}	\mathcal{L}	\mathcal{L}

Table 3

The trivalent interpretation τ is defined as in the previous section; $\tau(\mathbf{0}) = 0$.

Now we introduce a hybrid logic $\text{L}_{\mathbf{0}}$: we say that $\text{L}_{\mathbf{0}} \vdash \rightarrow G$ ($G \in \text{Tp}_{\mathbf{0}}$) if either $\tau(G) = 1$ for any trivalent interpretation τ , or for any trivalent interpretation τ we have $\tau(G) \in \{1, \mathcal{L}\}$, and $\text{L} \vdash \rightarrow G$ (in L-derivations $\mathbf{0}$ is considered an ordinary primitive type). We say that $\text{L}_{\mathbf{0}} \vdash F_1 \dots F_m \rightarrow G$ iff $\text{L}_{\mathbf{0}} \vdash \rightarrow (F_1 \cdot \dots \cdot F_m) \setminus G$.

Intuitively, the third truth value \mathcal{L} here means “I don’t know, ask the Lambek calculus”.

Theorem 3. *If $\text{L}_{\mathbf{0}} \vdash F_1 \dots F_m \rightarrow G$, then the sequent $F_1 \dots F_m \rightarrow G$ is true in all L-models.*

Proof. It is sufficient to consider only sequents with empty antecedents. If $\text{L} \vdash \rightarrow G$ (the second case in the definition of $\text{L}_{\mathbf{0}}$, then $\rightarrow G$ is true in all L-models due to Theorem 1.

Now let $\tau(G) = 1$ for any trivalent interpretation τ . Let $\mathcal{M} = \langle \Sigma, w \rangle$ be an L-model. For any p_i let $\tau(p_i) = \mathcal{L}$. One can prove by induction that for any G' if $\tau(G') = 0$, then $w(G') = \emptyset$, and if $\tau(G') = 1$, then $w(G') = \Sigma^*$. Therefore, $w(G) = \Sigma^* \ni \epsilon$.

This is a correctness theorem. Unfortunately, L_0 is incomplete: $L_0 \not\vdash p_2 \cdot ((\mathbf{0} \setminus \mathbf{0}) \setminus p_1) \rightarrow ((\mathbf{0} \setminus \mathbf{0}) \setminus p_1)$, but this sequent is true in all L-models (due to the fact that for any languages N and M we have $N \cdot (\Sigma^* \setminus M) \subseteq (\Sigma^* \setminus M)$).

Acknowledgments I am grateful to Prof. Mati Pentus for fruitful discussions and constant attention to my work. I am also grateful to the anonymous referees for their helpful comments and suggestions.

This research was supported by the Russian Foundation for Basic Research (grants 11-01-00281-a and 12-01-00888-a), by the Presidential Council for Support of Leading Research Schools (grant NSh-5593.2012.1), and by the Scientific and Technological Cooperation Programme Switzerland–Russia (STCP-CH-RU, project “Computational Proof Theory”).

References

1. Avron, A.: Natural 3-valued logics — characterization and proof theory. *Journal of Symbolic Logic*, Vol. 56, No. 1, 276–294 (1991)
2. Buszkowski, W.: Compatibility of categorial grammar with an associated category system. *Zeitschr. für Math. Logik und Grundle. der Math.* 28, 229–238 (1982)
3. Kleene, S. C.: *Introduction to metamathematics*. D. van Nostrand Company, New York — Toronto (1952)
4. Lambek, J.: The mathematics of sentence structure. *Amer. Math. Monthly*, Vol. 65, No. 3, 154–170 (1958)
5. Lambek, J.: Deductive systems and categories II. Standard constructions and closed categories. *Category Theory, Homology Theory and Their Applications I*, ed. by P. Hilton (Proc. of the Conference held at the Seattle Research Center of the Battelle Memorial Institute, June 24–July 19, 1968; Springer Lect. Notes Math., vol. 86), 76–122. Springer, Berlin (1969)
6. Morrill, G.: *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford University Press, Oxford (2011)
7. Pentus, M.: Free monoid completeness of the Lambek calculus allowing empty premises. *Proc. Logic Colloquium '96*, ed. by J. M. Larrazabal, D. Lascar, G. Mints (Lect. Notes Logic; vol. 12), 171–209. Springer, Berlin (1998)
8. Pentus, M.: Lambek calculus is NP-complete. *Theoretical Comput. Sci.*, Vol. 357, No. 1–3, 186–201 (2006)
9. Sobociński, B.: Axiomatization of a partial system of three-value calculus of propositions. *Journal of Computing Systems*, Vol. 1, 23–55 (1952)
10. Zvonkin, M. M.: Language models for the Lambek calculus with the unit (in Russian). Term paper, Dept. of Math. Logic and Theory of Algorithms, Moscow State University (2011)

A trivalent logic that encapsulates intuitionistic and classical logic

Tin Perkov

Polytechnic of Zagreb, Croatia
tin.perkov@tvz.hr

Abstract. A third truth-value is proposed with purpose to distinguish intuitionistic valid formulas within classical validities.

Keywords: trivalent logic, intuitionistic logic, Kripke semantics

1 Introduction

In the classical semantics for propositional logic, we interpret propositional variables by assigning a truth-value 0 or 1 to each of them. Then we extend this assignment to all well-formed formulas in a familiar way, by truth tables. If we introduce the third truth-value, say $1/2$, we need to include it in truth tables, and this has been done in various ways, resulting in various systems of trivalent logic, like Łukasiewicz [8], Kleene [6], Gödel logics [4] etc. On the other hand, even with only two truth values we can get diversity by considering semantics other than truth functional, e. g. Kripke semantics. In this way we get intuitionistic logic or, if we also enrich syntax by adding operators, modal logics and so on.

In this paper, a simple example of combining three-valued logic with Kripke semantics is presented, in which classical propositional tautologies and intuitionistic validities are characterized by a correspondence with truth values. Truth value 0 is read as *false*, 1 as *strongly true*, and new truth value $1/2$ is read *weakly true*. From the way in which semantics will be defined, it will follow that tautologies correspond to non-falsity (truth value 1 or $1/2$), while intuitionistic validities correspond to the truth value 1.

In Section 3 the same idea is applied to the first-order logic, using Kripke semantics for the intuitionistic predicate logic.

2 Propositional logic

We work with the basic countable propositional vocabulary. Fix a countable set $\text{PROP} = \{p_1, p_2, \dots\}$. Elements of PROP are called *propositional variables*. Sometimes we denote them by p, q , and so on, without indices, namely in examples in which only few of them occur. We denote by FORM a set of well-formed

formulas, the smallest set that contains $\text{PROP} \cup \{\perp\}$ and is closed under conjunctions, disjunctions and conditionals. That is, the syntax is given by

$$F ::= p \mid \perp \mid F_1 \wedge F_2 \mid F_1 \vee F_2 \mid F_1 \rightarrow F_2,$$

where $p \in \text{PROP}$. We abbreviate $\neg F := F \rightarrow \perp$.

Consider a *Kripke frame* $\mathfrak{F} = (W, \leq)$, where W is a non-empty set, and \leq a partial ordering on W called *accessibility relation*. A *local assignment* of truth values is a family $l = \{l_w : w \in W\}$ of functions $l_w : \text{FORM} \rightarrow \{0, 1/2, 1\}$ such that for all $w \in W$ we have:

1. if $l_w(p) = 1$ and $w \leq v$, then $l_v(p) = 1$, for all $p \in \text{PROP}$,
2. $l_w(\perp) = 0$,
3. $l_w(F_1 \wedge F_2) = \begin{cases} 1, l_w(F_1) = l_w(F_2) = 1, \\ 0, l_w(F_1) = 0 \text{ or } l_w(F_2) = 0, \\ 1/2, \text{ otherwise,} \end{cases}$
4. $l_w(F_1 \vee F_2) = \begin{cases} 1, l_w(F_1) = 1 \text{ or } l_w(F_2) = 1, \\ 0, l_w(F_1) = l_w(F_2) = 0, \\ 1/2, \text{ otherwise,} \end{cases}$
5. $l_w(F_1 \rightarrow F_2) = \begin{cases} 1, \text{ for all } v \geq w, \text{ if } l_v(F_1) = 1 \text{ then } l_v(F_2) = 1, \\ 1/2, \text{ if } l_w(F_1) \neq 0 \text{ then } l_w(F_2) \neq 0, \text{ and} \\ \text{there is } v \geq w \text{ such that } l_v(F_1) = 1 \text{ and } l_v(F_2) \neq 1, \\ 0, \text{ otherwise.} \end{cases}$

From 2. and 5. we conclude:

$$l_w(\neg F) = \begin{cases} 1, \text{ for all } v \geq w, l_v(F) \neq 1, \\ 1/2, \text{ there is } v \geq w \text{ such that } l_v(F) = 1, \text{ and } l_w(F) = 0, \\ 0, \text{ there is } v \geq w \text{ such that } l_v(F) = 1, \text{ and } l_w(F) \neq 0. \end{cases}$$

Furthermore, 1. is easily extended to all formulas by induction.

We call $\mathfrak{M} = (W, \leq, l)$ a *model*. Clearly, \mathfrak{M} can be viewed as a Kripke model for intuitionistic logic, by putting $w \Vdash F$ if and only if $l_w(F) = 1$, since all clauses considering truth value 1 are exactly the same as in the definition of Kripke semantics for intuitionistic logic (see [7]). From this viewpoint, 1 is a designated truth value, while 0 and 1/2 are identified as both false.

Let $F \in \text{FORM}$. We say that F is *strongly valid* if $l_w(F) = 1$ for any choice of model $\mathfrak{M} = (W, \leq, l)$, and any $w \in W$. We say that F is *weakly valid* if $l_w(F) \neq 0$, for any \mathfrak{M} and w . The following fact immediately follows from the previous remarks.

Proposition 1. *Let $F \in \text{FORM}$. Then F is strongly valid if and only if it is an intuitionistic validity.*

The truth value 1 corresponds with validities of propositional intuitionistic logic both on the local and global level, since the semantics is intentionally defined in a way that completely resembles Kripke semantics for intuitionistic logic. On the other hand, we do not have local correspondence of truth values 1 and 1/2 with classical propositional truth, namely since the statements involving truth value 0 clearly do not cover all cases of classical falsity.

Example 1. Let $\mathfrak{F} = (\{w\}, \{(w, w)\})$. Let p and q be two propositional variables and l a local assignment such that $l_w(p) = 1/2$ and $l_w(q) = 0$. If we consider $1/2$ a designated value, then $p \rightarrow q$ is false at w in the classical sense, but $l_w(p \rightarrow q) = 1$.

But in terms of validity, when we abstract away from a choice of a model, we have the result analogous to the previous proposition.

Proposition 2. *Let $F \in \text{FORM}$. Then F is weakly valid if and only if it is a propositional tautology.*

Proof. Assume F is not a tautology. Then there exists a (classical) assignment $a : \text{FORM} \rightarrow \{0, 1\}$ such that $a(F) = 0$. Now let $\mathfrak{F} = (\{w\}, \{(w, w)\})$. Define a local assignment l by putting $l_w(p) = a(p)$ for all $p \in \text{PROP}$. This satisfies the clause 1. of the definition of local assignment, so l is well defined. It is easy to show by induction that $l_w(A) = a(A)$ for all $A \in \text{FORM}$. Thus $l_w(F) = 0$. Hence by contrapositive, a weakly valid formula is a tautology.

The reverse implication follows immediately from the definition of local assignment. \square

3 Predicate logic

Let σ be a first-order vocabulary. For the sake of notational simplicity, assume σ has no function or constant symbols, so the only terms are individual variables, which we denote by x, y, z , with or without indices. We denote the set of variables, which is countable, by VAR.

So, σ has finitely or countably many relation symbols which we denote by R, R_1, R_2 and so on. Each relation symbol has a given arity. An *atomic formula* is $Rx_1x_2 \dots x_n$, where R is an n -ary relation symbol, and x_1, \dots, x_n variables. We denote the set of atomic formulas by AT. By FO we denote the set of all well-formed formulas over σ , which is the smallest set containing $\text{AT} \cup \{\perp\}$ that is closed under conjunctions, disjunctions, conditionals and quantifiers. In other words, the syntax is

$$F ::= A | \perp | F_1 \wedge F_2 | F_1 \vee F_2 | F_1 \rightarrow F_2 | \forall x F | \exists x F,$$

where $A \in \text{AT}$. As before, we abbreviate $\neg F := F \rightarrow \perp$.

A *model* based on a Kripke frame $\mathfrak{F} = (W, \leq)$ is $\mathfrak{M} = (W, \leq, M)$, where M is a family $\{M_w : w \in W\}$ of non-empty sets, together with three-valued interpretations of relation symbols, i. e. functions $R_w : M_w^n \rightarrow \{0, 1/2, 1\}$ for each $R \in \sigma$, where n is the arity of R , such that we have:

1. if $w \leq v$, then $M_w \subseteq M_v$,
2. if $w \leq v$ and $R_w(c_1, \dots, c_n) = 1$, then $R_v(c_1, \dots, c_n) = 1$.

Assigning truth values to the first-order formulas at $w \in W$ depends on a valuation of variables by elements of M_w . Note that, if $w \leq v$, since we have

$M_w \subseteq M_v$, a is also a valuation of variables by elements of M_v . We denote by $a(c/x)$ a valuation that assigns c to x , and agrees with a otherwise. Given a valuation $a : \text{VAR} \rightarrow M_w$, an *assignment* at w is a function $l_w^a : \text{FO} \rightarrow \{0, 1/2, 1\}$ defined as follows:

1. $l_w^a(Rx_1 \dots x_n) = R_w(a(x_1), \dots, a(x_n))$ for all $Rx_1 \dots x_n \in \text{AT}$,
2. $l_w^a(\perp) = 0$,
3. $l_w^a(F_1 \wedge F_2) = \begin{cases} 1, l_w^a(F_1) = l_w^a(F_2) = 1, \\ 0, l_w^a(F_1) = 0 \text{ or } l_w^a(F_2) = 0, \\ 1/2, \text{ otherwise,} \end{cases}$
4. $l_w^a(F_1 \vee F_2) = \begin{cases} 1, l_w^a(F_1) = 1 \text{ or } l_w^a(F_2) = 1, \\ 0, l_w^a(F_1) = l_w^a(F_2) = 0, \\ 1/2, \text{ otherwise,} \end{cases}$
5. $l_w^a(F_1 \rightarrow F_2) = \begin{cases} 1, \text{ for all } v \geq w, \text{ if } l_v^a(F_1) = 1 \text{ then } l_v^a(F_2) = 1, \\ 1/2, \text{ if } l_w^a(F_1) \neq 0 \text{ then } l_w^a(F_2) \neq 0, \text{ and} \\ \text{there is } v \geq w \text{ such that } l_v^a(F_1) = 1 \text{ and } l_v^a(F_2) \neq 1, \\ 0, \text{ otherwise,} \end{cases}$
6. $l_w^a(\exists x F) = \begin{cases} 1, \text{ there is } c \in M_w \text{ such that } l_w^{a(c/x)}(F) = 1, \\ 0, \text{ for all } c \in M_w, \text{ we have } l_w^{a(c/x)}(F) = 0, \\ 1/2, \text{ otherwise,} \end{cases}$
7. $l_w^a(\forall x F) = \begin{cases} 1, \text{ for all } v \geq w \text{ and all } c \in M_w, \text{ we have } l_v^{a(c/x)}(F) = 1, \\ 0, \text{ there is } c \in M_w \text{ such that } l_w^{a(c/x)}(F) = 0 \\ 1/2, \text{ otherwise,} \end{cases}$

To get a standard intuitionistic Kripke model (cf. [7]) from a model \mathfrak{M} , we just need to define (classical) interpretations of relation symbols for each $w \in W$. Obviously, we interpret an n -ary relation symbol R by the relation $\{(c_1, \dots, c_n) : R_w(c_1, \dots, c_n) = 1\}$. Clearly, model and assignment are defined in a way to ensure that $w \Vdash_a F$ (this denotes the intuitionistic satisfaction relation under a valuation a) in thus defined Kripke model if and only if $l_w^a(F) = 1$.

We say that a formula $F \in \text{FO}$ is *strongly valid* if $l_w^a(F) = 1$ for any model $\mathfrak{M} = (W, \leq, M)$, $w \in W$ and $a : \text{VAR} \rightarrow M_w$. We say that F is *weakly valid* if $l_w(F) \neq 0$, for any \mathfrak{M} , w and a . An analogue of Proposition 1 immediately follows.

Proposition 3. *Let $F \in \text{FO}$. Then F is strongly valid if and only if it is an intuitionistic validity.*

Similarly as in the case of propositional formulas, we also establish a correspondence between weak validity and classical validity.

Proposition 4. *Let $F \in \text{FO}$. Then F is weakly valid if and only if it is a valid formula of predicate calculus.*

Proof. Let F be a weakly valid formula. Assume that F is not a classical validity. Then there exists a σ -structure \mathfrak{M} and an assignment $a : \text{VAR} \rightarrow M$, where M

is the domain of \mathfrak{M} , such that $\mathfrak{M} \not\models_a F$. Let $\mathfrak{F} = (\{w\}, \{(w, w)\})$. To define a model based on \mathfrak{F} , we only need to define M_w and a three-valued interpretation of relation symbols on M_w . Put $M_w = M$ and for any $R \in \sigma$, with the arity n , define $R_w(c_1, \dots, c_n) = 1$ if $(c_1, \dots, c_n) \in R^{\mathfrak{M}}$, where $R^{\mathfrak{M}}$ denotes the relation that interprets R in \mathfrak{M} , and $R_w(c_1, \dots, c_n) = 0$ otherwise. Since \mathfrak{F} is singleton, this is a well-defined model.

Now, by induction we easily see that under the assignment a we have $\mathfrak{M} \models_a A$ if and only if $l_w^a(A) = 1$, and $\mathfrak{M} \not\models_a A$ if and only if $l_w^a(A) = 0$, for any $A \in \text{FO}$. Thus $l_w^a(F) = 0$, hence F is not weakly valid, which contradicts the assumption. So, F is in fact a classical validity.

For the converse, let $F \in \text{FO}$ be a classical validity. It is easy to see from the definition of l_w^a that we can not define a model and an assignment such that $l_w^a(F) = 0$, which proves the claim. \square

4 Concluding remarks

It should be noted that Baaz and Fermüller [1] systematically combine many-valued and intuitionistic logic, namely tableau calculi for both, to obtain intuitionistic many-valued logics. They obtain soundness and completeness results for these logics with respect to intuitionistic Kripke semantics generalized to many-valued interpretation in a somewhat similar way as in this paper. While the notion of Kripke model is essentially the same, although more general than the one presented here, semantics of connectives however differs, due to different intentions – authors of [1] develop intuitionistic semantics generalized to the many-valued context, without considering classical validities. In propositional case this has already been done by Rousseau [10], who developed sequent calculus for many-valued intuitionistic logic, which was further improved by Reznik and Curmin [9].

Furthermore, it may be worthwhile to explore applying similar idea to modal logic. Indeed, the Łukasiewicz's original idea behind introducing three-valued logic was to reason about possibility. But it has been shown that standard modal logic is not definable by any finitely many-valued logic with any truth table semantics (see [3]). However, Béziau [2] considers some non-standard modal logics based on four-valued truth tables. On the other hand, one might ask if there is a Kripke-like semantics for three (or more) valued logic that would capture modality, without using modal operators in a language as symbols. Consider the basic propositional modal language with one modal operator \Box . It should be clear that by slightly adapting the definition of local assignment from Section 2, we can characterize the truth of a formula of the form $\Box F$, where $F \in \text{FORM}$, by the strong truth of F . It is not momentarily clear if we can characterize the truth of more complex modal formulas, e. g. nested boxes, by generalizing this approach somehow, but it might be worthy to explore this further.

Also, Wójcicki [11, 12] uses similar approach to define Kripke semantics for Nelson logic, that is, an extension of propositional intuitionistic logic by a connective \sim called *strong negation*. Analogous semantics for predicate Nelson logic

is considered by Hasuo and Kashima [5]. Roughly, semantics of negation \neg corresponds to the strong truth of negation, while semantics of \sim corresponds to weak truth of negation as presented in this paper. We leave for further consideration a question if Nelson logic can be precisely captured by multi-valued approach without using the strong negation as a symbol of a language.

References

1. M. Baaz, C. G. Fermüller: Combining many-valued and intuitionistic tableaux. In: P. Miglioli et. al. (eds.) *Theorem Proving with Analytic Tableaux and Related Methods*, pp. 65–79. Springer (1996)
2. J. Y. Béziau: A new four-valued approach to modal logic. *Logique et Analyse*, 54, 109–121 (2011)
3. J. Dugundji: Note on a property of matrices for Lewis and Langford’s calculi of propositions. *Journal of Symbolic Logic*, 5, 150–151 (1940)
4. K. Gödel: Zum intuitionistischen Aussagenkalkül. *Anzeiger der Akademie der Wissenschaften in Wien*, 69, 65–66 (1932), English translation in: K. Gödel: *Collected Works Vol. I*, S. Feferman et al. (eds.), Oxford: Oxford University Press (1986)
5. I. Hasuo, R. Kashima: Kripke completeness of first-order constructive logics with strong negation. *Logic Journal of the IGPL* 11(6), 615–646 (2003)
6. S. C. Kleene: On notation for ordinal numbers. *Journal of Symbolic Logic*, 3, 150–155 (1938)
7. S. A. Kripke: Semantical analysis of intuitionistic logic. In: J. Crossley and M. A. E. Dummett (eds.) *Formal Systems and Recursive Functions*, pp. 92–130. Amsterdam: North-Holland Publishing (1965)
8. J. Łukasiewicz: O logice trójwartościowej, *Ruch Filozoficzny*, 5, 170–171 (1920), English translation in: J. Łukasiewicz: *Selected Works*, L. Borkowski (ed.), Amsterdam: North-Holland and Warsaw: PWN. (1970)
9. E. Reznik, P. Kurmin: Intuitionistic sequent calculi for finitely many-valued logics. *Logic Journal of the IGPL* 9(6), 793–812 (2001)
10. G. Rousseau: Sequents in many valued logic II. *Fundamenta Mathematicae*, 67, 125–131 (1970)
11. R. Wójcicki: *Lectures on Propositional Calculi*. Dordrecht: Kluwer Academic Publishers (1988)
12. R. Wójcicki: *Theory of Logical Calculi*. Wrocław: Ossolineum (1984)

TCS for presuppositions

Jérémy Zehr¹ and Orin Percus²

¹ Institut Jean Nicod, UMR 8128,
École Normale Supérieure, France
<http://www.institutnicod.org/>

² LLING, EA 3827,
Université de Nantes, France
<http://lling.univ-nantes.fr/>

Abstract. For several decades now, there have been attempts to account for vagueness and presuppositions in a uniform manner. These attempts are motivated by the fact that the two phenomena have something in common: the lack of a solid truth-value judgment that we find when a presuppositional statement is used in a presupposition failure situation, and when a vague predicate when used to describe a borderline case. In this paper, we extend this enterprise by considering the logical system named TCS³ and developed by Cobreros&al. [3] to account for vagueness. Cobreros&al. introduce two notions of truth — tolerant truth and strict truth — that allow propositions to be characterized as both true and false, or as neither true nor false. We explore to what degree this system can be applied to treat presuppositional phenomena. In particular, we use the notion of strict truth to derive the presupposition associated with a statement; this way of deriving presuppositions extends naturally to complex sentences and thus leads to an account for the phenomenon of presupposition projection.

Keywords: vagueness, presuppositions, presupposition projection, tolerant truth, strict truth, trivalence, logic, semantics

1 The TCS System

Cobreros&al. [3] build their system on the basis of a bivalent logic. They start from a language in which predicates may combine with every constant of the language, and where they are associated with functions that return 1 (truth) or 0. Adapting their terminology⁴, we can speak of these functions as representing the *classical* meaning of a predicate. The classical meanings of connectives — the standard ones — insure that, when we speak classically, all the principles of

³ TCS stands for *Tolerant Classical Strict*.

⁴ We will speak of classical (or strict or tolerant) *meaning*, even if strictly speaking Cobreros&al.'s definitions are in terms of classical (or strict or tolerant) *truth*. When we speak of the classical (or strict or tolerant) *meaning* of a predicate P , what we mean is a function that, given the value of a constant a , yields the classical (or strict or tolerant) value of Pa .

classical logic are preserved. In particular, the classical meaning of negation is such that a proposition's negation is true iff the proposition itself is not true, so we never find that both a proposition and its negation are true.

Importantly, however, classical meaning is not the only type of meaning an expression can have: there are also two other types of meaning, *strict* meaning and *tolerant* meaning. Depending on how we want to assert a statement, we can associate it with its classical, strict or tolerant meaning. Usually we speak classically, but if we want to strengthen our discourse, we will associate our statements with their strict meanings; and if we want to weaken our discourse, we will associate our statements with their tolerant meanings. We give below definitions for strict and tolerant truth and falsity, following the general lines of Cobreros&al. Note that the definitions in the case of falsity correspond to definitions of negation — to say that a proposition is tolerantly, classically or strictly false is to say that its negation is tolerantly, classically or strictly true, respectively.

Definition 1 (Strict truth and falsity of a vague predicate P).

*P(a) is strictly true for any a if and only if P is classically true of all individuals that are sufficiently like a with respect to the measure associated with P*⁵.

P(a) is strictly false for any a if and only if P is classically false of all individuals that are sufficiently like a with respect to the measure associated with P.

Definition 2 (Duality in TCS).

A proposition is tolerantly true if and only if it is not strictly false.

A proposition is tolerantly false if and only if it is not strictly true.

Definition 3 (Connectives in TCS).

Conjunctions of the form A & B are strictly true if and only if a is strictly true and b is strictly true.

Conjunctions of the form A & B are tolerantly true if and only if a is tolerantly true and b is tolerantly true.

The contributions to strict and tolerant truth of the disjunction \vee and the conditional \rightarrow are defined in terms of negation and conjunction in the ways familiar from classical logic.

It is important to observe the dualities that we find in this system: to say that a predicate is not *strictly true* of an individual is to say that it is *tolerantly false* of her; and reciprocally, to say that a predicate is not *tolerantly true* of an individual is to say that it is *strictly false* of her. Crucially, the dual of strict truth is not strict falsity nor is the dual of tolerant truth tolerant falsity.

⁵ In reality, Cobreros&al. imagine that vague predicates are associated with a relation of indifference between individuals. They go no further than this. But, assuming that each vague predicate is associated with a specific scale of measurement (e.g. as Kennedy [12]), we can see this relation as holding between two individuals if the difference of their measures on the scale falls below a certain threshold of “distinguishability”.

Here is an example. Consider a context of only four people (let's name them a , b , c and d) who are sorted by height: a is slightly smaller than b who is slightly smaller than c who is slightly smaller than d . It's worth noting that the relation slightly smaller is not transitive: a being slightly smaller than b and b being slightly smaller than c does not entail that a is slightly smaller than c . In fact we will assume that a is not slightly smaller than c , nor is b slightly smaller than d . Assuming that the distribution of classical truth for tall is as specified in the first line of Table 1, the distribution of classical falsity would be as in the second line. Assuming that two individuals count as sufficiently alike with respect to tallness only when one is at most slightly smaller than the other, we then have the distribution of strict truth and falsity in Table 2, and accordingly the distribution of tolerant truth and falsity in Table 3.

Table 1. Classical truth and falsity of *tall* for a , b , c and d

a	b	c	d
not true	not true	true	true
false	false	not false	not false

Table 2. Strict truth and falsity of *tall* for a , b , c and d

a	b	c	d
not true	not true	not true	true
false	not false	not false	not false

Table 3. Tolerant truth and falsity of *tall* for a , b , c and d

a	b	c	d
not true	true	true	true
false	false	false	not false

Notice that, in the case of b and c , tall is on the one hand neither strictly true nor strictly false, and on the other hand both tolerantly true and tolerantly

false. This situation arises because of the way in which Cobreros&al. establish their dualities. In their system, for someone *not to be strictly tall* doesn't mean for him *to be strictly not tall*; rather, it means that he's tolerantly not tall. At the same time, when someone (like a) is not (even) tolerantly tall, this means that he's strictly not tall (and necessarily tolerantly not tall too). These remarks suggest that the Cobreros&al. semantics can be recast in terms of a sort of modal semantics with a *strictly* operator — even if on their formulation there are two special kinds of meaning, tolerant and strict, their duality allows us to formulate one in terms of the other, by playing on the relative scope of the negation and a *strictly* operator.

The aspect of Cobreros&al.'s system that interests us here is that it gives us a way of characterizing borderline cases for vague predicates — that is, cases that trigger an unclear truth-value judgment. Borderline cases are simply those for which a vague predicate yields a proposition that is neither strictly true nor strictly false — or, equivalently, cases for which the proposition is both tolerantly true and tolerantly false. Alternatively stated, a borderline case for a vague predicate is a case that leads to different strict and tolerant values for the proposition that results when we apply that predicate. Note that, the Cobreros&al. approach seems to predict that, to the extent that strict readings are natural, we should be able to judge predicates like *neither – tall – nor – not – tall* to be true of borderline cases for *tall*; likewise, to the extent that tolerant readings are natural, we should be able to judge predicates like *both – tall – and – not – tall* to be true of the same cases. In support of their approach, Cobreros&al. cite the results of an experiment by Alxatib&Pelletier [1], who found judgments of precisely this kind.

Of related interest is the fact that, to the extent that tolerant readings are natural, the system provides an account for the sorites paradox: the paradox arises due to the tolerant way of reading the inductive premiss. Consider the following statement, meant to exemplify the kind of statement that appears as a premiss in versions of the sorites paradox.

Example 1. If a man is slightly smaller than a tall man, then he is a tall man too.

Strictly speaking, (1) conveys that any man who is slightly smaller than a tolerantly tall man is a strictly tall man — and this is false. But *tolerantly speaking*, (1) conveys that any man who is slightly smaller than a strictly tall man is a tolerantly tall man — and this is true. (To see this, suppose first that (1) should be represented as a formula of the form $\forall y\forall z[[Pz\&Ryz] \rightarrow Py]$ - or equivalently as a conjunction of conditionals of the form $[[Pa\&Rab] \rightarrow Pb]$, where *Rab* expresses that *b* is slightly smaller than *a*. Then it suffices to observe that a conditional of the form $[p \rightarrow q]$ is strictly true if and only if the tolerant truth of *p* entails the strict truth of *q*, and that it is tolerantly true if and only if the strict truth of *p* entails the tolerant truth of *q*.⁶). With this in mind, consider

⁶ Here is a proof of the first claim; the proof of the second is parallel.

For brevity, we abbreviate “ ϕ is strictly true” as $S[\phi]$, “ ϕ is not strictly true” as

the tolerant reading. If a is a strictly tall man and b is slightly smaller than a , accepting the inductive premiss in a tolerant way does not lead you to conclude that b is a *strictly* tall man; rather you have to conclude that b is a *tolerantly* tall man. And if c is slightly smaller than b , then, even if this means that he's slightly smaller than a tolerantly tall man, he is just trivially concerned by the inductive premiss if b happens not to be a strictly tall man — we don't have to conclude that c is a strictly or tolerantly tall. This is the way the TCS system handles the sorites paradox. The “trick” is that the vague predicate in the antecedent of the conditional has to be understood in the dual way of the one in the consequent.

In sum, the TCS system can be seen as a system that describes strengthening and weakening of meaning. On the one hand, there is a “classical” way of using statements on which a statement is necessarily exclusively true or false. But, on the other hand, we might wish to strengthen or weaken our way of speaking, and on these modified ways of using statements, a statement can be neither true nor false, or both true and false.

2 TCS for presuppositions

2.1 Definitions

In what follows, we will explore how the TCS system might be used to treat presuppositions. More specifically, we will try to answer these two questions: i) what would we consider to be the classical, strict and tolerant meanings of presuppositional predicative objects (PPOs) ⁷ and sentences containing them?; ii) given the classical, strict and tolerant meanings for PPOs, can we derive the presuppositions associated with the various statements in which they appear, and more specifically, can this system account for presupposition projection?

Let us start with the classical meaning of a presuppositional statement. In cases where the presupposition associated with a presuppositional statement is fulfilled, there is no debate about when the sentence is true and when it is false. For instance, take a sentence like (2), associated with the presupposition that Mary used to smoke. In cases where the presupposition is fulfilled, it is uncontroversial to say that the sentence is true if Mary does not smoke now and false otherwise.

$\neg S[\phi]$ and “ ϕ is strictly false” as $S[\neg\phi]$; similarly, we use $T[\phi]$ and $T[\neg\phi]$ to talk about tolerant truth and falsity.

To start with, note that it follows from Definition 2 that, for any ϕ , $T[\neg\neg\phi]$ iff $T[\phi]$ (since $T[\neg\neg\phi]$ iff $\neg S[\neg\phi]$ and $\neg S[\neg\phi]$ iff $T[\phi]$).

Now, by Definition 3, $S[p \rightarrow q]$ iff $S[\neg[p\&\neg q]]$. But $S[\neg[p\&\neg q]]$ iff $\neg T[\neg\neg[p\&\neg q]]$ (by Definition 2) iff $\neg T[p\&\neg q]$ (by the above reasoning) iff $\neg T[p]$ or $\neg T[\neg q]$ (by Definition 3) iff $\neg T[p]$ or $S[q]$ (by Definition 2).

⁷ We adopt the terminology of Charlow [4]. PPOs are predicates such as *quit/stop X* that are known to trigger the presupposition that the individual they are asserted of used to X or predicates such as *know that Y* that are known to trigger the presupposition that Y . The article “Presupposition” of the Stanford Encyclopedia [2] includes a list of lexical classes that can be seen as corresponding to PPOs.

Example 2. Mary has stopped smoking

At the same time, in cases where the presupposition is not fulfilled, there has been a controversy — ever since Russell [13] argued, contra an interpretation of Frege [9], that a presuppositional statement should be considered false when the presupposition is not fulfilled. In the spirit of Cobreros&al., we will simply imagine that the classical meaning of a presuppositional statement distinguishes those cases where the statement is uncontroversially true from those cases where it is not. We will represent a sentence like (2) as $not - smoke_{used-to-smoke}(Mary)$ and consider classical truth values to be defined as follows:

Definition 4 (Classical truth of a statement P_p). $P_p(a)$ is classically true for any a if and only if $p(a)$ and $P(a)$ are both classically true.

Definition 4 tells us that (2) is classically true if and only if Mary does not smoke but used to; otherwise it is classically false, given that classically a proposition’s negation is true iff the proposition itself is not. In other words, (2) is classically false not only if Mary used to smoke and has kept on smoking (in which case $not - smoke(Mary)$ would be classically false), but also if Mary never smoked (since in this case $used-to-smoke(Mary)$ would be classically false). In actual fact, it seems that, when people *have* to assign a truth value to a statement in the knowledge that its presupposition is not fulfilled — when they are forced to choose between “true” and “false” — they behave in a Russellian way and judge it false. This judgment reflects the classical truth value.

Now let us turn to the strict meanings for presuppositional statements. The hallmark of a presupposition is the lack of a clear truth value judgment in cases where the presupposition is known to be unfulfilled - this is the source of the controversy we alluded to above. Now, we also find lack of a clear truth judgment when borderline cases of vague predicates are concerned, and Cobreros&al. characterize those statements as statements that are neither strictly true nor strictly false. This suggests to us that we should adapt Cobreros&al.’s approach in such a way that presupposition failures come out as neither strictly true nor strictly false. We can do so as follows:

Definition 5 (Strict truth and falsity of a statement P_p ⁸).

$P_p(a)$ is strictly true for any a if and only if $p(a)$ is classically true and $P(a)$ is classically true.

$P_p(a)$ is strictly false for any a if and only if $p(a)$ is classically true and $P(a)$ is classically false.

We will go a little further than this. In the discussion until now, we have viewed sentences as literally encoding presuppositions, and treated representations like $not - smoke_{used-to-smoke}(Mary)$ as directly expressing the presupposition that Mary used to smoke. But another possible position is that the

⁸ The definition that Cobreros&al. give for the *strict* meaning of a vague predicate makes it *stronger* than its classical meaning: the truth (falsity) of a strict assertion of a vague predicate entails the truth (falsity) of its classical assertion, while the reverse is not true. Given our definitions of strict and classical meanings of a statement P_p , we observe the same relation of strength.

presupposition is derived, in the following way. Our definitions already establish that a form $P_p(a)$ is neither strictly true nor strictly false if $p(a)$ is classically false. One might take the position that there is a “bridge principle” that derives as presuppositions of a statement P those propositions whose classical falsity would prevent P from being strictly true or false. This is what we will assume⁹:

Principle 1 (Presupposition and strict truth-value). *Any proposition P is associated with the presupposition that $S[P]$ or $S[\neg P]$.*

Note that this implies that the use of a vague predicate is associated with the presupposition that we’re not talking about a borderline case; and thus that, whenever we use a vague predicate to describe a borderline case, there is a presupposition failure. We could then associate the lack of a clear truth judgment specifically with presupposition failure:

Principle 2 (Lack of truth-value judgment in cases of presupposition failure). *When confronted with a presupposition failure, a speaker will be uneasy giving a truth value judgment for the statement whose presupposition is unfulfilled.*

In sum, on our adaptation of the TCS system, simple sentences that are associated with presuppositions essentially encode a stipulation about their strict meaning. The presupposition is derived from this: the presupposition that is associated with a PPO is the condition that needs to be satisfied for the predication of this PPO to be either strictly true or strictly false. In this manner, $P_p(a)$ gives rise to the presupposition that $p(a)$ is classically true. For completeness, let us show this explicitly. (In what follows, for brevity we will write P instead of $P(a)$ and p instead of $p(a)$.) By Principle 1, P_p is associated with the presupposition that $S[P_p]$ or $S[\neg P_p]$. But by Definition 5, $S[P_p]$ iff $C[p]$ and $C[P]$, while $S[\neg P_p]$ iff $C[p]$ and $C[\neg P]$. So P_p is associated with the presupposition that $C[p]$ and either $C[P]$ or $C[\neg P]$, or in other words with the presupposition that $C[p]$ (as, for any ϕ , it is always the case that $C[\phi]$ or $C[\neg\phi]$).

Note that the condition in which a predication of a PPO is either strictly true or strictly false is the condition in which there would be no difference in the resulting classical or strict truth value for the sentence (since the strict meaning is stronger than the classical one — see note 5). In this light, qualifying an expression as “presuppositional” can be seen as anticipating its behavior in some contexts; rather, what would be intrinsically associated with a PPO would not be a presupposition anymore but an ambiguity between its classical and its strict meaning¹⁰.

⁹ As before, we write $S[\phi]$ as an abbreviation for “ ϕ is strictly true”.

¹⁰ This idea of associating a (regular) truth-value with an expression only when it wouldn’t give rise to different truth-judgments no matter how we interpret it — and observing a truth-value gap when different interpretations lead to different truth-judgments — is reminiscent of the idea of supervaluationism, which was originally proposed by van Fraassen [8] to account for presuppositions. Within this frame-

2.2 Judgments for simple predications

Our set-up, on which PPOs are associated with a particular kind of strict meaning, correctly predicts the fact that we naturally judge a predication of a PPO as neither true nor false when the associated presupposition is not fulfilled. For instance, it correctly predicts that, given the knowledge that Mary never smoked, it is natural to judge (2) in this way — this follows from the fact, that, if Mary never smoked, the sentence is neither strictly true nor strictly false. At the same time, it also potentially allows for a judgment of falsity in this kind of situation, for the sentence is classically false. And this judgment indeed seems to be possible as well. The fact that one can say of this sentence both that it is false and that it is not false is naturally accounted for given its strict and classical meanings.

A further successful prediction of the approach is that, in this same kind of situation, the negative counterpart of a sentence like (2) should naturally give rise to the same neither-true-nor-false judgment. We make this prediction to the extent that we would represent a sentence like (3) as $\neg \textit{not-smoke}_{\textit{used-to-smoke}}(\textit{Mary})$, the negation of (2). To see this, notice that to say that $S[\neg \textit{not-smoke}_{\textit{used-to-smoke}}(\textit{Mary})]$ or $S[\neg \neg \textit{not-smoke}_{\textit{used-to-smoke}}(\textit{Mary})]$ is just to say that $S[\textit{not-smoke}_{\textit{used-to-smoke}}(\textit{Mary})]$ or $S[\textit{not-smoke}_{\textit{used-to-smoke}}(\textit{Mary})]$ (due to Definition 2), so (3) is predicted to have the same presupposition as (2), in accordance with the common wisdom concerning the negative counterparts of presuppositional statements. At the same time, our approach also potentially allows for a judgment of truth in this kind of situation, for the sentence is classically true, and this judgment seems to be possible too. Parallel to the affirmative case, the fact that one can say of this sentence both that it is true and that it is not true is naturally accounted for given its strict and classical meanings.

Example 3. Mary has not stopped smoking

However, a word is in order about tolerant meanings, which we have not yet discussed in connection with presuppositional phenomena. Given Definition 2, in a situation in which Mary never smoked, (2) is *tolerantly true*, but it would certainly be quite unnatural to judge (2) true on the basis of the knowledge that Mary never smoked. The success of this approach thus interacts with the question of how natural it is to judge a sentence true on the basis of the fact that it is tolerantly true. Perhaps it is quite unnatural in other cases as well, and in this case there is no problem for the approach we have taken to presuppositions. But note that Cobreros&al. rely on the assumption that it *is* natural specifically in their treatment of the sorites paradox, and in their explanation of the both-true-and-false judgments associated with vague predicates as observed by Alxatib&Pelletier. The issue clearly deserves further examination.

work Fine [6] proposes that a vague expression is (*super*)true or (*super*)false if and only if, no matter how we precisify our language, its logical value would still be truth/falsity — and neither (definitely/super)true nor (definitely/super>false when its logical value depends on how we precisify our language.

2.3 Presupposition projection with connectives (*and*, *or*, *if*)

We have adopted an approach on which the presuppositions of a statement are derived from the conditions for its strict truth and strict falsity. Since our system as it stands makes predictions about the conditions under which complex sentences are strictly true and strictly false, this means that it also makes predictions about the presuppositions of complex sentences. In other words, it makes predictions about the patterns of “presupposition projection” that we should find — about the ways in which the presuppositions that we associate with complex sentences relate to the presuppositions that we associate with the sentences that they embed. We have already seen this in the case of negation.

A major challenge for theories of presuppositions is to account for presupposition projection effects in a non-stipulative way. In what follows, we review some more basic facts about presupposition projection and show to what extent our approach can account for them. Some can be accounted for without any further stipulation, and this alone seems to render our approach promising as the basis for a theory of presupposition. Others need additional stipulations.

2.4 Facts

We will be concerned here with presupposition projection effects in sentences containing the connectives *and*, *or* and *if* — and thus in sentences that we would formalize with the connectives $\&$, \vee and \rightarrow . For example, consider the following statements in (4) and their counterparts in (5) where the embedded clauses are reversed:

Example 4.

- (a) Mary has stopped smoking and she buys anti-smoking patches.
- (b) Either Mary has stopped smoking or she doesn't buy anti-smoking patches.
- (c) If Mary has stopped smoking, then she buys anti-smoking patches.

Example 5.

- (a) Mary buys anti-smoking patches and she has stopped smoking.
- (b) Either Mary doesn't buy anti-smoking patches or she has stopped smoking.
- (c) If Mary buys anti-smoking patches, then she has stopped smoking.

All of these statements embed the proposition [*she/Mary*] *has stopped smoking*, which is associated with the presupposition that Mary used to smoke. But this presupposition “projects” in different ways. The statements in (4), which embed the proposition “to the left”, retain the presupposition that Mary used to smoke (we say that here the presupposition projects globally). By contrast, those in (5), which embed the proposition “to the right,” seem to presuppose that, *if Mary buys anti-smoking patches*, then she used to smoke (we could say that here it projects conditionally or restrictively)¹¹. Adapting these sentences

¹¹ We intend our *if...then...* here to be understood as material implication: presupposition failure arises when Mary buys anti-smoking patches but has never smoked.

to our formalism ¹², it thus seems that sentences of the form in (1) presuppose that $C[p]$ while sentences of the form in (2) conditionalize this presupposition as below:

Form 1.

- a. $A_p \& B$ *presupposition: $C[p]$*
- b. $A_p \vee B$ *presupposition: $C[p]$*
- c. $A_p \rightarrow B$ *presupposition: $C[p]$*

Form 2.

- a. $B \& A_p$ *presupposition: $C[p]$*
- b. $B \vee A_p$ *presupposition: $C[p]$*
- c. $B \rightarrow A_p$ *presupposition: $C[p]$*

(In support of the claim that sentences that embed the proposition “to the right” should be viewed as behaving in this way, consider the sentences in (6). Unlike the sentences in (5), these seem to be associated with no presupposition at all. This effect can be explained if the process that derives presuppositions derives for (6a) — for example — the presupposition that, *if Mary used to smoke, then she used to smoke*. This presupposition is vacuous: it could never be false.)

Example 6.

- a. Mary used to smoke and she has stopped smoking.
- b. Either Mary never smoked or she has stopped smoking.
- c. If Mary used to smoke, then she has stopped smoking.

2.5 TCS predictions

As things stand, with no further modification, our system makes the following predictions.

It correctly derives the conditional presuppositions for sentences of the form in 2 above. To see this, recall that a proposition P is associated with the presupposition that $S[P]$ or $S[\neg P]$. Below we show that in this case this amounts to the conditional presupposition we specified above. (In what follows, it is useful to recall that classical truth and falsity are defined in such a way that, for any ϕ , $C[\neg\phi]$ iff $\neg C[\phi]$. Moreover, we assume that B in all of the cases we discuss is associated with no presupposition, and thus that $S[B]$ iff $C[B]$ and $S[\neg B]$ iff $C[\neg B]$. It is also useful to remember that it follows from Definition 2 that, for any ϕ , $S[\neg\neg\phi]$ iff $S[\phi]$.¹³)

¹² In what follows, we will use A_p to represent a proposition that is associated with the presupposition that p . In other words, strict truth and falsity are determined as follows for such a proposition: A_p is strictly true iff p is classically true and A is classically true; A_p is strictly false iff p is classically true and A is classically false.

¹³ This is because $S[\neg[\phi\&\psi]]$ iff $\neg T[\phi\&\psi]$ (by Definition 2) iff $\neg T[\phi]$ or $\neg T[\psi]$ (by Definition 3) iff $S[\neg\phi]$ or $S[\neg\psi]$ (by Definition 3).

Proof ($B \& A_p$ is associated with the presupposition that if $C(B)$ then $C(p)$).

- a. $B \& A_p$ is associated with the presupposition that $S[B \& A_p]$ or $S[\neg[B \& A_p]]$.
(Principle 1)
- b. $S[B \& A_p]$ iff $S[B]$ and $S[A_p]$ (Definition 3)
iff $C[B]$ and $C[p]$ and $C[A]$ (Definition 5)
- c. $S[\neg[B \& A_p]]$ iff $\neg T[B \& A_p]$ (Definition 2)
iff $\neg T[B]$ or $\neg T[A_p]$ (Definition 3)
iff $S[\neg B]$ or $S[\neg A_p]$ (Definition 2)
iff $C[\neg B]$ or ($C[p]$ and $C[\neg A]$) (Definition 5)
iff $\neg C[B]$ or ($C[p]$ and $\neg C[A]$)
- d. To say that $\frac{C[B] \text{ and } C[p] \text{ and } C[A]}{C[B]}$ or $\frac{\neg C[B] \text{ or } (C[p] \text{ and } \neg C[A])}{C[p]}$
is to say that $\neg C[B]$ or $C[p]$, or in other words that if $C[B]$ then $C[p]$.¹⁴

□

Proof ($B \vee A_p$ is associated with the presupposition that if $\neg C(B)$ then $C(p)$).

- a. $B \vee A_p$ is associated with the presupposition that $S[B \vee A_p]$ or $S[\neg[B \vee A_p]]$.
(Principle 1)
- b. $S[B \vee A_p]$ iff $S[\neg[\neg B \& \neg A_p]]$ (Principle of connectives)
iff $S[B]$ or $S[A_p]$ (Definition 2,3)
iff $C[B]$ or ($C[p]$ and $C[A]$) (Definition 5)
- c. $S[\neg[B \vee A_p]]$ iff $S[\neg[\neg[\neg B \& \neg A_p]]]$ (Principle 3)
iff $S[\neg B \& \neg A_p]$ (Definition 2)
iff $S[\neg B]$ and $S[\neg A_p]$ (Definition 3)
iff $C[\neg B]$ and $C[p]$ and $C[\neg A]$ (Definition 5)
iff $\neg C[B]$ and $C[p]$ and $\neg C[A]$
- d. To say that $\frac{C[B] \text{ or } (C[p] \text{ and } C[A])}{C[B]}$ or $\frac{\neg C[B] \text{ and } C[p] \text{ and } \neg C[A]}{C[p]}$
is to say that $C[B]$ or $C[p]$, or in other words that if $\neg C[B]$ then $C[p]$.

□

Proof ($B \rightarrow A_p$ is associated with the presupposition that if $C(B)$ then $C(p)$).

1. $B \rightarrow A_p$ is associated with the presupposition that $S[B \rightarrow A_p]$ or $S[\neg[B \rightarrow A_p]]$.
(Principle 1)
2. $S[B \rightarrow A_p]$ iff $S[\neg[B \& \neg A_p]]$ (Definition 3)
iff $S[\neg B]$ or $S[A_p]$ (Definitions 2,3)
iff $C[\neg B]$ or ($C[p]$ and $C[A]$) (Definition 5)
iff $\neg C[B]$ or ($C[p]$ and $C[A]$)
3. $S[\neg[B \rightarrow A_p]]$ iff $S[\neg[\neg[B \& \neg A_p]]]$ (Definition 3)
iff $S[B \& \neg A_p]$ (Definition 2)
iff $S[B]$ and $S[\neg A_p]$ (Definition 3)
iff $C[B]$ and $C[p]$ and $C[\neg A]$ (Definition 5)
iff $C[B]$ and $C[p]$ and $\neg C[A]$

¹⁴ To see this, note that we intend the *and* and *or* in our metalanguage to behave like the connectives of classical logic, and the formula $[[p \& [q \& r]] \vee [\neg p \vee [q \& \neg r]]]$ is equivalent to $[\neg p \vee q]$.

4. To say that $\neg C[B]$ or $(C[p]$ and $C[A])$ or $C[B]$ and $C[p]$ and $\neg C[A]$ is to say that $\neg C[B]$ or $C[p]$, or in other words that if $C[B]$ then $C[p]$.

□

On the other hand, the system also *incorrectly* derives conditional presuppositions for sentences of the form in 1. It is obvious that it derives exactly the same presuppositions for 1a,b that it does for 2a,b, where the embedded propositions are switched. This is because “&” is symmetric: for any ϕ, ψ , $S[\phi \& \psi]$ iff $S[\psi \& \phi]$ and $S[\neg[\phi \& \psi]]$ iff $S[\neg[\psi \& \phi]]$. In the case of 1c, we will derive the same presupposition that we derive for the disjunction 2b:

Proof ($A_p \rightarrow B$ is associated with the presupposition that if $\neg C(B)$ then $C(p)$).

1. $A_p \rightarrow B$ is associated with the presupposition that $S[A_p \rightarrow B]$ or $S[\neg[A_p \rightarrow B]]$.
2. $S[A_p \rightarrow B]$ iff $S[\neg[A_p \& \neg B]]$ iff $S[\neg A_p]$ or $S[B]$
 iff $(C[p]$ and $\neg C[A])$ or $C[B]$
 $S[\neg[A_p \rightarrow B]]$ iff $S[\neg\neg[A_p \& \neg B]]$ iff $S[A_p]$ and $S[\neg B]$
 iff $C[p]$ and $C[A]$ and $\neg C[B]$
 To say that $(C[p]$ and $\neg C[A])$ or $C[B]$ or $C[p]$ and $C[A]$ and $\neg C[B]$ is to say that $C[p]$ or $C[B]$, or in other words that if $\neg C[B]$ then $C[p]$.

□

A possible way of getting around this problem is to supplement this system with a principle that effectively projects by brute force presuppositions associated with the proposition on the left side of a connective. For example, one could supplement the definitions of strict truth and falsity with a clause that specifically concerns propositions built out of connectives with a presuppositional item on the left:

Definition 6 (“Incremental” strict truth and falsity). Let P_p be a proposition associated with the presupposition p , Q a proposition and \times any connective.

Then:

$P_p \times Q$ is strictly true if and only if p is classically true and $P \times Q$ is strictly true.

$P_p \times Q$ is strictly false if and only if p is classically true and $P \times Q$ is strictly false.¹⁵

This can be seen as a variant of a proposal by Fox [7] to account for presupposition projection within a trivalent (Strong Kleene) logic. (For his precise proposal, Fox draws his inspiration from Schlenker’s [14] pragmatically motivated

¹⁵ An anonymous reviewer complained that the need for this definition made the theory non predictive. Apparently, the reviewer felt that this definition implies that the projection behavior of a connective has to be lexically encoded. Although it is stated as a definition, one could arguably view this as a parsing principle, since it is designed to apply in full generality and full recursivity to a sequence of two propositions connected with *any* connective. Report to Fox [7] and Schlenker [14] for further details on this kind of implementation.

approach to presupposition projection, and one could entertain alternatives to definition 6 that arrive at the same effect in a different way.) These additions to the definitions of strict truth and falsity would have no effect as far as the forms in 2 are concerned, but they would associate the forms in 1 with an additional presupposition that $C[p]$. The result as a whole will then be that the forms in 1 presuppose that $C[p]$, as desired.

2.6 Comments and further extensions

We just saw that, with the addition of a further order-sensitive stipulation concerning the strict truth and falsity of propositions formed with connectives, the TCS system derives correct presuppositions for a variety of complex sentences. Is there a reason to favor it over other kinds of approaches of presupposition projection? At least we can say the following. For one thing, a TCS system that computes strict values alongside classical values seems to handle some presupposition facts more naturally than the simplest trivalent approaches that come to mind. Consider examples (2) and (3) again, and recall that they both give rise to two kinds of judgments in cases where it is known that Mary never smoked. The most natural judgment for both is a judgment of truth-valuelessness, but there is also another judgment possible, in particular when one is asked to choose between true and false: false for (2) and true for (3). Now, in order to account for the first, more natural judgment, a system that simply computes trivalent truth values would assign the sentences the third value #, but in that case it is not straightforward to account for the second kind of judgment: if, in order to account for the “false” judgment for (2), we were to say that in special circumstances a “false” judgment can correspond to #, then we would expect a “false” judgment for (3) as well, contrary to fact. By contrast, we have seen that the TCS system can account for the two kinds of judgment easily: the first concerns strict truth values, the second concerns classical truth values. The TCS system as we have presented it is also arguably immune to the kinds of criticisms that have been levelled against Heims [11] approach to presupposition projection in terms of context-change semantics. On Heims approach, there is a sense in which the aspects of a connectives semantics that are responsible for its contribution to presupposition projection are arbitrary and not intrinsically related to its contribution to truth conditions. On our approach, by contrast, there is a systematicity to the way in which a connective encodes its contribution to presupposition projection: a connectives contribution to presupposition projection derives from its strict meaning, which is related to its classical meaning.

Since our TCS system for presuppositions is inspired by an account for vagueness, and draws a connection between borderline cases and presupposition failure — the judgment of no clear truth value arises in the same way in the two cases — one might think that this approach generally predicts parallel behavior for the two kinds of predicates. In particular, one might think that it leads us to expect parallel “presupposition projection” effects. We would like to point out that, while systematically parallel behavior for the two kinds of predicates would certainly lend credence to an approach like ours, a difference in behavior could

still be compatible with it: we defined strict truth and falsity differently for the two kinds of predicates in the basic cases (compare Definition 1 and Definition 5), and in principle they could also be defined differently when more complex constructions are involved. To make this remark concrete, consider a specific case where we seem *not* to find parallel behavior for the two kinds of predicates. Benjamin Spector points out, in work in progress, that sentences like (7a) and (8a) below do not seem to behave in the same way. On the one hand, when we consider Sarah to be a borderline case for *tall* (i.e. when we consider *Sarah is tall* to lack a truth value), we naturally judge (7a) false if we consider Mary to be short (and thus answer “No” to (7b)); on the other hand, when we know that Sarah never smoked (i.e. when we consider *Sarah has stopped smoking* to lack a truth value), we are uneasy giving a truth judgment for (8a) (or answering “Yes” or “No” to (8b)) even if we know that Mary began smoking at some point in the past and never gave up.

Example 7.

- a. Sarah and Mary are both tall.
- b. Are both Sarah and Mary tall?

Example 8.

- a. Sarah and Mary have both stopped smoking.
- b. Have both Sarah and Mary stopped smoking?

Viewing things from the standpoint of our approach, it thus seems that the strict falsity of *Mary is tall* suffices to make (7a) strictly false while the strict falsity of *Mary has stopped smoking* does not on its own render (8a) strictly false — rather, in a sense, the presupposition associated with *Sarah has stopped smoking projects*. While we suspect that the factors behind this difference are complex, suppose for the sake of argument that the facts are as follows: quantifying over a PPO projects the presupposition “proportionally” (cf. Chemla [5], George [10]) while quantifying over a vague predicate results in a neither-true-nor-false judgment in those instances where, so to speak, one could only consider the sentence true by counting borderline cases as not being such. (So, in the case at hand where the quantifier is universal, we arrive for (8b) at the presupposition that both used to smoke. We do not arrive at a neither-true-nor-false judgment for (7a) because counting Sarah as tall would not serve to make the sentence true.) We could derive this result if we defined strict truth and falsity for quantified statements differently for cases involving PPOs and for cases involving vague predicates. A sketch follows (where we assume that the logical language contains restricted quantification). It will follow from Definition 7 that a proposition of the form $Qx[A(x)][Pp(x)]$ presupposes that $Qx[A(x)][p(x)]$ is classically true. It will follow from Definition 8 that a proposition of the form $Qx[A(x)][P(x)]$ (where P is a vague predicate) presupposes, in effect, that the proposition is not merely true by virtue of considering P tolerantly.¹⁶

¹⁶ For the sake of legibility, Definition 8 is formulated in terms of tolerant truth rather than strict falsity, but recall that a proposition is strictly false iff it is not tolerantly

Definition 7 (Quantification with Q over a PPO P_p).

A proposition of the form $Qx[A(x)][Pp(x)]$ is strictly true if and only if $Qx[A(x)][p(x)]$ is classically true and $Qx[A(x)][P(x)]$ is classically true. A proposition of the form $Qx[A(x)][Pp(x)]$ is strictly false if and only if $Qx[A(x)][p(x)]$ is classically true and $Qx[A(x)][P(x)]$ is classically false.

Definition 8 (Quantification with Q over a vague predicate P).

Let $S_{S:A}$ abbreviate the set of propositions $\{ P(z) \mid z \text{ is a constant and } A(z) \text{ is classically true} \}$ and let δ_Q abbreviate the proportion associated with the quantifier Q .

A proposition of the form $Qx[A(x)][P(x)]$ is strictly true if and only if δ_Q of the propositions in $S_{P:A}$ are strictly true.

A proposition of the form $Qx[A(x)][P(x)]$ is tolerantly true if and only if δ_Q of the propositions in $S_{P:A}$ are tolerantly true.

3 Conclusion

We have suggested here that the TCS system can be adapted to account for presuppositional phenomena in a coherent way, rendering it a legitimate alternative to other treatments of presuppositions. Seeing the TCS system as a theory of presuppositions, one can in fact view part of this theory — the treatment of connectives — as independently motivated, given its original role in accounting for phenomena involving vague predicates. However, we have also seen that some stipulations have to be added that seem to go beyond what is needed in order to account for vagueness.

Our initial motivation in extending TCS to presuppositions was the feeling that vagueness and presupposition have much in common, and more specifically the suspicion that expressions that lead to the lack of a clear truth-value judgment might be treated similarly by the grammar. (Which is not to imply that the lack of a truth-value judgment always arises for the same reason. We have suggested, though, that expressions that give rise to this phenomenon might have in common an essential ambiguity between two kinds of meaning.) At the same time, we touched on some tentative evidence that vague predicates and PPOs differ with respect to a detail of their “presupposition projection” behavior. Further empirical work will be needed to determine whether the facts really force us to different stipulations for vague predicates and for PPOs within a TCS system. We also mentioned in passing another area where the facts need to be better understood: facts concerning the naturalness of tolerant meanings in constructions involving vague predicates could have a bearing on the question of whether the same system should be used in order to treat presuppositions.

true. The result of Definition 8 is thus that $Qx[A(x)][P(x)]$ presupposes that either (i) δ_Q of the propositions in $S_{P:A}$ are strictly true; or (ii) it is not the case that δ_Q of the propositions in $S_{P:A}$ are tolerantly true. In other words, $Qx[A(x)][P(x)]$ presupposes that, if it is the case that δ_Q of the propositions in $S_{P:A}$ are tolerantly true, then it is also the case that δ_Q of the propositions in $S_{P:A}$ are strictly true — or, put another way, that borderline cases cannot be held responsible for truth.

References

1. Alxatib, S., Pelletier, J. F.: The psychology of vagueness: Borderline cases and contradictions, *Mind and Language* (2010)
2. Beaver, D. L., Geurt, B.: article "Presupposition", *Stanford Encyclopedia of Philosophy*, URL <http://plato.stanford.edu/entries/presupposition/> (2011)
3. Cobreros, P., Egre, P., Ripley, D., van Rooij, R., Tolerant, classical, strict, *Journal of Philosophical Logic*. (2011)
4. Charlow, S., 'Strong' predicative presuppositional objects, in Natahan Klinedinst and Daniel Rothschild (eds.), *Proceedings of ESSLLI 2009 Workshop: New Directions in the Theory of Presupposition*. (2009)
5. Chemla, E., Similarity: Towards a Unified Account of Scalar Implicatures, Free Choice Permission and Presupposition Projection. Ms., ENS & MIT. URL <http://www.emmanuel.chemla.free.fr/Material/Cemla-StandPres.pdf>. (2008)
6. Fine, K., Vagueness, truth and logic, *Synthese* 30, 265-300. D. Reidel Publishing Company, Dordrecht-Itolland (1975)
7. Fox, D., Two short notes on Schlenkers theory of presupposition projection, *Theoretical Linguistics* 34(3), pp. 237-252. (2008)
8. van Fraassen, B. C., Singular Terms, Truth-Value Gaps, and Free Logic, *The Journal of Philosophy* 63 (17), pp. 481-495. (1966)
9. Frege, G., *Uber Sinn und Bedeutung*, *Zeitschrift fur Philosophie und philosophische Kritik*, C: 25-50. English Translation: On Sense and Meaning, in McGuinness, B. (ed), *Frege: collected works*, Oxford: Basil Blackwell, 157-177. (1892)
10. George, B. R., Predicting Presupposition Projection: some alternatives in the strong Kleene tradition. M.A. thesis, UCLA. (2008)
11. Heim, I., On the Projection Problem for Presuppositions, in Michael Barlow, Daniel Flickinger & Michael Westcoat (eds.) *Second Annual West Coast Conference on Formal Linguistics*, 114-126. Stanford University. (1983)
12. Kennedy, C., Vagueness and grammar: The semantics of relative and absolute gradable predicates. in *Linguistics and Philosophy*, 30: 1-45. (2007)
13. Russell, B., On Denoting, *Mind*, 14: 479-493. (1905)
14. Schlenker, P., Local Contexts, *Semantics & Pragmatics Volume 2, Article 3: 1-78*, 2009. (2008)