– Final Manuscript for Logique & Analyse –

TERM NEGATION IN FIRST-ORDER LOGIC

Igor Sedlár Karel Šebela

Abstract

We provide a formalization of Aristotelian term negation within an extension of classical first-order logic by two predicate operators. The operators represent the range of application of a predicate and the term negation of a predicate, respectively. We discuss several classes of models for the language characterised by various assumptions concerning the interaction between range of application, term negation and Boolean complementation. We show that the discussed classes can be defined by sets of formulas. In our intended class of models, term negation of P corresponds to the complement of P relative to the range of application of P. It is an established fact about term negation that it does not satisfy the the principle of Conversion by Contraposition. This seems to be in conflict with the thesis, put forward by Lenzen and Berto, that contraposition is a minimal requirement for an operator to be a proper negation. We show that the arguments put forward in support of this thesis do not apply to term negation.

Keywords: Aristotle · Contraries · Contraposition · Law of Excluded Middle · Negation · Term negation

1 Introduction

It is well known that Aristotle distinguished two kinds of negation—*predicate denial* consisting of a predicate being denied of a subject, e.g. 'Stone is not ill' and *term negation* consisting of a negative predicate being affirmed of a subject, e.g. 'Stone is not-ill'; see Horn (1989, 15). Negative predicates, also translated as 'indefinite' or 'infinite' predicates, are obtained by prefixing a *term-negation operator* (in English usually 'not-', 'non-' or 'in-') to ordinary positive predicates. While 'Stone is ill' and the corresponding predicate denial 'Stone is not ill' are contradictories, i.e. precisely one of them is true, 'Stone is ill' and 'Stone is not-ill' are contraries, i.e. at most one of them is true but both can be false. According to Aristotle, both 'Stone is ill' and 'Stone is not-ill' are false if 'Stone' does not denote a person or an animal—but a stone, for example—or if it does not denote at all.[1] Hence, both the Law of Non-Contradiction and the Law of Excluded Middle—in (one of) their predicate version(s); see (Horn 1989, 20)—apply to predicate denial while only the former applies to term negation.

Although there are important differences, Boolean sentential negation carries some aspects of Aristotelian predicate denial; 'Stone is ill' and its Boolean negation 'It is not the case that Stone is ill' are contradictory in the Aristotelian sense.[2]

Keeping the distinction between Boolean and term negation is one of the main motivations of the *Logic of Terms* introduced by Fred Sommers (1967; 1970; 1982). The Logic of Terms, however, is a stark alternative to first-order logic and one might wonder if there is room for a compromise. Is it possible to formalise term negation within a framework that is still quite close to classical first-order logic? The present article answers affirmatively by formalising term negation within a simple extension of classical first-order logic.

We introduce an extension of first-order logic by *predicate operators* '$\overline{}$' (term negation) and '$\widehat{}$' (predicate range). The range $\widehat{P}$ of a predicate P can be seen as the *range of applicability* of P, denoting things that may be meaningfully—even if not truly—described by P. There are two basic assumptions concerning the relations between P, $\widehat{P}$ and $\overline{P}$ for which we find explicit textual evidence in Aristotle, namely, that the extensions of P and $\overline{P}$ are disjoint and that all objects in the extension of $\widehat{P}$ are in the extension of P or in the extension of $\overline{P}$ (we give the textual evidence below). It is also implicit in the notion of range that the extension of P be a subset of the extension of $\widehat{P}$. These three conditions characterise the class of *weakly Aristotelian* models and yield our basic logic. We give a set of axioms that define the class of Aristotelian models in the sense that the axioms are valid precisely in weakly Aristotelian models. Subsequently we define additional subclasses of weakly Aristotelian models. The first one is characterised by adding the requirement that the extension of $\widehat{P}$ be *equal* to the union of the extensions of P and $\overline{P}$; this requirement gives rise to the class of *Aristotelian* models. Although we have not found explicit textual evidence in Aristotle that supports the requirement, we will argue that it is implicit in his discussion of the failure of the Law of Excluded Middle in the context of

[1] Aristotle does not consider the possibility that such sentences lack truth value if the subject term does not refer: 'For if Socrates exists, one will be true and the other false, but if he does not exist, both will be false; for neither "Socrates is ill" nor "Socrates is well" is true, if Socrates does not exist at all' (*Cat.* 13b17–19). According to Aristotle, sentences lacking truth value are beyond the scope of logical investigations (cf. *De Int.* 16b33–17a8).

[2] Horn concurs: '[predicate denial] is semantically analogous to garden-variety external or weak negation in a modern propositional logic' (Horn 1989, 138); where by a 'weak negation' he means sentential Boolean negation.

term negation. Another motivation is that, unlike in Aristotelian models, term negation is not uniquely characterised in weakly Aristotelian models—we may have a predicate that satisfies all the axioms for term negation of P and still fails to be coextensional with $\overline{P}$. The second subclass discussed is the class of Aristotelian models satisfying the double (term) negation law. We give axioms characterising all the additional subclasses.

In addition to a formalisation of term negation in an extension of classical first-order logic, we discuss the relation of term negation to the thesis, put forward by Lenzen (1996) and Berto (2015), that every reasonable negation has to satisfy the Contraposition Principle. It is well known that Aristotelian term negation does not satisfy contraposition—e.g. Jean Buridan discussed counterexamples based on existential import in *Summulae de Dialectica* (Buridan 2001, 55). We put forward two observations concerning this issue. Firstly, we show that counterexamples to contraposition can be devised independently of the question of existential import. Secondly, we show that term negation does not belong to the classes of negation operators to which Berto and Lenzen's arguments apply.

The article is organised as follows. Section 2 sets the notion of term negation into historical context and discusses some formalizations known from the literature. A formalization of term negation within an extension of first-order logic is provided in Section 3 where we also discuss different classes of models for the language with term negation and their informal interpretations. In Section 4 we discuss the thesis that contraposition is a minimal requirement to be satisfied by any negation operator its consequences for the status of term negation as a negation operator. Section 5 concludes the article.

2 Term negation

This section sets the notion of term negation into a wider historical context (Sect. 2.1) and discusses some influential modern formalizations related to the notion (Sect. 2.2)

2.1 A historical overview

Term negation can be traced back to Aristotle's *De Interpretatione* (16a 30n, 16b 11n) where so called *indefinite* names and verbs are studied. These names and verbs are obtained (in English) by adding a prefix such as 'not-', 'un-' or 'non-' to regular names and verbs. For instance, 'not-man' and 'not-ill' are indefinite names. Aristotle observed that sentence pairs containing a name (verb) in one sentence and the corresponding indefinite name (verb) in the other are *contrary*, i.e. they cannot be simultaneously true:

> If it is true to say 'it is not-white', it is true also to say 'it is not white': for it is impossible that a thing should simultaneously be white and be not-white. (*Pr. An.* 51b42–52a4)

Nevertheless, such pairs are not necessarily *contradictory*, i.e. one may be false without the other being true:

> [...] everything is equal or not equal, but not everything is equal or unequal, or if it is, it is only within the sphere of that which is receptive of equality. (*Met.* 1055 10n)

Hence, statements of the form 'S is not-P' are not equivalent to 'S is not P':

> In establishing or refuting, it makes some difference whether we suppose the expressions 'not to be this' and 'to be not-this' are identical or different in meaning, e.g., 'not to be white' and 'to be not-white'. For they do not mean the same thing, nor is 'to be not-white' the negation of 'to be white', but 'not to be white' [is]. (*Pr. An.* 51b5–10)

It is clear from *Met.* 1055 that Aristotle thinks of term negation as being connected to the range of applicability of a predicate—there may be things that are neither P nor not-P, but no such thing can be found within the range of applicability of P. The connection between term negation and range of applicability is invoked in some modern formalisations of term negation (Sect. 2.2) and it will be central to our approach (Sect. 3).

The notion of term negation played an important role in the post-Aristotelian literature. Following Aristotle's discussion of contraries that 'have no intermediate' (*Cat.* 12a1–7; see Horn (1989, 7)) Boethius divides contraries into mediate ('black' and 'white', 'good' and 'bad'—in short, contraries that have an intermediate alternative, such as 'grey' in the former case and 'morally indifferent' in the latter) and immediate ('odd' and 'even', 'left-hander' and 'right-hander'), (Horn 1989, 39). According to Boethius, the term negation of a given predicate is the immediate contrary of the predicate; see Horn (1989, 7, 39). This reading of term negation is assumed by most modern formalizations and we will build on it as well.

Scholastic followers of Aristotle called sentences of the form 'S is not P' *negatio negans* while sentences of the form 'S is non-P' were called *negatio infinitas* (whence the label 'infinite negation' for the sentences with indefinite predicates). This kind of negation became an integral part of logical tradition, as we can see e.g. in Kant's theory of judgement, where in the category of Quantity

Kant distinguishes affirmative, negative and infinite judgements ('The soul is immortal', for example).

Modern mathematical logic, however, diverges from this tradition. G. Frege, in his article *Negation* (Frege 1919), famously advocates the view that all kinds of negation can be reduced to Boolean sentential negation 'It is not true that ...'. The role of term negation in contemporary logic, if any, differs from the role of Boolean negation. We will see in the next section that there is only a handful of logical systems explicitly designed to formalize a contrary-forming negation and these are mostly scattered in the literature on non-classical logic. The properties of contrary-forming negation in these systems vary and neither is commonly considered as *the* sentential cousin of Aristotle's term negation (in the sense in which Boolean negation is considered to be the cousin of predicate denial). Moreover, the coherence of the informal interpretations of these systems and their faithfulness to the Aristotelian model have been repeatedly questioned. We turn to these issues now.

2.2 Some modern formalizations

It appears that Georg Henrik Von Wright (1959) developed the first logic directly aiming at formalizing both Aristotelian negations in a propositional setting. Von Wright uses a propositional language with two negation operators for weak and strong negation, the strong one being a sentential cousin of Aristotelian term negation.

Von Wright's interpretation of strong negation uses the Aristotelian notion of a *genus*, related to range of application. The basic assumption is that among all properties we may distinguish special properties, called genera, under which specific families of properties are subsumed. A genus S is said to be *appropriate* to a property P iff, roughly, it makes sense to ask if an arbitrary member of S has the property P. For instance, 'man' is appropriate to 'English-speaking', but 'mammal' is not. There are mammals for which the question whether they speak English does not arise 'naturally'.[3] The interpretation of strong negation is given explicitly in terms of genera:

> If x belongs to *some* genus which is appropriate to P but x is not P, then and then only shall I say that the proposition 'x is P' is *false* or, which I regard as meaning the same, that the proposition 'x is not-P' is true. (Von Wright 1959, 10).

Von Wright relates the genera appropriate to a given P to the notion of the range of applicability of P:

[3] These examples are Von Wright's. He distinguishes several ways in which a genus can be in/appropriate for a property, see (Von Wright 1959, 8). We do not need to go into detail.

> If x actually does not belong to any genus at all which is [...] appropriate to P, I shall say that x is *removed from the range of application* of the predicate P. (Von Wright 1959, 9)

Hence, (the extension of) the range of applicability of P may be defined as the union of (the extensions of) all the genera appropriate to P and x is not-P iff it belongs to the range of applicability of P, but it is not P.

Barnes (1969), in his discussion of contrariety in Aristotle, invokes a similar notion when he writes that 'RF—the range of F—is, roughly, the class of objects which can be called "F" without commission of a category mistake' (p. 303). One notion of contrariety he discusses comes close to Von Wright's notion of strong negation. Barnes calls P and Q 'contradictory' (or 'contrary$_2$') iff (i) P and Q have the same range, (ii) nothing is both P and Q and (iii) everything in the range of P (and Q) is either P or Q.

The notion of range invoked by Von Wright and Barnes is, in turn, related to Sommers' notion of *span* and Bergmann's *sortal range*.[4]

McCall (1967) outlines a formalization of a contrary-forming sentential operator within propositional modal logic (the connection with modal logic is pointed out by Von Wright (1959) as well). He is criticised by Geach (1969) for formalising the notion of a contrary as an *operator*, and by Englebretsen (1974) for formalising contraries by a *sentential* operator. Geach points out that, strictly speaking, 'the contrary of P' is a misnomer as there are many contraries of any given P. [5] Englebretsen's objection, roughly speaking, is that Aristotelian term negation is not a sentential operator but an operator on predicates. We will show below that our approach avoids both of these objections.

A contrary-forming operator in propositional logic is also studied by Humberstone (2008; 2011). A contrary-forming sentential operator within first-order logic is studied by Wessel (1998) (first edition published in 1983), whose approach is set in the context of propositional modal logic by Wojciechowski (1997).

Fred Sommers (1967; 1970; 1982) introduced the Logic of Terms, an upgrade of Aristotelian syllogistic that captures term negation as an operator on predicates.[6] It is an interesting observation that term negation is a central notion of the Logic of Terms also because Sommers thinks of terms as 'coming in pairs': 'any term is positive or negative with respect to another term that is logically contrary to it' (Sommers 1970, 4).

[4]'A predicate will be said to span a thing if it is predicated of it either truly or falsely but not absurdly'. (Sommers 1963, 329); 'The extension of a predicate is the collection of all those entities of which the predicate is true, while the sortal range consists of all those entities to which the predicate is significantly applicable'. (Bergmann 1977, 61).

[5]Similar criticism may be applied to Von Wright. Just for the record, McCall is well aware of the fact pointed out by Geach.

[6]See also (Englebretsen 1981; 1996) and (Sommers and Englebretsen 2000).

3 Formalizing term negation

This section provides a formalization of term negation within a variant of first-order logic with predicate operators (Sect. 3.1). Then we discuss classes of first-order models corresponding to various properties of term negation endorsed either by Aristotle himself, or assumed within the modern formalisations of term negation. We show that all these classes of models can be defined by (sets of) formulas of our language (Sect. 3.2).

The basic idea of our formalization is to extend the standard language of first-order logic (without individual functions) with two *operators on predicates*—$\widehat{P}$ will stand for the range of applicability of predicate P and $\overline{P}$ will stand for the term negation of P. In full generality, the logic obtained by this modification is just classical first-order logic (over an extended language). However, Aristotle's discussion of term negation motivates us to focus on a narrower class of *weakly Aristotelian* models—in these models, (extensions of) P and $\overline{P}$ are disjoint (contrariety), $\widehat{P}$ is contained in the union of P and $\overline{P}$ (Law of Excluded Middle restricted to the appropriate range) and P is contained in $\widehat{P}$ (a natural requirement implicit in the notion of range). We show that term negation is not characterised uniquely in weakly Aristotelian models and, therefore, it does not avoid Geach's objection. This leads us to considering even narrower classes of models; primarily *Aristotelian models* where $\widehat{P}$ is *equal* to the union of P and $\overline{P}$ (an assumption implicit in Aristotle's discussion of term negation and corresponding to Von Wright's definition of strong negation) and *involutive* Aristotelian models where P and $\overline{P}$ have the same range—cf. Barnes' condition (i).

We note that our formalization is close in spirit to Priest's semantics for quantified First Degree Entailment (Priest 1987). Priest distinguishes between the *extension* and the *antiextension* of a predicate—the former is the class of objects of which P is true and the latter the class of objects of which it is false—and uses the latter in specifying the falsity condition of atomic formulas. For instance, $P(a)$ is false iff a belongs to the antiextension of P. However, Priest then represents falsity by means of a strong *paraconsistent sentential* negation. We assume that 'extensions' and 'antiextensions' are disjoint, so our term negation is a *non-paraconsistent predicate* operator.[7]

[7] This is in spirit of Aristotle's defence of the Law of Non-Contradiction: 'There are some, however, as we have said, who both state themselves that the same thing can be and not be, and say that it is possible to hold this view [...] But we have just assumed that it is impossible at once to be and not to be, and by this means we have proved that this is the most certain of all principles. Some, indeed, demand to have the law proved, but this is because they lack education'. *Met.* 1005b-1006a.

3.1 First-order logic with predicate operators

For all $k \in \omega$, let $\mathcal{P}_0^k$ be a countable set of *primitive* k-ary predicates. Let $\mathcal{P}^k$ be the smallest superset of $\mathcal{P}_0^k$ such that if $P \in \mathcal{P}^k$, then $\overline{P} \in \mathcal{P}^k$ and $\widehat{P} \in \mathcal{P}^k$. The predicate $\widehat{P}$ is called the *range* of predicate P and $\overline{P}$ is called the *term negation* of P. The *signature* Σ contains the union of $\mathcal{P}^k$ for all $k \in \omega$ and a countable set $\mathcal{C}$ of individual constants. We also assume a countable set $\mathcal{V}$ of individual variables. First-order formulas over Σ are defined in the usual way; we take $\wedge, \neg$ and $\exists$ to be our primitive operators and we assume the standard definitions of the other operators. In what follows, we use 'P', 'Q' etc. as variables ranging over the predicates of Σ and 'a', 'b' etc. as variables ranging over $\mathcal{C}$.

We note that Σ contains 'nested' predicates such as $\overline{\overline{P}}$, $\widehat{\widehat{P}}$ and $\overline{\widehat{P}}$. On the other hand, our language does not express term negations or ranges of complex predicates such as 'corrupt politician', corresponding to non-atomic open formulas such as $C(x) \wedge P(x)$.[8]

Our semantics uses standard first-order models. A *model (of* Σ*)* is a couple $\mathcal{M} = \langle \mathcal{U}, \mathcal{I} \rangle$, where $\mathcal{U}$ is a non-empty set ('universe') and $\mathcal{I}$ is a function that assigns (i) members of the universe to propositional constants, i.e. $\mathcal{I}(a) \in \mathcal{U}$, and (ii) k-tuples of members of the universe to k-ary predicates, i.e. $\mathcal{I}(P) \subseteq \mathcal{U}^k$ for $P \in \mathcal{P}^k$. Valuations of individual variables and the satisfaction relation are defined as usual, i.e.

- $\mathcal{M}, v \vDash P(x_1, \ldots, x_n)$ iff $\langle v(x_1), \ldots v(x_n) \rangle \in \mathcal{I}(P)$, for all $P \in \mathcal{P}^k$;
- $\mathcal{M}, v \vDash P(a_1, \ldots, a_n)$ iff $\langle \mathcal{I}(a_1), \ldots \mathcal{I}(a_n) \rangle \in \mathcal{I}(P)$, for all $P \in \mathcal{P}^k$;
- $\mathcal{M}, v \vDash \neg X$ iff $\mathcal{M}, v \nvDash X$;
- $\mathcal{M}, v \vDash X \wedge Y$ iff $\mathcal{M}, v \vDash X$ and $\mathcal{M}, v \vDash Y$;
- $\mathcal{M}, v \vDash \exists x.Z(x)$ iff there is a valuation u such that $u(y) = v(y)$ for all $y \neq x$ and $\mathcal{M}, u \vDash Z(x)$.

X is valid in $\mathcal{M}$, notation $\mathcal{M} \vDash X$, iff $\mathcal{M}, v \vDash X$ for all v. A set of formulas Γ is valid in $\mathcal{M}$, notation $\mathcal{M} \vDash \Gamma$, iff $\mathcal{M} \vDash X$ for all $X \in \Gamma$. $Mod(\Gamma)$ is the set of models in which Γ is valid. It is clear that the class of formulas valid in all models is the classical first-order logic over signature Σ.

Notation '$Z(x_1, \ldots, x_n)$' implies that every free variable in formula Z is to be found in the set $\{x_1, \ldots, x_n\}$. We shall write simply '$\forall x(Px \to Qx)$' instead '$\forall \vec{x}(P(\vec{x}) \to Q(\vec{x}))$' ($P, Q$ may be k-ary for arbitrary $k \in \omega$).

[8] The reasons for this restriction are twofold. Firstly, this approach is technically simpler. Secondly, allowing to term-negate expressions such as $C(x) \wedge P(x)$ comes very close to taking term negation as a sentential operator. We note, however, that predicates such as '*un*-corrupt politician' can be formalized as $\overline{C}(x) \wedge P(x)$.

3.2 Special classes of models

The informal interpretation of $\widehat{P}$ and $\overline{P}$ as the range of P and the term negation of P, respectively, motivates specific assumptions concerning the 'intended' models for our language. We study some of these classes of models in this section.

DEFINITION 3.1 A model of Σ is *weakly Aristotelian* iff

$$\mathcal{I}(\overline{P}) \cap \mathcal{I}(P) = \emptyset \tag{1}$$

$$\mathcal{I}(P) \subseteq \mathcal{I}(\widehat{P}) \tag{2}$$

$$\mathcal{I}(\widehat{P}) \subseteq \mathcal{I}(P) \cup \mathcal{I}(\overline{P}) \tag{3}$$

The requirement (1) states that, roughly, 'S is P' and 'S is $\overline{P}$' are *contrary*, i.e. no thing is both P and $\overline{P}$.[9] This is in line with explicit textual evidence from Aristotle:

> If it is true to say 'it is not-white', it is true also to say 'it is not white': for it is impossible that a thing should simultaneously be white and be not-white. (*Pr. An.* 51b42–52a4)

The condition (2) states the natural requirement that each thing that is P belongs also to the range of applicability of P. Although this is not explicitly stated by Aristotle in this very form (to the best of our knowledge), it is implicit in his *Dictum de omni et nullo*:

> Whenever one thing is predicated of another as of a subject, all things said of what is predicated will be said of the subject also. For example, man is predicated of the individual man, and animal of man; so animal will be predicated of the individual man also—for the individual man is both a man and an animal. (*Cat.* 1b10)

Indeed, the requirement is implicit in the very notion of range.

Finally, (3) states a restricted form or the Law of Excluded Middle. As we noted in the Introduction, 'S is P' and 'S is $\overline{P}$' may both be false but, according to Aristotle, LEM holds *within the range of the given predicate*. For instance,

> [...] everything is equal or not equal, but not everything is equal or unequal, or if it is, it is only within the sphere of that which is receptive of equality. (*Met.* 1055 10n)

[9] In fact, the condition states this for P of arbitrary arity—we restrict our informal discussion to unary predicates with the understanding that similar interpretations apply to predicates of arbitrary arity.

If we look at pairs of things that are 'receptive of equality' ($\widehat{P}$), things in each such pair are either equal (P) or unequal ($\overline{P}$). The requirement is also related to the reading of term negation in terms of immediate contraries:

> Those contraries which are such that the subjects in which they are naturally present, or of which they are predicated, must necessarily contain either the one or the other of them, have no intermediate. Thus disease and health are naturally present in the body of an animal, and it is necessary that one or the other should be present in the body of an animal. (*Cat.* 12a1–7)

Definition 3.2 A set of formulas Γ is said to *define* a class of models K in case $\mathcal{M} \models \Gamma$ iff $\mathcal{M} \in K$ (i.e. $K = Mod(\Gamma)$).

Theorem 3.3 The class of weakly Aristotelian models is defined by the set WA, the smallest set of formulas containing each instance of

- $\forall x(\overline{P}x \to \neg Px)$
- $\forall x(Px \to \widehat{P}x)$
- $\forall x(\widehat{P}x \to Px \vee \overline{P}x)$

Theorem 3.3, which follows straightforwardly from the definition of a weakly Aristotelian model, may also be seen as yielding a completeness result for a 'weakly Aristotelian' version of first-order logic with predicate operators. Let *FOLWA* denote a fixed axiomatization of first-order logic (over Σ) extended by WA as extra axioms.

Theorem 3.4 There is no finite set of formulas that defines the class of weakly Aristotelian models.

Proof (sketch). Our language does not contain predicate variables, nor predicate quantifiers. □

Hence, the class of weakly Aristotelian models is not finitely axiomatizable.

We have seen that $FOLWA$, the logic of all weakly Aristotelian models, is consistent with Aristotle's discussion of term negation. It also clearly avoids Englebretsen's objection against the sentential approaches—our term negation is explicitly an operator on predicates. It is also *as close as it gets* to classical first order logic, and so it is an alternative available to those who would like to work with a formalisation of Aristotelian term negation but are not willing to go as far as The Logic of Terms, for example.

However, weakly Aristotelian models are defective in one crucial aspect, namely, the requirements stated in their definition fail to determine term negation uniquely.

Proposition 3.5 There is a weakly Aristotelian model $\mathcal{M}$ and (unary) predicates P, Q such that versions of (1)–(3) with $\overline{P}$ replaced by Q hold, but $\mathcal{I}(\overline{P}) \neq \mathcal{I}(Q)$.

Proof. We assume that constants a, b and c have distinct interpretations and we refer to the three distinct objects in the universe using the constants. Assume that the relevant extensions are given as follows:

$$P : a \qquad Q : b, c \qquad \overline{P} : b \qquad \widehat{P} : a, b$$

It is clear that Q-versions of (1) and (3) are true, but $\mathcal{I}(\overline{P}) \neq \mathcal{I}(Q)$. □

Hence, *FOLWA* does not avoid Geach's objection, discussed in Section 2.2, that there is typically no unique contrary of a given predicate. This objection is avoided in a special subclass of weakly Aristotelian models.

Definition 3.6 A model for Σ is *Aristotelian* iff it is weakly Aristotelian and it satisfies

$$\mathcal{I}(\overline{P}) \subseteq \mathcal{I}(\widehat{P}) \tag{4}$$

Hence, in Aristotelian models, both P and the term negation of P are contained in the range of P.

Proposition 3.7 A model is Aristotelian iff

$$\mathcal{I}(P) \cap \mathcal{I}(\overline{P}) = \emptyset$$
$$\mathcal{I}(P) \cup \mathcal{I}(\overline{P}) = \mathcal{I}(\widehat{P})$$

Aristotelian models correspond, for example, to Von Wright's approach to strong negation—something is $\overline{P}$ *precisely* if belongs to the range of P but it is not P. We submit that $\mathcal{I}(\overline{P}) \subseteq \mathcal{I}(\widehat{P})$ is also implicit in the way how Aristotle argues for the failure of LEM in case the subject does not belong to the range of application of the predicate. If it were possible for some particular thing s to be $\overline{P}$ without being $\widehat{P}$, it would not be possible to claim *generally* that LEM fails for subjects not in $\widehat{P}$. For, in that case, s could satisfy $P(x) \vee \overline{P}(x)$ even if it failed to satisfy $\widehat{P}(x)$.

Theorem 3.8 The class of Aristotelian models is defined by the set of formulas A, the smallest superset of *WA* containing

$$\forall x(\overline{P}x \rightarrow \widehat{P}x)$$

for all $P \in \Sigma$.

Theorem 3.9 There is no finite set of formulas defining the class of Aristotelian models.

Proposition 3.10 For all Aristotelian models and all predicates P, Q, if Q satisfies all the properties of $\overline{P}$ given in (1)–(3), then $\mathcal{I}(P) = \mathcal{I}(Q)$.

Proof. If (the object denoted by) a is Q, then it is $\widehat{P}$ but it is not P, so a is $\overline{P}$ (by Prop. 3.7). If a is $\overline{P}$, then it is $\widehat{P}$ but it is not P. Hence, a is Q by the assumption. □

Proposition 3.10 shows that the logic of all Aristotelian models (which we may call *FOLA*) avoids Geach's objection as well as Englebretsen's—the requirements concerning the interplay between a predicate, its range and its term negation determine (the extension of) term negation of a predicate uniquely.

Definition 3.11 A model for Σ is called *involutive* iff it is Aristotelian and satisfies

$$\mathcal{I}(\widehat{P}) = \mathcal{I}(\widehat{\overline{P}}) \tag{5}$$

In involutive models, the range of a predicate and the range of its term negation have the same extension. The name 'involutive' is justified by the next proposition.

Proposition 3.12 An Aristotelian model for Σ is involutive iff

$$\mathcal{I}(\overline{\overline{P}}) = \mathcal{I}(P)$$

Proof. We show that $\mathcal{I}(\overline{\overline{P}}) \subseteq \mathcal{I}(P)$ iff $\mathcal{I}(\widehat{\overline{P}}) \subseteq \mathcal{I}(\widehat{P})$, the equivalence between the converse inclusions is established similarly. Firstly, assume $\mathcal{I}(\overline{\overline{P}}) \subseteq \mathcal{I}(P)$. If $a \in \mathcal{I}(\widehat{\overline{P}})$, then $a \in \mathcal{I}(\overline{P})$ or $a \in \mathcal{I}(\overline{\overline{P}})$. By the assumption, $a \in \mathcal{I}(\overline{P})$ or $a \in \mathcal{I}(P)$. Hence, $a \in \mathcal{I}(\widehat{P})$. Secondly, assume $\mathcal{I}(\widehat{\overline{P}}) \subseteq \mathcal{I}(\widehat{P})$. If $a \in \mathcal{I}(\overline{\overline{P}})$, then $a \in \mathcal{I}(\widehat{\overline{P}})$, and so $a \in \mathcal{I}(\widehat{P})$ by the assumption. Now if $a \notin \mathcal{I}(P)$, then $a \in \mathcal{I}(\overline{P})$, so $x \notin \mathcal{I}(\overline{\overline{P}})$. This is a contradiction, so it has to be the case that $a \in \mathcal{I}(P)$. □

Theorem 3.13 The class of involutive models is defined by *IA*, the smallest superset of A containing

$$\forall x(\widehat{\overline{P}}x \leftrightarrow \widehat{P}x)$$

for all $P \in \Sigma$.

Theorem 3.14 There is no finite set of formulas defining the class of involutive models.

Von Wright (1959) assumes (a variant of the idea) that P and $\overline{P}$ should have the same range but—interestingly enough—he also argues that double negation introduction should not be a valid principle for a strong negation. Von Wright's argument invokes the notion of *ultimate genera*, i.e. genera for which no genus is appropriate.[10] McCall (1967) shows that a modal rendering of the double negation introduction principle is inconsistent with other plausible modal principles, but he argues that both double negation laws should hold for a specific kind of contraries he calls 'strong contraries' (also known as 'polar contraries', see (Horn 1989, 39).)

In any case, we leave (5) as an optional addition. In response to Von Wright we could say that we do not share his notion of ultimate genera. In our setting, we might have $\mathcal{I}(P) = \mathcal{I}(\widehat{P})$, but this does not mean that P *has no* range. In fact, the notion of ultimate genus has an unintuitive feature—since the range of applicability of P is the union of (extensions) of all the genera appropriate to P, the extension of the range of an ultimate genus is the empty set. In that case, however, the extension of an ultimate genus, if it is non-empty, is not a subset of the extension of its range.

Note that if we *defined* $\overline{P}x$ as $\widehat{P}x \wedge \neg Px$, then the double term negation of a predicate, $\overline{\overline{P}}x$, would not be a formula of our language (recall that '$\widehat{\ }$' is an operator on predicates in Σ; also see footnote 8). We want to keep the option of expressing double negations, so we consider both predicate operators as primitive.

Proposition 3.15 The class of Aristotelian models is a proper superclass of the class of weakly Aristotelian models and a proper subclass of the class of involutive models.

In a sense, term negation may be seen as a generalization of predicate denial. Let us call an Aristotelian model *classical* iff

$$\mathcal{I}(\widehat{P}) = \mathcal{U}^k \tag{6}$$

for all $P \in \mathcal{P}^k$. Note that, in classical models,

$$\forall x(\overline{P}x \leftrightarrow \neg Px)$$

is valid. The property (6) corresponds to the assumption that the range of *any* predicate is the trivial property satisfied by everything.

The property (6) might be seen as reflecting a setting where we dispense with considering ranges of applicability of predicates altogether. This setting is related to the debate about the so called *genus generalissimum*, see (Audi 1999, 343). If $\widehat{P}$ is read as *the smallest genus subsuming P*, then (6) corresponds to the assumption that there is just one genus, genus generalissimum, subsuming all the predicates and that this genus corresponds to the trivial property—we may call it 'being'— satisfied by all objects.

[10]His argument runs, roughly, as follows. Assume that P is a non-empty ultimate genus, so there is some x that is P. If x is also not-not-P, then x belongs to some genus appropriate to not-P. By the assumption, this genus is also appropriate to P. But this is impossible since P is assumed to be a ultimate genus.

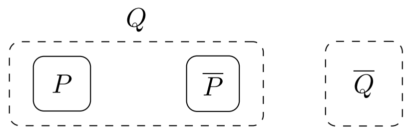

Figure 1: A counterexample to CC

4 Contraposition

We have shown so far that there is a simple formalisation of Aristotelian term negation that avoids the major objections against related approaches, but does not go as far away from classical first-order logic as, say, The Logic of Terms. Hence, our formalisation, we hope, may facilitate a 're-incorporation' of term negation to the mainstream of contemporary logic.

Nevertheless, there is a thesis, advocated by Lenzen (1996) and Berto (2015) for instance, that may render this re-incorporation problematic—in fact, the thesis seems to entail that term negation is not a 'proper' negation. The thesis is that all proper negation operators satisfy the Principle of Contraposition.[11] It is well known that term negation does not satisfy the principle of 'Conversion by Contraposition', the principle saying that every S is P only if every non-P is non-S—for instance, Buridan discussed counterexamples based on existential import in *Summulae de Dialectica* (Buridan 2001, 55). In fact, simple counterexamples can be devised even without assuming existential import. Each smoker, for instance, is a man, but, presumably, not each non-man is a non-smoker. For example, gorillas are non-men, but it does not make sense to call them non-smokers.

This general observation is independent of our formalisation, but let us turn it into a statement about our models for the sake of later discussion. We formalise the Conversion by Contraposition principle as

$$\forall x(Px \to Qx) \to \forall x(\overline{Q}x \to \overline{P}x) \tag{CC}$$

Proposition 4.1 There is an involutive model where (CC) is invalid.

Proof. See Fig. 1 and fill the details so that the model is involutive (there is nothing that would prevent this); moreover, assume that the extension of $\overline{Q}$ is non-empty. Every P is Q, but obviously not every $\overline{Q}$ is $\overline{P}$. □

[11] See also (Dunn and Hardegree 2001, 90), where an algebraic version of the contraposition principle is regarded as one of the 'unopposed [principles that] constitute the bare-bones notion of complementation, as we conceive it'. Contraposition fails in some paraconsistent logics, for example in paraconsistent Nelson's logic (Almukdad and Nelson 1984). Wansing (2001) provides a nice discussion of these issues.

Since each involutive model is (weakly) Aristotelian, (CC) fails in all the major kinds of models for term negation. Does this mean that term negation is not really a negation or that the arguments supporting the thesis put forward by Berto (2015) and Lenzen (1996) are flawed?

Let us first take a look at Berto and Lenzen's thesis in more detail. Berto claims that (notation adjusted)

> nothing can be a negation unless it satisfies Minimal Contraposition: if A entails B, then $\sim B$ entails $\sim A$. (Berto 2015, 776)

Similarly, Lenzen's claim is that

> A unary operation $\sim$ is a negation of the logic L only if it satisfies **CP1**: If $p \vdash_L q$, then $\sim q \vdash_L \sim p$. (Lenzen 1996, 46)

Both Berto and Lenzen focus on sentential negation, so we could defend term negation by saying that they simply deal with a different topic. It might be the case that proper sentential negations have to satisfy contraposition, but term negation is not a sentential negation. Nevertheless, we think that this defence is not entirely satisfactory, if only because it discourages a deeper consideration of the issue.

Notice that the difference between predicate negations and sentential negations, especially the ones of the modal kind discussed by Berto (2015), is not that big. Formulas of a propositional language can be seen as predicates expressing properties of valuations or, more generally, 'worlds'. The fact that a valuation (a world) satisfies a formula X can be seen as the valuation (the world) exemplifying the property expressed by X. To take an informal example, 'London is the capital of the UK' can be seen as expressing the property of possible worlds exemplified by worlds in which London is the capital of the UK; similarly 'London is the capital of the UK and the largest city in Europe' expresses a complex property of worlds exemplified by worlds in which London is the capital of the UK *and* the largest city in Europe. So, in a sense, sentential negations can be seen as operators generating complex properties of worlds or, as we shall put it in what follows, world-predicate negations. Interestingly, Boolean negation then corresponds to predicate denial ('w does not satisfy the property expressed by p'). On the other hand, modal negation is often motivated by an 'internal perspective' on which, roughly, all formulas express predicate ascription—negation corresponds to a negative predicate, not to predicate denial (see the discussion in (Restall 2000, ch. 16.1), for example).

Given this perspective on sentential negation, the requirements put forward by both Berto and Lenzen come very close to (CC). Note that formula A entails

B iff, in every model, each individual (i.e. world, valuation etc.) that satisfies the property expressed by A satisfies also the property expressed by B. Hence, for the sake of discussion, we may construe Berto and Lenzen's thesis as the claim that a variant of (CC) should hold for all world-predicates and their predicate negations:

$$\forall x(Ax \to Bx) \to \forall x(\widetilde{B}x \to \widetilde{A}x) \tag{CC′}$$

We shall now take a look at *why* Berto and Lenzen think that (CC′) should hold (again, the world-predicate perspective is assumed for the sake of discussion). This will allow us to explain how exactly term negation avoids their arguments. Berto (2015) infers (CC′) from the thesis that the extension of $\widetilde{A}$ is to be characterised in terms of a *compatibility* relation between individuals in the universe:

For all x, x is $\widetilde{A}$ iff no y such that $C(x, y)$ is A (C)

In fact, it is a simple exercise to show that (CC′) is valid in each model satisfying (C), independently of any particular interpretation of $C(x, y)$. Hence, thus far, Berto's argument is impeccable. Yet, the failure of (CC) shows that term negation *cannot* be characterised by a binary relation between individuals. Rather, term negation (at least in Aristotelian models) is characterised by a relation of compatibility between individuals *and predicates*—we might say that x is compatible with P if x belongs to the range of P (if x can be called 'P' without 'commission of a category mistake' as Barnes would have put it; or if x is 'receptive' of P in Aristotle's words).[12]

We do not consider Berto's claim that compatibility is a basis for negation to be false, though. Term negation only shows that one has to take a wider notion of compatibility into account, namely, compatibility between individuals and predicates; and, importantly, that interesting negations can be characterised by using other logical constructs than the 'for all...not...' one used by Berto. The construction used in the case of term negation does not even contain quantifiers; we may say (speaking of Aristotelian models) that x is $\overline{P}$ iff x is compatible with P but x is not P.

Let us now look at the main argument in favour of (CC′) put forward by Lenzen (1996). The argument runs as follows. If A entails B, then, necessarily, A is true only if B is true. So by classical contraposition, necessarily, B is *not true* only if A is *not true*. But then, Lenzen argues (notation adjusted),

[12]This notion of compatibility can be used to characterise term negation by a modified version of (C), namely, the principle that, for all x, x is $\overline{P}$ iff $C'(x, P)$ and x is not P, where '$C'(x, P)$' means that x is $\widehat{P}$. This version of (C) hold in Aristotelian models (cf. Proposition 3.7), but not in all weakly Aristotelian models, a fact that may be seen as a reason to consider Aristotelian models as the intended models for term negation.

> to say that some proposition A is not 'true' (either in the sense of classical, two-valued semantics or in the sense of some other distinguished true-like value) appears to be tantamount to *negating* A (in some way or another). Thus [... if A entails B], then, necessarily (if B is 'false', then A must be 'false', too). (Lenzen 1996, 45)

Lenzen seems to assume here that 'A is not true' and 'A is false' are equivalent. This equivalence is, at least in the case of term negation, far from being clear. Let us recast Lenzen's argument from the world-predicate perspective (where 'necessarily' means 'for all objects in the universe', assuming the S5 picture of modality for the sake of simplicity):

1. $\forall x(Ax \rightarrow Bx)$ — Assumption
2. $\forall x(\neg Bx \rightarrow \neg Ax)$ — 1., Classical contraposition
3. $\forall x(\widetilde{B}x \rightarrow \widetilde{A}x)$ — 2., Equivalence of 'not true' and 'false'

It is clear that Lenzen's argument does not apply to term negation—the difference between predicate denial ('not true') and affirmation of the negative predicate ('false') was the crucial distinction constituting the notion of term negation in the first place. The only thing that could be correctly concluded in the case of term negation is

3′. $\forall x(\overline{B}x \rightarrow \neg Ax)$ — 2., 'false' implies 'not true'

Again, our conclusion is a modest one. Lenzen's argument correctly applies to negations for which 'not true' and 'false' coincide. Term negation, however, is not such a negation.

To conclude the discussion of whether term negation is a negation in light of Berto's and Lenzen's arguments, we claim that their arguments apply to special classes of negations (ones that can be characterised in terms of a binary compatibility relation on individuals and ones for which 'p is not true' and '$\sim p$ is true' are equivalent, respectively) and that term negation does not belong to either of these classes. If a positive argument as to why term negation *should* be called negation is desired, we cannot do much more than to argue by authority, in a sense. The fact that Aristotle thought about term negation as being a proper and important kind of negation could give us reasons to adopt the view until compelling reasons against it are produced.

We conclude this section by two propositions concerning circumstances under which (CC) holds and a valid variant of (CC).

Proposition 4.2 (CC) holds in Aristotelian models where $\forall x(\widehat{Q}x \rightarrow \widehat{P}x)$.

Proof. Assume that $\forall x(Px \to Qx)$. Now $\overline{Q}a$ implies $\widehat{Q}a$ implies $\widehat{P}a$; moreover, $\overline{Q}a$ implies $\neg Qa$ implies $\neg Pa$. But $\widehat{P}a \wedge \neg Pa$ implies $\overline{P}a$. □

PROPOSITION 4.3 The Strong Conversion by Contraposition principle

$$\forall x((Px \to Qx) \to (\overline{Q}x \to \overline{P}x)) \tag{SCC}$$

is not valid in either of the classes of models discussed above (as it entails (CC), a principle not valid in these classes), but the Restricted SCC principle

$$\forall x(\widehat{P}x \to ((Px \to Qx) \to (\overline{Q}x \to \overline{P}x))) \tag{RSCC}$$

is valid in all of the classes.[13]

Proof. If $\widehat{P}a$, $\overline{Q}a$ and $\neg\overline{P}a$, then Pa and $\neg Qa$. □

5 Conclusion

In this article we discussed a formalization of Aristotelian term negation within an extension of classical first-order logic. The extension is a simple one, adding two operators on predicates. The upshot is that one does not have to go as far as, e.g., the Logic of Terms, to have a logical system with an operator representing Aristotelian term negation. We discussed classes of models that correspond to various assumptions concerning the nature of term negation. We have seen that in the weakest class of models motivated by textual evidence, term negation is not characterised uniquely. This observation motivated the definition of so-called Aristotelian models where term negation of P turns out to be the weakest contrary of P relative to the range of applicability of P. We gave (infinite) sets of formulas defining all the classes of models discussed in the article. It is well known that term negation does not satisfy the principle of Conversion by Contraposition, closely related to the principle of Minimal Contraposition. The latter principle is often pointed out as a necessary requirement for an operator to be a negation (Lenzen 1996, Berto 2015). We have noted that arguments supporting this thesis are valid for special classes of operators, yet, ones term negation does not belong to.

Acknowledgements

The authors thank the reviewers for useful feedback on a previous version of the article. Their comments have helped to improve the article significantly.

[13]We owe this observation to a reviewer.

Igor Sedlár acknowledges the support of the grant JSPS-16-08 for the project 'Nonclassical Mathematics Based on Fuzzy, Paraconsistent and Substructural Logics,' provided jointly by the Czech Science Foundation and the Japanese Society for the Promotion of Science. While working on the paper, Sedlár has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 689176.

References

Almukdad, A. and Nelson, D. (1984), 'Constructible falsity and inexact predicates', *The Journal of Symbolic Logic* **49**(1), 231–233.

Audi, R. (1999), *The Cambridge Dictionary of Philosophy*, 2 edn, Cambridge University Press, Cambridge.

Barnes, J. (1969), 'The law of contradiction', *The Philosphical Quarterly* **19**(77), 302–309.

Bergmann, M. (1977), 'Logic and sortal incorrectness', *The Review of Metaphysics* **31**(1), 61–79.

Berto, F. (2015), 'A modality called 'negation'', *Mind* **124**(495), 761–793.

Buridan, J. (2001), *Summulae de Dialectica*, Yale University Press, New Heaven. English translation by G. Klima.

Dunn, J. M. and Hardegree, G. M. (2001), *Algebraic Methods in Philosophical Logic*, Oxford University Press, Oxford.

Englebretsen, G. (1974), 'A note on contrariety', *Notre Dame Journal of Formal Logic* **15**(4), 613–614.

Englebretsen, G. (1981), *Logical Negation*, Van Gorcum, Assen.

Englebretsen, G. (1996), *Something to Reckon With: The Logic of Terms*, University of Ottawa Press, Ottawa.

Frege, G. (1919), Die Verneinung. Eine Logische Untersuchung., *in* 'Beiträge zur Philosophie des deutschen Idealismus, I', pp. 143–157. (Translated as 'Negation', by P. Geach and R. Stoothoff, in McGuinness (ed.) *Collected Papers on Mathematics, Logic, and Philosophy*, Oxford: Blackwell, 1984, pp. 373–389.).

Geach, P. T. (1969), 'Contradictories and contraries', *Analysis* **29**(6), 187–190.

Horn, L. (1989), *A Natural History of Negation*, Chicago University Press, Chicago.

Humberstone, L. (2008), 'Contrariety and Subcontrariety: The Anatomy of Negation (with Special Reference to an Example of J.-Y. Béziau)', *Theoria* **71**(3), 241–262.

Humberstone, L. (2011), *The Connectives*, MIT Press, Cambridge, MA.

Lenzen, W. (1996), Necessary conditions for negation operators, *in* H. Wansing, ed., 'Negation: A Notion in Focus', De Gruyter, New York and Berlin, pp. 37–58.

McCall, S. (1967), 'Contrariety', *Notre Dame Journal of Formal Logic* **8**(1-2), 121–132.

Priest, G. (1987), *In Contradiction*, Martinus Nijhoff, Dordrecht. (Second expanded edition Oxford, Oxford University Press, 2006).

Restall, G. (2000), *An Introduction to Substrucutral Logics*, Routledge, London.

Sommers, F. (1963), 'Types and ontology', *The Philosophical Review* **72**(3), 327–363.

Sommers, F. (1967), On a Fregean dogma, *in* I. Lakatos, ed., 'Problems in the Philosophy of Mathematics. Proceedings of the International Colloquium in the Philosophy of Science, London, 1965', Vol. 1, North-Holland Publishing Company, Amsterdam, pp. 47–81.

Sommers, F. (1970), 'The calculus of terms', *Mind* **79**(313), 1–39.

Sommers, F. (1982), *The Logic of Natural Language*, Clarendon Press, Oxford.

Sommers, F. and Englebretsen, G. (2000), *An Invitation to Formal Resoning: The Logic of Terms*, Ashgate, Farnham.

Von Wright, G. H. (1959), 'On the logic of negation', *Commentationes Physico-Mathematicae* **22**(4), 1–30.

Wansing, H. (2001), Negation, *in* L. Goble, ed., 'The Blackwell Guide to Philosophical Logic', Blackwell, Malden, MA and Oxford, pp. 415–436.

Wessel, H. (1998), *Logik*, 4th edn, Logis, Berlin.

Wojciechowski, E. (1997), 'External and internal negation in modal logic', *Conceptus* **30**(76), 57–66.

IGOR SEDLÁR
The Czech Academy of Sciences, Institute of Computer Science
sedlar@cs.cas.cz

KAREL ŠEBELA
Palacký University in Olomouc, Department of Philosophy
karel.sebela@upol.cz