Aristotelian Diagrams for Semantic and Syntactic Consequence

Lorenz Demey

Abstract

Various authors have recently studied Aristotelian diagrams for metatheoretical notions from logic. However, all these metalogical Aristotelian diagrams focus on the semantic (model-theoretical) perspective on logical consequence, thus ignoring the complementary, and equally important, syntactic (proof-theoretical) perspective. In this paper, I propose an explanation for this discrepancy, by arguing that the metalogical square of opposition for semantic consequence exhibits a natural analogy to the well-known square of opposition for the categorical statements from syllogistics, but that this analogy breaks down once we move from semantic to syntactic consequence. I then show that despite this difficulty, one can indeed construct metalogical Aristotelian diagrams from a syntactic perspective, which have their own, equally elegant interpretation in terms of the categorical statements. Finally, I construct a metalogical Aristotelian diagram that incorporates both semantic and syntactic consequence (and their interaction), and study how it is influenced by the underlying logical system's soundness and/or completeness.

1 Introduction

Aristotelian diagrams are visual representations of a set of propositions, concepts or expressions, and the logical relations holding between them. Without a doubt, the oldest and most widely used example is the so-called 'square of opposition', but there also exist many larger, more complex Aristotelian diagrams. Throughout the history of philosophy and logic, distinguished authors such as William of Ockham, Gottfried Leibniz and Gottlob Frege have made use of such diagrams to illustrate and explain their theorizing (Kienzler 2013, Lenzen 2016, Parsons 2017). In 20th century analytic philosophy, Aristotelian diagrams have been used in fields as diverse as philosophy of action (Kenny 1963), ethics (Chisholm 1963), philosophy of language (Hare 1967), philosophy of law (Hart 1982), epistemology (Sosa 1964) and philosophy of religion (Hess 2017). They are also widely used to study families of logics such as modal logic (Rini and Cresswell 2012), relevant logic (Beall et al. 2006), deontic logic (Barcan Marcus 1966) and probabilistic logic (Pfeifer and Sanfilippo 2017). Furthermore, because of the ubiquity of the logical relations that they visualize, Aristotelian diagrams are nowadays also frequently used beyond philosophy and logic, in disciplines such as computer science (Ciucci et al. 2016), law (Vranes 2006) and linguistics (Ziegeler 2017).

Most of the Aristotelian diagrams that have appeared in the logical literature thus far are situated at the *object-logical* level: the propositions appearing in these diagrams come from the object language of the logical system for which the diagram is drawn. However, in recent years there has been a surge of interest in Aristotelian diagrams at the *metalogical* level, which contain statements or notions belonging to the metatheory of a given logical system. For example, Béziau (2012, 2013) and Diaconescu (2015) have studied Aristotelian diagrams for metalogical notions such as tautology, contradiction, satisfiability, etc., and Demey (2017a) has argued that these diagrams can be fruitfully used in specific pedagogical contexts. Furthermore, Béziau (2012, 2016), Seuren (2014) and Demey and Smessaert (2016) have constructed diagrams that represent the Aristotelian relations holding between the Aristotelian relations themselves; Demey and Smessaert (2014) and Demey (2017b) have used these diagrams to study the specific language used in metalogic from the perspective of (neo-)Gricean pragmatics.

Despite this recent surge of interest, all metalogical Aristotelian diagrams studied thus far focus exclusively on the *semantic* (*model-theoretical*) perspective on logic and logical consequence (cf. Demey 2017a, Footnote 5). So far, no one seems to have studied metalogical Aristotelian diagrams for the complementary, and equally important, *syntactic (proof-theoretical)* perspective on logic and logical consequence.[1] Therefore, Aristotelian diagrams involving the central metalogical notions of soundness and completeness have not been studied either. (The soundness and completeness theorems establish a connection between syntactic and semantic

[1] Diaconescu (2015) studies several metalogical Aristotelian diagrams for abstract (Tarski-style) consequence relations, which transcend the distinction between syntax and semantics. However, he does not deal with any Aristotelian diagrams for *syntactic* consequence in particular.

logical consequence, so if one does not take the syntactic perspective into account to begin with, then, *a fortiori*, one cannot deal with soundness and completeness either.)

My main aim in this paper is to address this important lacuna in the literature on (metalogical) Aristotelian diagrams. This naturally leads to a sequence of interrelated subgoals. First of all, I will propose an explanation as to why all research in this area has thus far focused exclusively on the semantic perspective. In particular, I will argue that the metalogical square of opposition for semantic consequence exhibits a natural analogy to the oldest and most well-known square of opposition, viz. that for the categorical statements from Aristotelian syllogistics. By contrast, this analogy breaks down quite spectacularly once we move from semantic to syntactic consequence. The second subgoal is to show that despite this *prima facie* difficulty, one can indeed construct metalogical Aristotelian diagrams from a syntactic perspective. Furthermore, these new diagrams have their own, equally theoretically elegant interpretation in terms of the categorical statements. We thus obtain Aristotelian diagrams for semantic consequence on the one hand, and for syntactic consequence on the other. The paper's third and final subgoal is therefore to construct a metalogical Aristotelian diagram that incorporates both semantic and syntactic consequence (and their interaction), and to study how this diagram is influenced by the underlying logical system's soundness and/or completeness.

The paper is organized as follows. In Section 2, I briefly rehearse the most important metalogical Aristotelian diagram for semantic consequence that has been studied in the literature, and argue that it can naturally be characterized in terms of the categorical statements from syllogistics. Next, in Section 3, I show that one can construct an analogous metalogical Aristotelian diagram for syntactic consequence. I also consider the most straightforward way of interpreting this new diagram in terms of the categorical statements, and argue that this fails for various reasons. In Section 4, I then propose an alternative way of understanding the new diagram for syntactic consequence in terms of the categorical statements, and show that this new interpretation is indeed successful. I also argue that this alternative characterization of the diagram for syntactic consequence is in fact the closest possible analogue of the characterization of the diagram for semantic consequence. Building on these results, in Section 5, I construct a metalogical Aristotelian diagram that incorporates both semantic and syntactic consequence (and any potential interaction between them). I also study how the soundness and/or completeness of the underlying logical system influence the metalogical properties of this diagram. Finally, in Section 6, I briefly summarize the paper, and offer some suggestions for future work.

2 The Metalogical Square of Opposition for Semantic Consequence

In this section I will discuss the metalogical Aristotelian diagram for semantic consequence that has been used most widely in the literature. First, however, I will briefly make some basic remarks about the Aristotelian relations and Aristotelian diagrams in general.

The Aristotelian relations can be defined on various levels of generality and abstractness. The most general definition is formulated in terms of arbitrary Boolean algebras; this definition enables us to capture very precisely the similarities and differences between Aristotelian relations between object-logical notions and those between metalogical notions (Demey and Smessaert 2016, Demey 2017b). However, for our current purposes it will suffice to define the Aristotelian relations in the usual, more informal way. Two statements are said to be:

- *contradictory* iff they cannot be true together, and they cannot be false together;
- *contrary* iff they cannot be true together, but they can be false together;
- *subcontrary* iff they cannot be false together, but they can be true together;
- in *subalternation* iff the first statement entails the second one, but the second statement does not entail the first one.

An Aristotelian diagram visually represents a finite number of statements, together with the Aristotelian relations holding between those statements. The Aristotelian relations are visualized according to the code shown in Figure 1. The oldest example of an Aristotelian diagram is the square of opposition for the four categorical statements from syllogistics, which is also shown in Figure 1. In particular, the A-statement ‘all S are P’ and the O-statement ‘some S are not P’ are contradictory; the A-statement ‘all S are P’ and the E-statement ‘no S are P’ are contrary, and so on.[2] Note that we are implicitly assuming the principle of existential import – i.e., that there exists at least one S.[3] Without this assumption, all Aristotelian relations, except for the two contradiction relations (A/O and E/I), would fail to hold, and hence, the classical square of opposition for the categorical statements would turn into a 'degenerate square' (or 'X of opposition'; Béziau and Payette 2012, pp. 11-12), in which only the two contradiction relations are left. This constitutes a prime example of a fact that is well-known in logical geometry – i.e. the systematic study of Aristotelian diagrams –, viz. the fact that these diagrams are highly sensitive with respect to certain background assumptions (Demey 2015, Demey and Smessaert 2017).

<INCLUDE FIGURE 1 HERE>
<CAPTION: Figure 1: (left) visual code for representing the Aristotelian relations; (right) square of opposition for the categorical statements from syllogistics, under the assumption of existential import.>

We now turn to the metalogical square of opposition for semantic consequence. To this end, we consider a logical system S, with an object language $\mathcal{L}_S$, which is supposed to have a classical negation ($\neg$) and a model-theoretic semantics. This involves a collection $\mathcal{C}_S$ of models[4] and a binary relation $\vDash$ between the models in $\mathcal{C}_S$ and the formulas in $\mathcal{L}_S$, where $M \vDash \varphi$ means that the formula φ is true in the model M. We write $M \nvDash \varphi$ to abbreviate that it is not the case that $M \vDash \varphi$. The classicality of negation means that $M \vDash \neg\varphi$ iff $M \nvDash \varphi$, for every $M \in \mathcal{C}_S$. Furthermore, if Γ is a set of $\mathcal{L}_S$-formulas, we write $M \vDash \Gamma$ to abbreviate that $M \vDash \gamma$ for every $\gamma \in \Gamma$. We can now define the notions of semantic logical consequence and satisfiability in the logical system S. Given any set Γ of $\mathcal{L}_S$-formulas, and $\mathcal{L}_S$-formula φ, we say that:

- φ is a *semantic consequence* of Γ (notation: $\Gamma \vDash \varphi$) iff
  for all models $M \in \mathcal{C}_S$, it holds that if $M \vDash \Gamma$, then $M \vDash \varphi$;
- Γ is *satisfiable* iff there exists at least one model $M \in \mathcal{C}_S$ such that $M \vDash \Gamma$.

[2] Throughout this paper I will make use of the well-known mnemonic vowels (A, I, E, O) for the categorical statements. These are the first two vowels from each of the Latin verb forms 'affirmo' and 'nego'.

[3] The precise interpretation of the principle of existential import in syllogistics has been a matter of substantial scholarly debate (Parsons 2017), but this need not concern us here.

[4] In particular, these models can be first-order models, relational structures, etc. These details do not matter here.

Finally, we write $\Gamma \nvDash \varphi$ to abbreviate that it is not the case that $\Gamma \vDash \varphi$.

Béziau (2012, 2013) and Diaconescu (2015) show that if Γ is satisfiable, then one can construct a metalogical square of opposition for semantic consequence, as shown in Figure 2. For example, $\Gamma \vDash \varphi$ is contrary to $\Gamma \vDash \neg\varphi$, since these statements cannot be true together (if simultaneously $\Gamma \vDash \varphi$ and $\Gamma \vDash \neg\varphi$, then Γ would not be satisfiable), but they can be false together. Note that in order to establish any of the Aristotelian relations in this square of opposition (except for the two contradiction relations), we need to rely on the assumption that Γ is satisfiable. Without this assumption, the metalogical square for semantic consequence would thus not be a classical square of opposition, but rather a degenerate square. This can be seen as a metalogical manifestation of the sensitivity of Aristotelian diagrams with respect to certain background assumptions (cf. supra).

<INCLUDE FIGURE 2 HERE>
<CAPTION: Figure 2: metalogical square of opposition for semantic consequence, under the assumption that Γ is satisfiable.>

It has recently been pointed out that the definitions of the semantic consequence statements appearing in Figure 2 can all be interpreted as categorical statements (Demey 2017a). In particular, recall that the definition of $\Gamma \vDash \varphi$ is:

for all models $M \in \mathcal{C}_S$, it holds that if $M \vDash \Gamma$, then $M \vDash \varphi$.

This definition can be interpreted as a categorical A-statement: it is of the form 'all S are P', with the subject term S standing for 'being a model $M \in \mathcal{C}_S$ such that $M \vDash \Gamma$', and the predicate term P standing for 'being a model $M \in \mathcal{C}_S$ such that $M \vDash \varphi$'. Furthermore, consider the definition of $\Gamma \vDash \neg\varphi$:

for all models $M \in \mathcal{C}_S$, it holds that if $M \vDash \Gamma$, then $M \vDash \neg\varphi$.

Since $M \vDash \neg\varphi$ iff $M \nvDash \varphi$ (for every $M \in \mathcal{C}_S$), this can be reformulated as follows:

$$\text{for all models } M \in \mathcal{C}_S \text{, it holds that if } M \vDash \Gamma \text{, then } M \nvDash \varphi.$$

The definition of $\Gamma \vDash \neg\varphi$ can thus be interpreted as a categorical E-statement: it(s reformulated version) is of the form ‘all S are not P’ (i.e.: ‘no S or P’), with S and P exactly as above. Similarly, the definition of $\Gamma \nvDash \varphi$ is:

$$\text{there exists at least one model } M \in \mathcal{C}_S \text{ such that } M \vDash \Gamma \text{ and } M \nvDash \varphi.$$

This definition can be interpreted as a categorical O-statement: it is of the form ‘some S are not P’, with S and P as above. Finally, consider the definition of $\Gamma \nvDash \neg\varphi$:

$$\text{there exists at least one model } M \in \mathcal{C}_S \text{ such that } M \vDash \Gamma \text{ and } M \nvDash \neg\varphi.$$

Since $M \nvDash \neg\varphi$ iff $M \vDash \varphi$ (for every $M \in \mathcal{C}_S$), this can be reformulated as follows:

$$\text{there exists at least one model } M \in \mathcal{C}_S \text{ such that } M \vDash \Gamma \text{ and } M \vDash \varphi.$$

The definition of $\Gamma \nvDash \neg\varphi$ can thus be interpreted as a categorical I-statement: it(s reformulated version) is of the form ‘some S are P’, with S and P again as above.

The definitions of the four statements regarding semantic consequence can thus be interpreted as the four categorical statements. The subject and predicate term in these categorical statements involve the most essential semantic notion, viz. *truth (in a model)*. This interpretation establishes a direct link between the squares of opposition in Figures 1 and 2. For example, $\Gamma \vDash \varphi$ and $\Gamma \vDash \neg\varphi$ are contrary to each other in the square of opposition for semantic consequence in Figure 2, and the definitions of these two statements can be interpreted as A- and E-statements, respectively, which are contrary to each other in the square of opposition for the categorical statements in Figure 1. Similarly, $\Gamma \nvDash \varphi$ and $\Gamma \nvDash \neg\varphi$ are subcontrary to each other in the square in Figure 2, and the definitions of these two statements can be interpreted as O- and I-statements, respectively, which are subcontrary to each other in the square in Figure 1. Furthermore, this interpretation also shows that the assumption that Γ is satisfiable (i.e. there exists at least one model $M \in \mathcal{C}_S$ such that $M \vDash \Gamma$) is analogous to the assumption of existential import in syllogistics (i.e. there exists at least one S). Based on the definitions of its notions, the metalogical square of opposition for semantic consequence thus turns out to be perfectly analogous to the most widely known square of opposition in the literature, viz. that for the categorical statements from Aristotelian syllogistics.

Finally, it should be pointed out that this interpretation in terms of categorical statements crucially relies on the fact that we can shift negation between the object- and the metalogical level ($M \vDash \neg\varphi$ iff $M \nvDash \varphi$). This is essential for reformulating the definitions of $\Gamma \vDash \neg\varphi$ and $\Gamma \nvDash \neg\varphi$, which are statements about models that make the negated formula $\neg\varphi$ true or not, into statements about models that make the formula φ itself true or not. The latter can then be interpreted as E- and I-statements, on a par with the A- and O-statements that correspond to $\Gamma \vDash \varphi$ and $\Gamma \nvDash \varphi$, respectively.

3 The Metalogical Square of Opposition for Syntactic Consequence

I will now introduce a metalogical square of opposition for syntactic consequence. To this end, we assume that our logical system S also has a proof system $\mathcal{D}_S$, which allows us to construct derivations that lead from (sets of) $\mathcal{L}_S$-formulas to other $\mathcal{L}_S$-formulas.[5] We can now define the notions of syntactic logical consequence and consistency in S. Given any set Γ of $\mathcal{L}_S$-formulas, and $\mathcal{L}_S$-formula φ, we say that:

- φ is a *syntactic consequence* of Γ (notation: $\Gamma \vdash \varphi$) iff
  there exists at least one derivation in $\mathcal{D}_S$ that starts from formulas in Γ and ends in φ;
- Γ is *consistent* iff there does not exist a formula $\gamma \in \mathcal{L}_S$ such that $\Gamma \vdash \gamma$ and $\Gamma \vdash \neg\gamma$;
- Γ is *consistent with* φ iff the set $\Gamma \cup \{\varphi\}$ is consistent.

We will write $\Gamma \nvdash \varphi$ to abbreviate that it is not the case that $\Gamma \vdash \varphi$. Furthermore, we will assume that the proof system $\mathcal{D}_S$ has the property that if Γ is not consistent with φ, then $\Gamma \vdash \neg\varphi$; this will play an important role in (the proof of) the theorem in Section 4. Most classical proof systems indeed have this property (as well as its converse).

[5] In particular, the proof system can be an axiomatic proof system, a (Gentzen-style or Fitch-style) natural deduction system, etc. Again, these details do not matter here. (Also see Footnote 4.)

Under ideal circumstances, semantic and syntactic consequence should be closely related to each other – after all, both are meant to capture a single, informal notion of logical consequence. Given the square of opposition for semantic consequence from Section 2, one would therefore expect that there exists an analogous square of opposition for syntactic consequence, as shown in Figure 3. It is easy to prove that if Γ is consistent, then all the Aristotelian relations in Figure 3 indeed hold, i.e. we indeed obtain a square of opposition for syntactic consequence. For example, $\Gamma \vdash \varphi$ is contrary to $\Gamma \vdash \neg\varphi$, since these statements cannot be true together (if simultaneously $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$, then Γ would not be consistent), but they can be false together. Note that in order to establish any of the Aristotelian relations in this square of opposition (except for the two contradiction relations), we need to rely on the assumption that Γ is consistent. Without this assumption, the metalogical square for syntactic consequence would not be a classical square of opposition, but rather a degenerate square. This can be seen as yet another metalogical manifestation of the sensitivity of Aristotelian diagrams with respect to certain background assumptions (cf. supra).

<INCLUDE FIGURE 3 HERE>
<CAPTION: Figure 3: metalogical square of opposition for syntactic consequence, under the assumption that Γ is consistent.>

Let us now have a closer look at the definitions of the syntactic consequence statements that appear in Figure 3. First of all, recall that the definition of $\Gamma \vdash \varphi$ is:

there exists at least one derivation in $\mathcal{D}_S$ that starts from formulas in Γ and ends in φ.

Completely analogously, the definition of $\Gamma \vdash \neg\varphi$ is:

there exists at least one derivation in $\mathcal{D}_S$ that starts from formulas in Γ and ends in $\neg\varphi$.

Furthermore, here are three equivalent formulations of the definition of $\Gamma \nvdash \varphi$:

there exists no derivation in $\mathcal{D}_S$ that starts from formulas in Γ and ends in φ,
no derivation in $\mathcal{D}_S$ that starts from formulas in Γ, ends in φ,
all derivations in $\mathcal{D}_S$ that start from formulas in Γ, do not end in φ.

Finally, consider three equivalent formulations of the definition of $\Gamma \nvdash \neg\varphi$:

there exists no derivation in $\mathcal{D}_S$ that starts from formulas in Γ and ends in $\neg\varphi$,
no derivation in $\mathcal{D}_S$ that starts from formulas in Γ, ends in $\neg\varphi$,
all derivations in $\mathcal{D}_S$ that start from formulas in Γ, do not end in $\neg\varphi$.

If we try to interpret these definitions of the four syntactic consequence statements in terms of the four categorical statements from syllogistics, we immediately encounter several problems. First of all, there is a 'quantificational mismatch'. For example, since $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$ are contraries in the square of opposition for syntactic consequence in Figure 3, we should expect their definitions to correspond to the universally quantified A- and E-statements, which are contraries in the square of opposition for the categorical statements in Figure 1. However, we have just seen that the definitions of $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$ involve an *existential*, rather than a *universal* quantification ('there exists at least one derivation…'). Similarly, since $\Gamma \nvdash \varphi$ and $\Gamma \nvdash \neg\varphi$ are subcontraries in the square in Figure 3, we should expect their definitions to correspond to the existentially quantified O- and I-statements, which are subcontraries in the square in Figure 1. However, we have just seen that the definitions of $\Gamma \nvdash \varphi$ and $\Gamma \nvdash \neg\varphi$ involve a *universal*, rather than an *existential* quantification ('all derivations…').[6]

This quantificational mismatch can also be discerned in the assumption that is needed to prove that the square is indeed a classical square of opposition (rather than a degenerate square). In the case of the categorical statements in Figure 1, this is the assumption of *existential import*, which, as the name already suggests, is an *existential* claim ('there exists at least one S'). In the case of the semantic consequence statements in Figure 2, the assumption concerns the *satisfiability* of Γ, which is also an *existential* claim ('there exists at least one model M…'). By contrast, in the case of the syntactic consequence statements in Figure 3, the assumption

[6] The problematic nature of this quantifier mismatch should not be exaggerated. Ultimately, $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$ are effectively contrary to each other, regardless of whether their definitions can be interpreted as universal (A- and E-) statements. In the literature there exist other, object-logical examples of squares of opposition that exhibit a similar quantifier mismatch. For example, in the square of opposition for public announcement logic, the formulas $\langle !p\rangle q$ and $\langle !p\rangle\neg q$ are contrary to each other, although the semantics of these formulas involves an existential, rather than a universal quantification (over public announcements of p) (Demey 2012, 2017c).

concerns the *consistency* of Γ, which is a *universal*, rather than an existential claim ('there exists no formula γ…').

The final, and perhaps most serious problem, concerns the behavior of negation. Regardless of their quantificational status, the definitions of all four syntactic consequence statements share the same subject term, viz. 'being a derivation in $\mathcal{D}_S$ that starts from formulas in Γ'. However, they do *not* share the same predicate term: the predicate term in the definitions of $\Gamma \vdash \varphi$ and $\Gamma \nvdash \varphi$ is 'being a derivation in $\mathcal{D}_S$ that ends in φ', whereas the predicate term in the definitions of $\Gamma \vdash \neg\varphi$ and $\Gamma \nvdash \neg\varphi$ is 'being a derivation in $\mathcal{D}_S$ that ends in $\neg\varphi$'. The latter predicate term cannot be reduced to the former, since ending in $\neg\varphi$ is strictly stronger than not ending in φ. (If a derivation ends in $\neg\varphi$, it trivially does not end in φ; however, the converse does not hold: a derivation might end in some third formula ψ, in which case it does not in φ, and not in $\neg\varphi$ either.) This should be contrasted with the definitions of the semantic consequence statements in Section 2, which do share a single (subject term and) predicate term, since $M \vDash \neg\varphi$ is indeed equivalent to $M \nvDash \varphi$.

The overall argument that has been developed in this section and the previous one, can now be summarized as follows:

1. Both semantic and syntactic consequence give rise to a metalogical square of opposition (under analogous metalogical conditions, viz. the satisfiability and consistency of Γ).
2. The definitions of the four semantic consequence statements can straightforwardly be interpreted in terms of the four categorical statements from syllogistics. Hence, the square of opposition for semantic consequence in Figure 2 (and its assumption of the satisfiability of Γ) exhibits a strong analogy to the square of opposition for the categorical statements in Figure 1 (and its assumption of existential import).
3. By contrast, the definitions of the four syntactic consequence statements cannot be interpreted in terms of the four categorical statements. Hence, the square of opposition for syntactic consequence in Figure 3 (and its assumption of the consistency of Γ) does *not* exhibit a strong analogy to the square of opposition for the categorical statements in Figure 1 (and its assumption of existential import).

Taking into consideration that the square of opposition for the categorical statements from syllogistics is by far the oldest and most well-known Aristotelian diagram, and thus serves as a kind of 'golden standard', I would like to suggest that this discrepancy between the semantic and the syntactic square of opposition might be precisely the reason why the former has been studied quite frequently in the recent literature on metalogical Aristotelian diagrams, whereas the latter has not been studied at all thus far.

4 An Alternative Characterization of the Metalogical Square for Syntactic Consequence

In the previous section I have argued that although the syntactic consequence statements constitute a metalogical square of opposition, their definitions cannot be understood in terms of the categorical statements. In this section I will propose an alternative characterization of the syntactic consequence statements, and show that it does enable a direct analogy to the categorical statements.[7] We therefore consider the following:

Theorem. Given any set Γ of $\mathcal{L}_S$-formulas, and $\mathcal{L}_S$-formula φ, it holds that:

1) $\Gamma \vdash \varphi$	iff	all consistent $\Gamma' \supseteq \Gamma$ are consistent with φ;
2) $\Gamma \vdash \neg\varphi$	iff	no consistent $\Gamma' \supseteq \Gamma$ are consistent with φ;
3) $\Gamma \nvdash \varphi$	iff	some consistent $\Gamma' \supseteq \Gamma$ are not consistent with φ;
4) $\Gamma \nvdash \neg\varphi$	iff	some consistent $\Gamma' \supseteq \Gamma$ are consistent with φ;
5) Γ is consistent	iff	there exists at least one consistent $\Gamma' \supseteq \Gamma$.

Proof. Item 1, left to right. Consider an arbitrary consistent $\Gamma' \supseteq \Gamma$ and suppose, toward a contradiction, that Γ' is not consistent with φ. Hence $\Gamma' \vdash \neg\varphi$. Furthermore, since $\Gamma \vdash \varphi$ and $\Gamma \subseteq \Gamma'$, we also have $\Gamma \vdash \varphi$. This violates the consistency of Γ'.

Item 1, right to left is proved by contraposition, so we assume that $\Gamma \nvdash \varphi$, and show that there exists a consistent $\Gamma' \supseteq \Gamma$ that is not consistent with φ. Consider $\Gamma' \coloneqq \Gamma \cup \{\neg\varphi\}$. It trivially holds that $\Gamma' \supseteq \Gamma$; furthermore, Γ' is consistent because $\Gamma \nvdash \varphi$; finally, Γ' is not consistent with φ, since the set $\Gamma' \cup \{\varphi\} = \Gamma \cup \{\neg\varphi\} \cup \{\varphi\}$ is trivially not consistent.

[7] Thanks to an anonymous reviewer of Demey (2017a) for some useful discussion on this.

Item 2, left to right. Consider an arbitrary consistent $\Gamma' \supseteq \Gamma$; we show that Γ' is not consistent with φ. Since $\Gamma \vdash \neg\varphi$ and $\Gamma \subseteq \Gamma' \subseteq \Gamma' \cup \{\varphi\}$, it follows that $\Gamma' \cup \{\varphi\} \vdash \neg\varphi$. But trivially also $\Gamma' \cup \{\varphi\} \vdash \varphi$, and hence $\Gamma' \cup \{\varphi\}$ is not consistent, i.e. Γ' is not consistent with φ.

Item 2, right to left is again proved by contraposition, so we assume $\Gamma \nvdash \neg\varphi$, and show that there exists a consistent $\Gamma' \supseteq \Gamma$ that is consistent with φ, viz. Γ itself. After all, since $\Gamma \nvdash \neg\varphi$ it follows that Γ is consistent with φ, and thus, *a fortiori*, that Γ itself is consistent.

Items 3 and 4 follow from items 1 and 2, respectively (after all, if two statements A and B are equivalent to each other, then their negations ¬A and ¬B are also equivalent to each other).

Item 5, left to right. If Γ is consistent, there trivially exists at least one consistent $\Gamma' \supseteq \Gamma$, viz. Γ itself.

Item 5, right to left is proved by contraposition, so we assume that Γ is not consistent, and show that there does not exist a consistent $\Gamma' \supseteq \Gamma$. Since Γ is not consistent, there exists a formula γ such that $\Gamma \vdash \gamma$ and $\Gamma \vdash \neg\gamma$; hence, for every $\Gamma' \supseteq \Gamma$ it follows that also $\Gamma' \vdash \gamma$ and $\Gamma' \vdash \neg\gamma$, and thus Γ' is not consistent either.

QED

The first four items of the theorem provide characterizations of the syntactic consequence statements that correspond to the categorical statements, which all have the same subject term, viz. 'being a consistent set $\Gamma' \subseteq \mathcal{L}_S$ such that $\Gamma \subseteq \Gamma'$', as well as the same predicate term, viz. 'being a consistent set $\Gamma' \subseteq \mathcal{L}_S$ such that Γ' is consistent with φ'.

This correspondence does not suffer from a quantifier mismatch or problems regarding negation (cf. Section 3). For example, $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$ are contraries in the square of opposition for syntactic consequence in Figure 3, and the theorem's first two items show that these statements correspond to the universally quantified A- and E-statements, respectively, which are contraries in the square of opposition for the categorical statements in Figure 1. Similarly, $\Gamma \nvdash \varphi$ and $\Gamma \nvdash \neg\varphi$ are subcontraries in the square in Figure 3, and items 3 and 4 of the theorem show that these statements correspond to the existentially quantified O- and I-statements, respectively, which are subcontraries in the square in Figure 1. Furthermore, consider the assumption that is needed to prove that the square is indeed a classical square of opposition (rather than a degenerate square). In the case of the syntactic consequence statements, this is the assumption that Γ is consistent, and the theorem's fifth item provides a characterization of this assumption as an existential claim, which is completely analogous to the existential import assumption in the case of the categorical statements. Finally, note that there is no problem with negation, as all four characterizations share the same predicate term, viz. 'being a consistent set $\Gamma' \subseteq \mathcal{L}_S$ such that Γ' is consistent with φ', including the characterizations of the syntactic consequence statements that involve the negated formula $\neg\varphi$ (cf. items 2 and 4 of the theorem).

The theorem thus provides a successful characterization of the syntactic consequence statements in terms of the categorical statements. Hence, just like the square for semantic consequence in Figure 2, the square for syntactic consequence in Figure 3 also closely corresponds to the well-known square for the categorical statements in Figure 1. However, the former correspondence is based on the *definition* of semantic consequence (cf. Section 2), whereas the latter correspondence is not based on the *definition* of syntactic consequence (cf. Section 3), but rather on the *alternative characterization* provided by the theorem (cf. this section).[8] Nevertheless, in a sense, this alternative characterization of syntactic consequence can be understood as the closest possible analogue of the definition of semantic consequence. To see this, recall that the definition of semantic consequence involves categorical statements with subject term 'being a model $M \in \mathcal{C}_S$ such that $M \vDash \Gamma$' and predicate term 'being a model $M \in \mathcal{C}_S$ such that $M \vDash \varphi$', which are based on the core semantic notion of *truth (in a model)*. The theorem's alternative characterization of syntactic consequence involves categorical statements with subject term 'being a consistent set $\Gamma' \subseteq \mathcal{L}_S$ such that $\Gamma \subseteq \Gamma'$' and predicate term 'being a consistent set $\Gamma' \subseteq \mathcal{L}_S$ such that Γ' is consistent with φ', which are based on the corresponding syntactic notion of *consistency (with a consistent set of formulas)*.

5 The Role of Soundness and Completeness

[8] Furthermore, the theorem cannot be used to provide a *definition* of syntactic consequence, since that would clearly result in circularity: the theorem characterizes syntactic consequence in terms of consistency, but as we have seen in the beginning of Section 3, the latter notion is itself defined in terms of syntactic consequence.

In the previous sections I have focused exclusively on squares of opposition for semantic and syntactic consequence. However, in the literature we also find many larger, more complex metalogical Aristotelian diagrams (see especially Demey and Smessaert 2016 for a comprehensive overview). For example, Béziau (2012, 2013) and Diaconescu (2015) show that the square of opposition for semantic consequence (cf. Figure 2) can be extended to a so-called Jacoby-Sesmat-Blanché (JSB) hexagon of opposition, by adding two further semantic consequence statements, viz. the disjunction of $\Gamma \vDash \varphi$ and $\Gamma \vDash \neg\varphi$, and the conjunction of $\Gamma \nvDash \varphi$ and $\Gamma \nvDash \neg\varphi$. It is straightforward to show that the square of opposition for syntactic consequence (cf. Figure 3) can also be extended to a JSB hexagon, in a completely analogous fashion. However, in this section I want to focus on another type of metalogical Aristotelian diagram, in which the semantic and syntactic consequence statements appear simultaneously, and which is thus also able to capture their interaction.

Let us start by assuming that the logical system S is sound, but not necessarily complete – or more precisely: that the proof system $\mathcal{D}_S$ is sound, but not necessarily complete, with respect to the class of models $\mathcal{C}_S$. (One could also start by making no assumptions at all regarding the soundness/completeness of S, but this would not be very realistic, since soundness is usually taken to be a 'minimal criterium' that has to be met by any serious candidate proof system.) Based on this assumption (along with the assumptions regarding the satisfiability and consistency of Γ from Sections 2 and 3), one can construct an octagon of opposition for semantic and syntactic consequence, as shown in Figure 4. This type of Aristotelian diagram is sometimes called a 'Lenzen octagon', after Lenzen (2012).

<INCLUDE FIGURE 4 HERE>
<CAPTION: Figure 4: metalogical octagon of opposition for semantic and syntactic consequence, under the assumptions that Γ is satisfiable and consistent, and that S is sound.>

This octagon incorporates (a horizontally stretched version of) the square for semantic consequence, as well as (a vertically stretched version of) the square for syntactic consequence as subdiagrams. Furthermore, it also shows the Aristotelian relations capturing the *interaction* between these two squares. For example, there is a subalternation from $\Gamma \vdash \varphi$ to $\Gamma \vDash \varphi$: because of soundness, $\Gamma \vdash \varphi$ indeed entails $\Gamma \vDash \varphi$, but since the proof system is not assumed to be complete, the other direction does not generally hold. The same considerations also show that there is a subalternation from $\Gamma \nvDash \varphi$ to $\Gamma \nvdash \varphi$, that $\Gamma \vdash \varphi$ is contrary to $\Gamma \nvDash \varphi$, and that $\Gamma \vDash \varphi$ is subcontrary to $\Gamma \nvdash \varphi$. Furthermore, by composing the subalternation from $\Gamma \vdash \varphi$ to $\Gamma \vDash \varphi$ with the subalternation from $\Gamma \vDash \varphi$ to $\Gamma \nvDash \neg\varphi$ in the semantic square, we also obtain the subalternation from $\Gamma \vdash \varphi$ to $\Gamma \nvDash \neg\varphi$. All the Aristotelian relations shown in Figure 4 can straightforwardly be established in this manner.

The Lenzen octagon in Figure 4 can thus be partitioned into three parts, each of which depends on its own assumption: (i) the square of opposition for semantic consequence, which depends on the satisfiability of Γ, (ii) the square of opposition for syntactic consequence, which depends on the consistency of Γ, and (iii) the Aristotelian relations capturing the interaction between these two squares, which depend on the soundness of S.

Let's now examine what happens to this octagon if we assume that S is not only sound, but also complete (while still maintaining the assumptions regarding the satisfiability and consistency of Γ). Obviously, this additional assumption will not have an impact on the squares of opposition for semantic and syntactic consequence separately, but only on their interaction. However, it influences the Aristotelian relations capturing this interaction in a highly heterogeneous fashion. For example, the previous subalternation from $\Gamma \vdash \varphi$ to $\Gamma \vDash \varphi$ will turn into an equivalence: because we now assume soundness *and* completeness, these two statements entail each other. (The same consideration also shows that the previous contrariety between $\Gamma \vdash \varphi$ and $\Gamma \nvDash \varphi$ now turns into a contradiction.) By contrast, the previous subalternation from $\Gamma \vdash \varphi$ to $\Gamma \nvDash \neg\varphi$ remains a subalternation: even under the additional assumption of completeness, $\Gamma \nvDash \neg\varphi$ still does not entail $\Gamma \vdash \varphi$. (The same consideration also shows that the previous contrariety between $\Gamma \vdash \varphi$ and $\Gamma \vDash \neg\varphi$ remains a contrariety.)

All these changes can be summarized as follows: under the additional assumption of the completeness of S, the Lenzen octagon from Figure 4 'collapses' into the square of opposition shown in Figure 5. Note that each vertex of this square is occupied by two metalogical statements that are equivalent to each other (under the assumption of soundness and completeness, of course). Each Aristotelian relation between two vertices holds between either formula in one vertex and either formula in the other. Essentially, the square of opposition in Figure 5 can be seen as the result of laying the squares for semantic consequence (Figure 2) and for syntactic consequence (Figure 3) on top of each other.

<INCLUDE FIGURE 5 HERE>

<CAPTION: Figure 5: metalogical octagon of opposition for semantic and syntactic consequence, under the assumptions that Γ is satisfiable and consistent, and that S is sound and complete.>

In Sections 2 and 3, I already pointed out that Aristotelian diagrams are sensitive with respect to certain background assumptions: adding or dropping the assumption that Γ is satisfiable/consistent makes the Aristotelian diagrams for semantic/syntactic consequence switch between two types of squares (viz. a classical square of opposition vs. a degenerate square). In this section, however, we have encountered a much more radical manifestation of this general phenomenon: adding or dropping the assumption that S is complete makes the Aristotelian diagram for the interaction between semantic and syntactic consequence switch between two very different types of diagrams (viz. a Lenzen octagon vs. a classical square of opposition).

6 Conclusion

In this paper I have shown that syntactic logical consequence gives rise to a metalogical square of opposition, just like semantic logical consequence. Unlike the definitions of the semantic consequence statements, the definitions of the syntactic consequence statements cannot be interpreted in terms of the categorical statements from Aristotelian syllogistics. Nevertheless, the theorem proved in Section 4 provides an alternative characterization of syntactic consequence, thereby yielding another, equally elegant interpretation in terms of the categorical statements. I have also constructed a metalogical Aristotelian diagram that incorporates both semantic and syntactic consequence, as well as the interaction between them (as determined by the soundness and/or completeness of the underlying logical system). Along the way, I have pointed out several metalogical manifestations of a broader phenomenon that is well-known in logical geometry, viz. the idea that Aristotelian diagrams can be sensitive with respect to certain background assumptions. We encountered relatively mild cases of this phenomenon (classical square of opposition vs. degenerate square), but also a more radical case (Lenzen octagon vs. classical square of opposition).

As was already mentioned in the introduction, some of the metalogical Aristotelian diagrams for semantic consequence have been studied with specific pedagogical purposes in mind, viz. teaching (the semantic side of) metalogic to certain groups of students, who do not have much previous experience in formal logic, but are thoroughly familiar with Aristotelian syllogistics. In future research I will explore whether the new metalogical diagrams that have been developed in this paper (cf. Figures 3, 4 and 5) can also be put to similar pedagogical use, viz. to teach about the syntactic side of metalogic, and about soundness and completeness. The elegant characterization of syntactic consequence in terms of the categorical statements (cf. the theorem proved in Section 4) will probably turn out to be of crucial importance in this respect.

Acknowledgements

Thanks to Hans Smessaert and Margaux Smets for their valuable feedback on an earlier version of this paper. The author holds a Postdoctoral Fellowship of the Research Foundation–Flanders (FWO). The research reported in this paper was partially carried out during a research stay at the Institut für Philosophie II of the Ruhr-Universität Bochum, which was financially supported by an FWO travel grant.

Bibliography

R. Barcan Marcus. Iterated Deontic Modalities. *Mind* 75 (1966), 580 – 582.

J. C. Beall, R. T. Brady, A. P. Hazen, G. Priest, and G. Restall. Relevant Restricted Quantification. *Journal of Philosophical Logic* 35 (2006), 587 – 598.

J.-Y. Béziau. The Power of the Hexagon. *Logica Universalis* 6 (2012), 1 – 43.

J.-Y. Béziau. The Metalogical Hexagon of Opposition. *Argumentos* 5 (2013), 111 – 122.

J.-Y. Béziau. Disentangling Contradiction from Contrariety via Incompatibility. *Logica Universalis* 10 (2016), 157 – 170.

J.-Y. Béziau and G. Payette. Preface. In: J.-Y. Béziau and G. Payette (eds.), *The Square of Opposition. A General Framework for Cognition*, Peter Lang, 2012, p. 9 – 22.

R. Chisholm. Supererogation and Offence: A Conceptual Scheme for Ethics. *Ratio* 5 (1963), 1 – 14.

D. Ciucci, D. Dubois, and H. Prade. Structures of Opposition Induced by Relations. The Boolean and the Gradual Cases. *Annals of Mathematics and Artificial Intelligence* 76 (2016), 351 – 373.

L. Demey. Structures of Oppositions in Public Announcement Logic. In: J.-Y. Béziau and D. Jacquette (eds.), *Around and Beyond the Square of Opposition*, Springer, 2012, p. 313 – 339.

L. Demey. Interactively Illustrating the Context-Sensitivity of Aristotelian Diagrams. In: H. Christiansen, I. Stojanovic, and G. Papadopoulous (eds.), *Modeling and Using Context*, Lecture Notes in Computer Science 9405, Springer, 2015, p. 331 – 345.

L. Demey. Using Syllogistics to Teach Metalogic. *Metaphilosophy* 48 (2017a), 575 – 590.

L. Demey. Metalogic, Metalanguage, and Logical Geometry. Submitted (2017b).

L. Demey. The Logical Geometry of Russell's Theory of Definite Descriptions. Submitted (2017c).

L. Demey and H. Smessaert. Logische geometrie en pragmatiek. In: F. Van de Velde, H. Smessaert, F. Van Eynde, and S. Verbrugge (eds.), *Patroon en argument*, Leuven University Press, 2014, p. 553 – 564.

L. Demey and H. Smessaert. Metalogical Decorations of Logical Diagrams. *Logica Universalis* 10 (2016), 233 – 292.

L. Demey and H. Smessaert. Combinatorial Bitstring Semantics for Arbitrary Logical Fragments. Forthcoming in *Journal of Philosophical Logic* (2017).

R. Diaconescu. The Algebra of Opposition (and Universal Logic Interpretations). In: A. Koslow and A. Buchsbaum (eds.), *The Road to Universal Logic*, Springer, 2015, p. 127 – 143.

R. M. Hare. Some Alleged Differences between Imperatives and Indicatives. *Mind* 76 (1967), 309 – 326.

H. L. A. Hart. *Essays on Bentham. Jurisprudence and Political Theory*. Clarendon Press, 1982.

W. Kienzler. The Logical Square and the Table of Oppositions. Five Puzzles about the Traditional Square of Opposition Solved by Taking Up a Hint from Frege. *Logical Analysis and History of Philosophy* 15 (2013), 398 – 413.

W. Lenzen. How to Square Knowledge and Belief. In: J.-Y. Béziau and D. Jacquette (eds.), *Around and Beyond the Square of Opposition*, Springer, 2012, p. 305 – 311.

W. Lenzen. Leibniz's Logic and the "Cube of Opposition". *Logica Universalis* 10 (2016), 171 – 189.

T. Parsons. The Traditional Square of Opposition. In: E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy (Summer 2017 Edition)*, 2017.

N. Pfeifer and G. Sanfilippo. Probabilistic Squares and Hexagons of Opposition under Coherence. *International Journal of Approximate Reasoning* 88 (2017), 282 – 294.

A. Rini and M. Cresswell. *The World-Time Parallel*. *Tense and Modality in Logic and Metaphysics*. Cambridge University Press, 2012.

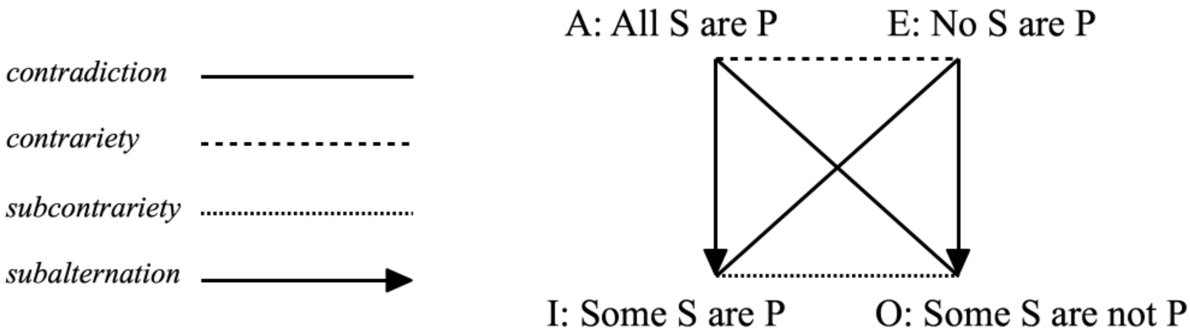
P. Seuren. Metalogical Hexagons in Natural Logic. Unpublished manuscript (2014).

E. Sosa. The Analysis of 'Knowledge that P'. *Analysis* 25 (1964), 1 – 8.

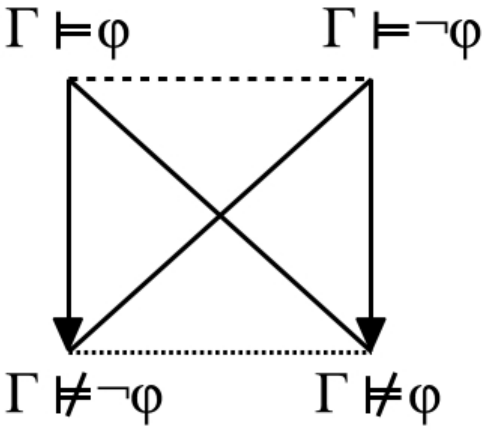
E. Vranes. The Definition of 'Norm Conflict' in International Law and Legal Theory. *European Journal of International Law* 17 (2006), 395 – 418.

D. Ziegeler. On the Empty O-Corner of the Aristotelian Square: A View from Singapore English. *Journal of Pragmatics* 115 (2017), 1 – 20.

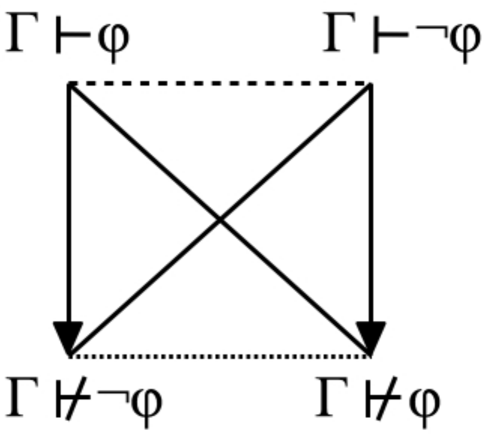
<FIGURE 1>



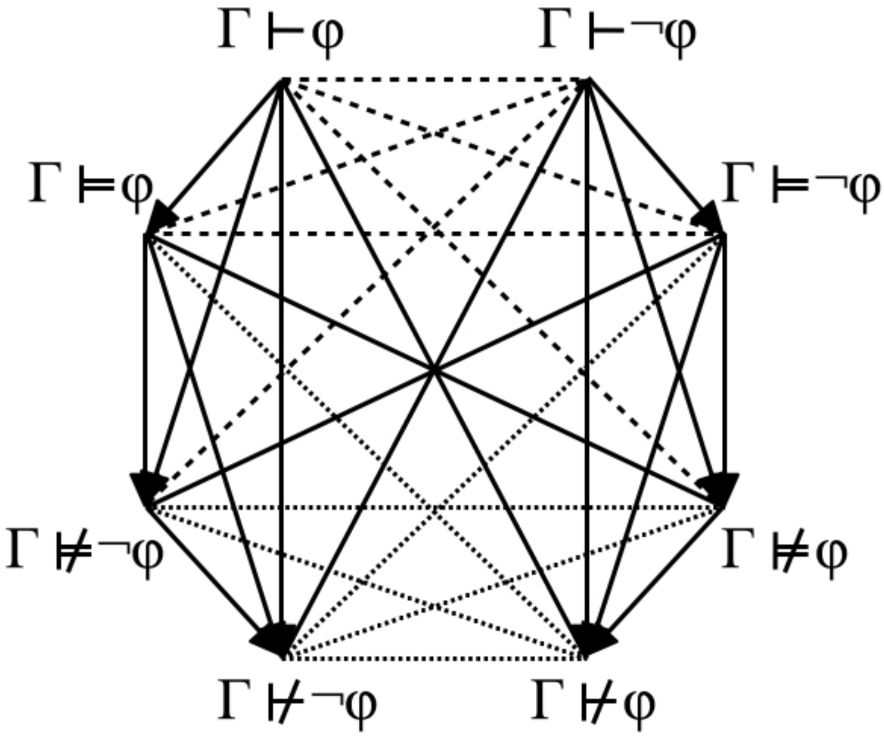
<FIGURE 2>



<FIGURE 3>



<FIGURE 4>



<FIGURE 5>