A causal analysis of modal syllogisms

van Rooij, R.; Xie, K.

DOI
10.1007/978-3-662-62843-0_10

Publication date
2020

Document Version
Author accepted manuscript

Published in
Monotonicity in Logic and Language

Citation for published version (APA):
van Rooij, R., & Xie, K. (2020). A causal analysis of modal syllogisms. In D. Deng, F. Liu, M. Liu, & D. Westerståhl (Eds.), *Monotonicity in Logic and Language: Second Tsinghua Interdisciplinary Workshop on Logic, Language and Meaning, TLLM 2020, Beijing, China, December 17-20, 2020 : proceedings* (pp. 183-206). (Lecture Notes in Computer Science; Vol. 12564), (FoLLI Publications on Logic, Language and Information). Springer. https://doi.org/10.1007/978-3-662-62843-0_10

A causal analysis of modal syllogisms

Robert van Rooij   Kaibo Xie

Abstract

It is well known that in his *Prior Analysis*, Aristotle presents the system of syllogisms. Although many commentators consider Aristotle's system of modal syllogisms almost impossible to understand from a modern point of view or even inconsistent, many philosophers still tried to account for these claims by looking for a consistent semantics of it. In this paper we will argue for a causal analysis of modal categorical sentences based on the notion of *causal power*. According to Cheng (1997), the causal power of A to produce B can be measured probabilistically. Based on Cheng's hypothesis, we will derive a qualitative semantics for modal categorical sentences. We will argue that our approach fits well with Aristotle's analysis of real definition in the *Posterior Analytics*, and that in this way we can account in a relatively straightforward way (using just Venn diagrams) for several puzzling aspects of Aristotle's system of modal syllogisms.

1 Introduction

In his *Prior Analytics* Aritstotle made a distinction between assertoric and modal syllogistics. The crucial difference between the two syllogistics is that only the latter makes use of two different types of predicative relations: accidental versus essential predication. 'Animal' is essentially predicated of 'men', but 'walking' is not. Although both (a) 'Every man walks' and (b) 'Every man is an animal' can be true, it is natural to say that the 'reasons' for their respective truths are different. Sentence (a) is true by accident, just because every actual man happens to (be able to) walk. The sentence (b), on the other hand, is true because manhood necessarily involves being animate. In traditional terms it is said that (b) is true *by definition*, although this notion of 'definition' should not be thought of nominalistically: it is the *real* definition. A natural way to account for accidental predication is to say that a sentence of the form 'Every S is P' is true just in case every actual S-*individual* is also a P-*individual*. But how should we account for essential predication? The answer to this question is important for logic, because it is by now generally assumed (e.g. Malink, 2013; van Rijen, 1989; Thom, 1991; Vecchio, 2016) that Aristotle's system of modal syllogisms, which is almost impossible to understand from a modern standard modal logic point of view, should be understood in terms of the difference between accidental and essential predication.

In this paper we will argue for a *causal* analysis of essential predication. We will argue that this fits well with Aristotle's analysis of real definition in the *Posterior Analytics*, and that in this way we can account in a relatively straightforward way for several puzzling aspects of Aritotle's system of modal syllogisms presented in his *Prior Analytics*.

2 Standard and Modal Syllogistics

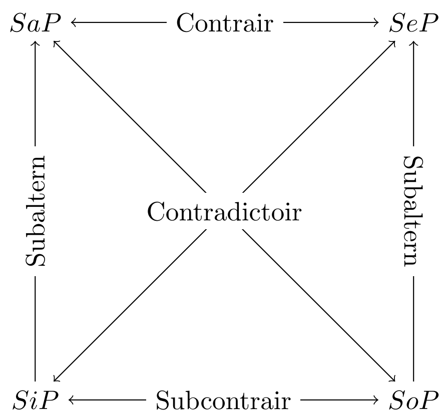
Syllogisms are arguments in which a categorical sentence is derived as conclusion from two categorical sentences as premisses. A categorical sentence is always one of four kinds:

1. *a*-type: Universal and affirmative ('All men are mortal')

2. *i*-type: Particular and affirmative ('Some men are philosophers')

3. *e*-type: Universal and negative ('No philosophers are rich')

4. *o*-type: Particular and negative ('Some men are not philosophers').



A categorical sentence always contains two *terms*. In the *a*-sentence, for instance, the terms are 'men' and 'mortal', while in the *e*-sentence they are 'philosopher' and 'rich'. Thus, the *syntax* of categorical sentences can be formulated as follows: If S and P are terms, SaP, SiP, SeP, and SoP are categorical sentences. Because a syllogism has two categorical sentences as premisses and one as the conclusion, every syllogism involves only three terms, each of which appears in two of the statements. The first term of the conclusion is called the *subject term*, or *minor term*, the last term, the *predicate term*, or *major term*, and the term that does not occur in the conclusion is called the *middle term*. The premiss in which the major term occurs together with the middle term is called the *major premiss*, the other one the *minor premiss*. The *quality* of a proposition is whether it is *affirmative* (in *a*- and *i*- sentences, the predicate is affirmed of the subject), or *negative* (in *e* and *o*-sentences, the predicate is denied of the subject). Thus 'every man is mortal' is affirmative, since 'mortal' is affirmed of 'man'. 'No men is immortal' is negative, since 'immortal' is denied of 'man'. The *quantity* of a proposition is whether it is *universal* (in *a*- and *e*-sentences the predicate is affirmed or denied of "the whole" of the subject) or *particular* (in *i* and *o*-sentences, the predicate is affirmed or denied of only 'part of' the subject).

Medieval logicians used the letters '*a*', '*i*', '*e*', and '*o*' for coding the various forms of syllogisms. The *mood* of a syllogism was given by a triple of letters like *aeo*. This triple, for instance, indicates that the major premiss is of type *a*, the minor premiss of type *e*, and the conclusion of type *o*. But apart from the mood, what is important as well is the *figure*. The figure of a syllogism says whether the major and minor terms occur as subject or predicate in their respective premisses. This gives rise to four possibilities, i.e., four figures:

1st	2nd	3rd	4th
MP	PM	MP	PM
SM	SM	MS	MS
SP	SP	SP	SP

A *valid* syllogism is a syllogism that cannot lead from true premisses to a false conclusion. It is well-known that by a set theoretic semantic analysis, we can account for syllogistic reasoning. For now we will interpret terms just as sets of individuals and equate for simplicity the interpretation of a term with the term itself. Then we say that SaP is true iff $S \subseteq P$, SiP is true iff $S \cap P \neq \emptyset$, SeP is true iff $S \cap P = \emptyset$, and SoP is true iff $S \not\subseteq P$.[1]

This semantic interpretation accounts for many valid syllogisms, but not all of them. In particular, not for the valid syllogisms for which it is required that SaP entails SiP. This can be easily accounted for by assuming that for the truth of SaP it is not only required that $S \subseteq P$, but also that $S \neq \emptyset$. It is well-known that with such an interpretation of categorical sentences, all and only all of the following syllogisms are predicted to be valid that Aristotle considered to be valid as well.

Barbara$_1$	Baroco$_2$	Bocardo$_3$	Camenes$_4$
Celarent$_1$	Festino$_2$	Disamis$_3$	(Fesapo$_4$)
Darii$_1$	Camestres$_2$	Ferison$_3$	Dimaris$_4$
Ferio$_1$	Cesare$_2$	Datisi$_3$	Fresison$_4$
(Barbari$_1$)	(Camestrop$_2$)	(Felapton$_3$)	(Bramantip$_4$)
(Celaront$_1$)	(Cesaro$_2$)	(Darapti$_3$)	(Camenop$_4$)

The syllogisms between brackets are only valid in case one assumes existential import, meaning that the extension of the subject term is non-empty. The above semantic analysis of categorical sentences is nice, because with the help of Venn-diagrams, one can now easily check the validity of any syllogistic argument.[2] For later in the paper, note that we could interpret Aristotle's standard categorical sentences probabilistically as well with equivalent predictions: SaP is true iff the conditional probability of P given S is 1, $P(P|S) = 1$, SeP is true iff $P(S \cap P) = 0$, SiP is true iff $P(S \cap P) \neq 0$ and SoP is true iff $P(P|S) \neq 1$. Notice that on this probabilistic interpretation SaP presupposes that $P(S) > 1$, which immediately accounts for Aristotle's subalternation inference: $SaP \models SiP$. This alternative semantics is the one we are going to use in our analysis of modal syllogisms. Therefore, we provide the following definition:

Definition 1 Truth conditions of Categorical sentences

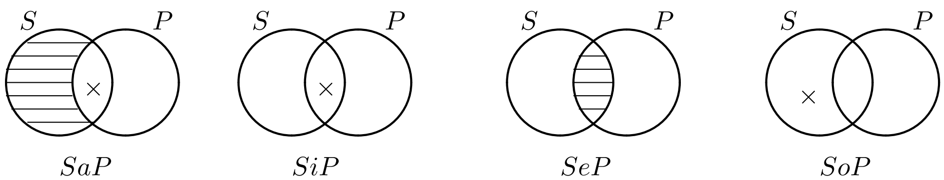
- *SaP is true iff $P(P|S) = 1$,*
- *SiP is true iff $P(S \cap P) \neq 0$,,*
- *SeP is true iff $P(S \cap P) = 0$,, and*

[1] Warning: in the literature categorical sentences of the form XaY and XiY are read many times in the converse order as we read them and mean that all/some Y belong to X.

[2] On the other hand, it is well-known that we don't need the full power of Boolean algebra to account for Syllogistic validity; semi-lattices will do.

- *SoP is true iff* $P(P|S) \neq 1$

With this definition we can give the truth definitions of categorical sentences with the following Venn diagrams (where an area has a cross when we know that it has at least one element, an area is shaded when we know it has no element, and if the area is empty we don't know whether the area has elements or not).



It is well-known that by drawing Venn diagrams one can give a *decision procedure* to determine which syllogisms are valid. Medieval logicians didn't make use of Venn diagrams, but developed another decision procedure to determine which syllogisms are valid. This procedure made crucial use of the so-called distribution-value of the terms involved. Whether a term is distributed or not is really a *semantic* question: a term is said to be distributed when it is actually applied to *all* the objects it can refer to, and undistributed when it is explicitly applied to only part of the objects to which it can refer. This formulation has been criticised by Geach (1962) and other modern logicians, but as noted by van Benthem (1983) and van Eijck (1985), it can be redefined in terms of **monotonicity**. A term occurs distributively when it occurs monotone decreasingly/negatively within a sentence, and undistributively when it occurs monotone increasingly/positively within a sentence. Denoting a distributed term by $-$ and an undistributed term by $+$, the following follows at once: $S^{-}aP^{+}$, $S^{+}iP^{+}$, $S^{-}eP^{-}$, and $S^{+}oP^{-}$, which we might think of now as a *syntactic* characterisation. In terms of the distribution values of terms, we can now state the laws of *quantity* or *distribution*, (R1) and (R2), and of *quality*, (R3). Together, they constitute the rules of the syllogism:[3]

(R1) The middle term must be distributed at least once.

(R2) Every term that is distributed in the conclusion is also distributed in one of the premises.

(R3) The number of negative conclusions must equal the number of negative premises.

The above rules assume existential import. Without this assumption, we have to strengthen (R2) to (R2'):

(R2') Every term that is (un)distributed in the conclusion is (un)distributed in one of the premises.

Medieval logicians and their followers standardly assumed that of all the reasoning schemas stated in syllogistic style, all and only all forms are valid that satisfy those

[3] Standardly, more rules are stated, but these can be derived from the rules below. One of the rules normally assumed, for instance, is that at least one of the premisses must be affirmative. But this follows immediately from (R3).

roles. As far as we know, the first one who explicitly *proved* this was Leibniz (1966).

Let us now come back to the question what is the natural interpretation of Aristotle's *modal* syllogistics. Let us assume that $Ba^{\Box}C$ means that all Bs are *necessary*/essentially C. Aristotle claims that the following modal syllogisms are valid and invalid, respectively:

1. $Ba^{\Box}C, Aa^{\Box}B \therefore Aa^{\Box}C$ Valid Barbara LLL
2. $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$ Valid Barbara LXL
3. $BaC, Aa^{\Box}B \therefore Aa^{\Box}C$ Invalid Barbara XLL

Although Aristotle had intuitions about which modal syllogistic inferences are valid and which not, he did not base that on a standard semantics. As it turns out, it is already hard enough to account semantically for the intuitions concerning 1-3. But what makes the task especially challenging is that Aristotle also claims that not only conversion inference 4 is valid, but that the same holds for the modal conversion inferences 5 and 6:

4. $BeC \therefore CeB$ Valid
5. $Be^{\Box}C \therefore Ce^{\Box}B$ Valid
6. $Bi^{\Box}C \therefore Ci^{\Box}B$ Valid

Of course, it is easy to account for inferences 5 and 6 if we assume that the modal should be interpreted in a *de dicto* way. But it is equally easy to see that on such an analysis inference 2 is *not* predicted to be valid. A *de re* analysis of sentences like $Ba^{\Box}C$, on the other hand, would make inference 2 valid, but such an analysis cannot account for the modal conversion inferences 5 and 6. So neither a standard *de dicto* nor a standard *de re* analysis of modal statements would work to account for Aristotle's intuitions.

Some commentators (e.g. Lukasiewicz, 1967; Patzig, 1968; Hintikka, 1973) concluded that the combination of these statements just doesn't make any sense and that Aristotle must have been confused. Others, however, tried to account for these claims by looking for a consistent semantics of Aristotle's system (e.g. Thomason, 1993; Uckelman & Johnston, 2010). The most interesting of these latter accounts build on the idea that Aristotle's modal syllogistics was based on his metaphysics and philosophy of science (e.g. Rescher, 1964; van Rijen, 1989; Patterson, 1995; Malink, 2013, Vecchio, 2016).[4] Unfortunately, most of these authors have difficulty making many predictions of valid modal syllogistic reasoning that correspond with Artstotle's intuitions. Recently, however, Malink (2013) has shown that it is actually possible to come up with a systematic analysis of modal syllogistic sentences such that it gives rise to predictions almost exactly in accordance with Aristotle's claims.[5] As we will see in section 5, however, on his analysis the validity of Barbara LXL, for example, is reduced to the validity of Barbara LLL, which we think is unexpected. One wonders

[4]Some (van Rijen (1989)) have claimed that $Ba^{\Box}C$ can hold only if 'B' is a substance term. This won't quite be enough (cf. Rini, 1998). Malink (2013) demands on top that a substance term can only be predicated of another substance term. We take this to follow naturally from a causal view.

[5]Vecchio (2016), building on Malink (2013), even slighly improves on Malink's predictions.

whether another analysis is not possible that interprets the second premiss of the argument not as a necessity statement. We think such an analysis possible, if we make use of a *causal* analysis of modal categorical statements.

In this paper we will argue for a *causal* analysis of Aristotle's modal claims. We will argue that this fits well with Aristotle's analysis of demonstrative inferences in the *Posterior Analytics*, and that in this way we can account in a relatively straightforward way for several puzzling aspects of Aritotle's system of modal syllogisms presented in his Prior Analytics. Although we don't see how something like the medieval distribution theory that is just based on monotonicity can be used as a decision procedure to check whether modal syllogisms are valid, to our surprise Venn diagrams can be used for this purpose, or at least for the fragment of Apodeictic syllogisms. In fact, we will see that just making use of the distribution rules, which can be thought of as a monotonicity calculus, cannot work on our causal analysis, because the rule of *right upward monotonicity* won't be valid anymore. In fact, we take this as a crucial insight behind the above problem of the three Barbara's.

3 Causal analysis and Aristotelian demonstrations

3.1 Causal dependence and causal models

Consider the following two sentences:

(1) a. Aspirin causes headaches to diminish.
    b. Aspirin relieves headaches.

Intuitively, (1-a) says that there exists a causal connexion between Aspirin and diminishing headaches: the intake of Aspirin *tends to* diminish headaches. Remarkably, (1-a) seems to express the same content as the *generic* sentence (1-b). This strongly suggests that also the generic sentence (1-b) should be given a causal analysis. Thus, not only (1-a), but also (1-b) expresses the fact that particular intakes of Aspirin *tend* to cause particular states of headache to go away, because of *what it is* to be Aspirin. Or, as we will say, because of the *causal power* of Aspirin to relieve headaches.

Causality is a kind of dependence. A number of authors have recently argued for a dependency analysis of conditionals, which is most straightforwardly done using probabilities: C depends on A iff $P(C|A) > P(C)$.[6] However, Douven (2008) has argued that dependence is not enough, 'If A, then C' is acceptable only if both $P(C|A) > P(C)$ *and* $P(C|A)$ are high.

We can implement Douven's proposal by requiring that $P(C|A) - P(C|\neg A)$ is close to $1 - P(C|\neg A)$. Since $P(C|A) > P(C)$ iff $P(C|A) > P(C|\neg A)$, we can demand that the conditional is acceptable iff $\frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$ is high. This can only be the case if both $P(C|A) - P(C|\neg A)$ and $P(C|A)$ are high, so it derives Douven's demands.

The measure $\frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$ is interesting from a causal perspective. Especially among philosophers dissatisfied with a Humean metaphysics, **causal powers** have recently become en vogue (again). Indeed, a growing number of philosophers (Harré & Madden, 1975; Cartwright, 1989; Shoemaker 1980; Bird 2007) have argued that causal powers, capacities or dispositions are the truth-makers of laws and other non-accidental generalities. Cheng (1997) hypothesises the existence of stable, but unobservable causal powers (Pearl (2000) calls them 'causal *mechanisms*') p_{ac} of (objects

[6]For a discussion of some qualitative variants, see Spohn (2013) and Rott (2019).

or events of kind) A to produce C. Cheng then *derives* a way how this objective but unobservable power can be estimated by an observable quantity, making use of standard probability theory and assuming certain natural independence conditions. It turns out that this quantity is exactly the above measure: $p_{ac} = \frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$. Cheng's notion has been used for the analysis of conditionals, generics and disposition statements, in van Rooij and K. Schulz (2019, 2020).

Dispositions and causal powers are things that (kinds of) objects have, independently of whether they show them. It is standardly assumed, though, that these (kinds of) objects *would* show them, if they *were* triggered sufficiently. Thus, there should be a relation with counterfactuals. Pearl (2000) provides a causal analysis of counterfactuals. He defines he 'probability of causal sufficiency of A to produce C', abbreviated by PS^C_A, as $P(C_A \mid \neg C, \neg A) = \frac{P(C_A, \neg C, \neg A)}{P(\neg C, \neg A)}$, with C_A the property that is true of an object if after making the object an A-object by intervention, the object would be a C-object.

Pearl (2000, chapter 9) shows that under natural conditions PA^C_A reduces to Cheng's notion of causal power, $\frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$. The first of these natural conditions is a **consistency assumption** used for counterfactuals,

(i) $A \Rightarrow (C_A = C)$.

This assumption is natural: if A already holds, an intervention to make A true leaves everything as is.[7] Pearl also assumes a notion of *exogeneity*, i.e., that C_A is *independent* of learning A (and thus also that $\neg C_{\neg A}$ is *independent* of $\neg A$).

(ii) A variable A is said to be **exogenous** relative C in model M iff $P(C_A \wedge C_{\neg A}|A) = P(C_A \wedge C_{\neg A})$.

Pearl's assumption that A is exogenous to C is very similar to Cheng's (1997) assumption that the potential causes of C are *independent* of one another (the Noisy-OR assumption). It rules out that learning A influences the probability of C via an indirect way, for instance that if B is another potential cause of C, there is a common cause of A and B.

Making use of these two assumptions, Pearl (2000) shows that $PS^C_A = \frac{P(C_A \wedge \neg C_{\neg A})}{1-P(C|\neg A)}$. On the additional assumption of *monotonicity*,

(iii) C is **monotonic** relative to A iff for all u: $C_A(u) \geq C_{\neg A}(u)$,

Pearl derives that

(2) $PS^C_A \quad = \quad \frac{P(C|A)-P(C|\neg A)}{1-P(C/\neg A)}$.

Thus, PS^C_A can be thought of as the causal power of A to produce C, i.e., p_{ac}. Notice that if all involved causal powers have value 1, a sequence of such causal powers is **transitively closed**: if $PS^B_A = 1$ and $PS^C_B = 1$, then also $PS^C_A = 1$. Obviously, also $PS^A_A = 1$, meaning that causal power is **reflexive**, and that demanding PS to be 1 gives rise to a pre-order.

[7] If we would analyse the counterfactual $A \Box\!\!\rightarrow C$ by C_A, this consistency rule would validate *modus ponens* and the inference $A, C \therefore A \Box\!\!\rightarrow C$, also known as conjunctive sufficiency. Both inference rules are accepted by almost everyone working on counterfactuals, although, to be honest, not by everyone.

In this paper we are going to make crucial of of the following interesting about the probabilistic measures $\frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$ and .[8]

Fact 1 $\frac{P(C|A)-P(C|\neg A)}{1-P(C|\neg A)}$ *has its* <u>*maximal value 1*</u> *iff* $P(C|A) = 1$ *and* $P(C|\neg A) \neq 1$.

Similarly, we predict that $PS_A^{\neg C} = 1$ and $p_{a\neg c} = 1$ holds only if $P(C|A) = 0$ and $P(C|\neg A) \neq 0$. This is due to the following fact

Fact 2 $\frac{P(\neg C|A)-P(\neg C|\neg A)}{1-P(\neg C|\neg A)}$ *is equal to* $\frac{P(C|\neg A)-P(C|A)}{P(C|\neg A)}$ *and has its* <u>*maximal value 1*</u> *just in case* $P(C|A) = 0$ *and* $P(C|\neg A) \neq 0$.

Interestingly, $p_{a\neg c}$ corresponds with Cheng's (1997) notion of *causal power* of A to *prevent* C. We propose that these notions might help us to provide a natural semantics for Aristotle's modal categorical sentences in order to illuminate Aristotle's hard to understand system of modal syllogisms.

3.2 A causal analysis of Aristotelian demonstrations

Many dialogues of Plato focus on questions of the form 'What is X?', where X is typically some moral property like *virtue* or *courage*, a natural kind of thing like *human*, or *water*, or a mathematical object like *a triangle*. A good answer to this kind of question must consist of a set of features all and only all individuals of type X have. Aristotle, a pupil of Plato, was interested in the same kind of questions. But he also was more ambitious. If all (and only all) individuals or objects of type X share certain features, Aristotle also wanted to know *why*. Indeed, for Aristotle, scientific inquiry is an attempt to answer 'why' questions. A scientific explanation of a fact about the world consists of a valid syllogistic argument with some fundamental true claims as its premises and this fact as the conclusion. But not any old valid syllogism would do, for the premises must express *fundamental* true claims. A valid syllogism that satisfies this extra requirement Aristotle calls a *demonstration*. A typical Aristotelian demonstration is the following:

(3)
a. All animals are living things.
b. All humans are animals.
c. Therefore, all humans are living things.

In this demonstration, the two premisses are taken to express essential features of animals and humans, respectively. They follow from Aristotle's theory of *real definitions* of objects of type X in terms of (i) an immediately higher type Y, and a differentia Z. If X is 'human', for instance, then Y would be 'animal', and Z would be 'rational': a man is a rational animal. Thus, in 'All humans are animals', 'being animal' is essentially predicated of humans, and the second premise of the above syllogism can be expressed by $Sa^{\Box}P$. However, not all true sentences of the form $Sa^{\Box}P$ can be read off directly from Aristotle's theory of real definitions. Some have to be indirectly derived. This is what happens in the above syllogism. In the above syllogistic argument, the premisses can be directly read off from Aristotle's theory of definition, but to reach the conclusion an additional argument is needed. This is provided by the syllogism,

[8]Of course, the causal notions PS_A^C and p_{ac} demand this as well in case their values are 1, but in addition they demand that A is a cause of C, and not that A is uniquely caused by C. If we limit ourselves to values that are 1 or not, the probabilistic measure is antisymmetric, and thus gives rise to a partial order.

that can be stated as being of the form $Ba^{\Box}C, Aa^{\Box}B \therefore Aa^{\Box}C$. For Aristotle, this argument *explains why* humans are living things. The argument turns a fact into a *reasoned fact*.[9]

What has this all to do with causality? Well, Aristotle had a somewhat wider notion of causality than many moderns have. For him, it is necessary for humans to be able to learn grammar. But being able to learn grammar is not an essential property of humans or of any higher kind. It just *causally follows by necessity* from being rational (according to Aristotle). Thus, even though all and only all objects of type X have feature f and g, it can be that one of the features is still only a derived feature, causally derived.

So far, it seems that scientific demonstrations must consist of two premisses that are both necessary. But this is not exactly what Aristotle seems to assume. In fact, in his *Posterior Analytics* Aristotle discusses the following two valid syllogisms:

(4) a. All objects that are near the earth do not twinkle
b. All (the) planets are near the earth
c. Therefore, (all) the planets do not twinkle.

and

(5) a. All objects that do not twinkle are near the earth.
b. All (the) planets do not twinkle.
c. Therefore, (all) the planets are near the earth.

In these arguments, the premises (4-b) and (5-a) are not taken to express necessary truths. Although the second syllogism is not taken to be a scientific demonstration, Aristotle claims that the first syllogistic inference is. It leads to a 'reasoned fact', because the middle term 'being near the earth' *causally explains* the conclusion, something that is not the case for the middle term in the other inference 'objects that do not twinkle'. If we would translate the above arguments in modal syllogistic terms, they would be of the forms $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$ and $BaC, Aa^{\Box}B \therefore Aa^{\Box}C$, respectively. Note that they are thus of types Barabara LXL and Barbara XLL, respectively.[10] Note also that in his *Prior Analytics*, Aristotle took only the first type of argument valid. So, there seems to be a close relation between what Aristotle claims in his two *Analytics*.

4 Causality and modal syllogisms

Causal links need not only connect propositions, they can connect properties, or features, as well. In fact, Danks (2014) argues that all prominent theories of concepts could be represented by graphical causal models. Although not explicitly discussed, the essentialistsâ version is one: features of birds are connected (and thus caused) in various strengths to the essence of the kind, i.e., by what it is to be a bird.

Let us now come back to the question what the natural interpretation of Aristotleâs *modal* syllogistics is. Recall that $Aa^{\Box}B$ means that all As are *necessary/essentially B*

[9]For much more detailed and sophisticated analyses of Aristotelian demonstration see Crager (2015) and Vecchio (2016).

[10]According to Vlecchio (2016), the argument in (9) explains why planets do not twinkle, by using a fact is which part of the nominal definition of a planet ('being near the earth'), but which is not a part of its real definition.

and that Aristotle claimed that the following modal syllogisms are valid and invalid, respectively:

1. $Ba^{\Box}C, Aa^{\Box}B \therefore Aa^{\Box}C$ Valid Barbara LLL
2. $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$ Valid Barbara LXL
3. $BaC, Aa^{\Box}B \therefore Aa^{\Box}C$ Invalid Barbara XLL

Similarly, Aristotle claims that the following modal syllogism is valid, where $Be^{\Box}C$ means that by (*de re*) necessity no B is a C:

4. $Be^{\Box}C, AaB \therefore Ae^{\Box}C$ Valid Celarent LXL

Moreover, Aristotle claims that not only conversion inference 5 is valid, but that the same holds for the modal conversion inferences 6 and 7:

5. $AeB \therefore BeA$ Valid
6. $Ae^{\Box}B \therefore Be^{\Box}A$ Valid
7. $Aa^{\Box}B \therefore Bi^{\Box}A$ Valid

We claim that Aristotle's claims make perfect sense once we understand $Aa^{\Box}B$ as causally explaining *why* B. More in particular, we would like to say that $Aa^{\Box}B$ just means that A has complete causal power to make B to hold, i.e., $PS^B_A = 1$ (or $p_{ab} = 1$) and that $Ae^{\Box}B$ just means that both A or B has complete causal powers to prevent the other to hold, i.e., $PS^{\neg B}_A = 1$ and $PS^{\neg A}_B = 1$ (or $p_{a\neg b} = 1$ and $p_{b\neg a} = 1$).[11]

Definition 2 Truth conditions of universal modal sentences.

- $Aa^{\Box}B$ *is true iff* $PS^B_A = 1$
- $Ae^{\Box}B$ *is true iff* $PS^{\neg B}_A = 1$ *and* $PS^{\neg A}_B = 1$

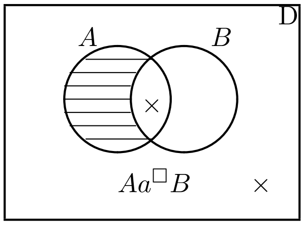
A simple fact about probabilities is that if $P(A), P(B) \neq 0$, then $P(\neg B|A) = 1$ iff $P(B|A) = 0$ iff $P(A \wedge B) = 0$ iff $P(\neg A|B) = 1$. Because of this, and if we assume the consistency assumption for counterfactuals, exogeneity and monotonicity, and assume in addition that $P(A), P(B) \neq 0$, we can derive immediately the following facts from the above proposed analysis of modal categorical sentences:

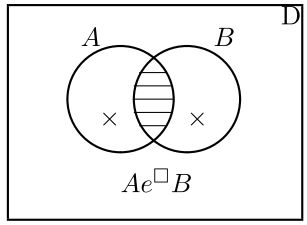
Fact 3 *Facts about truth conditions of universal modal sentences.*

- $Aa^{\Box}B$ *is true iff* $P(B|A) = 1$ *and* $P(B|\neg A) \neq 1$
  iff $P(A \wedge \neg B) = 0$ *and* $P(\neg A \wedge \neg B) \neq 0$
- $Ae^{\Box}B$ *is true iff* $P(A \wedge B) = 0$ *and* $P(\neg A \wedge B) \neq 0$ *and* $P(\neg B \wedge A) \neq 0$

[11] Aristotle's (hyperintensional) distinction between necessity and essentiality suggests that the analysis of $Aa^{\Box}B$ as $p_{ab} = 1$ is still too coarse-grained. Notice, however, that even if A and B are necessary co-extensive, it will typically be (causally speaking) that either $p_{ab} = 1$ and $p_{ba} = 0$, or $p_{ab} = 0$ and $p_{ba} = 1$. We take the former to be the case if A is a substantive term and B an adjectival one.

This fact shows that these truth conditions can be captured in terms of Venn diagrams. However, besides circles for A and B, we now also need to have a domain of discourse, D, to account for negation:





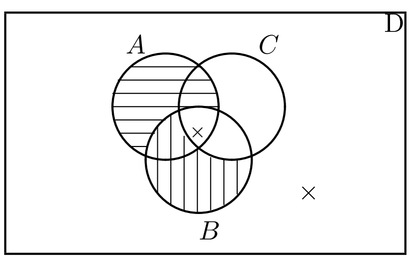
We will assume the interpretation rule for non-modal universal categorical sentences as in definition 1 repeated below

Definition 3 Truth conditions of non-modal Categorical sentences

- *AaB is true iff $P(B|A) = 1$,*
- *AiB is true iff $P(A \cap B) \neq 0$,*
- *AeB is true iff $P(A \cap B) = 0$, and*
- *AoB is true iff $P(B|A) \neq 1$*

Inference 1 is valid on this interpretation, because if the premisses are true the following will hold (i) $P(C|B) = 1$, (ii) $P(C|\neg B) \neq 1$, (iii) $P(B|A) = 1$ and (iv) $P(B|\neg A) \neq 1$. Obviously, by (i) and (iii) it follows that $P(C|A) = 1$. From (ii) and (iv) it follows that (a) there are some $\neg C$s among the $\neg B$s, and (b) that there are some $\neg B$s among the $\neg A$s. By (a) and (b) this means that $P(C|\neg A) \neq 1$. Thus, $P(C|A) = 1$ and $P(C|\neg A) \neq 1$ which means that $Aa^{\Box}C$.

The validity of the inference can be checked by the following Venn diagram:

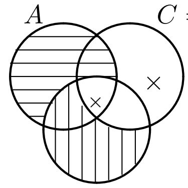


Inference 2 is also valid on this interpretation, because if the premisses are true it means that the following will hold (i) $P(C|B) = 1$, (ii) $P(C|\neg B) \neq 1$ and (iii) $P(B|A) = 1$. Obviously, by (i) and (iii) it follows again that $P(C|A) = 1$. From (ii) it follows that there are some $\neg C$s among the $\neg B$s. But because AaB, it holds that all $\neg B$s are $\neg A$s, and thus there must also be some $\neg C$s among the $\neg A$s. Thus, $P(C|A) = 1$ and $P(C|\neg A) \neq 1$ which means that $Aa^{\Box}C$. The validity of this inference follows from the same Venn diagram as the one that illustrates inference 1.

Inference 3, however, is not valid. The important thing to observe is that this is just an instance of 'right weakening',[12] or **right upward monotonicity**, an inference which should (and does) **not** hold on our causal analysis. In particular, the

[12] In condtional terms, right weakening means that if $A \Rightarrow B$ and $B \models C$, then also $A \Rightarrow C$.

inference has a counterexample, in case the domain consists only of C individuals. The counterexample is illustrated by the following Venn diagram.



Similarly, we can account for Aristotle's intuition that inference 4 is valid. Using the above interpretation of non-modal statements, we account for inference 5. The validity of inference 6 is obvious given the truth conditions of $Ae^{\Box}B$. As for inference 7, this immediately follows from the semantic analysis of statements like $Bi^{\Box}A$ to be given in a minute.

Our predictions agree with all Aristotle's claims of (in)validities of universal modal syllogisms with modality $\Box$. For instance, we correctly predict Aristotle's claimed validity of Cesare LXL, Camestres XLL, and his claim of invalidity of Camester LXL. The latter one – $Ba^{\Box}A, CeA \not\models Be^{\Box}C$ – is particularly interesting. It is easy to see that this inference would be predicted as valid, if we analysed CeA as true iff $P(A|C) = 0$, which presupposes that $P(C) \neq 0$. However, we have analysed CeA as true iff $P(C \wedge A) = 0$, and on this interpretation Camestres LXL is *not* predicted to be valid, in accordance with Aristotle's intuitions.[13] More in particular, our analysis makes the right predictions for the modal Barbara and Celarent syllogisms of the first figure.

As for the second figure, and limiting ourselves to universal statements, we have to explain why (according to Aristotle)

(6) a. $Ae^{\Box}B, CaB \models Ce^{\Box}A$ Cesare LXL
b. $AeB, Ca^{\Box}B \not\models Ce^{\Box}A$ Cesare XLL

and

(7) a. $Aa^{\Box}B, CeB \not\models Ce^{\Box}A$ Camestres LXL
b. $AaB, Ce^{\Box}B \models Ce^{\Box}A$ Camestres XLL

As for (6-a), this follows immediately from our semantics. For (6-b) this follows because AeB is true $P(A \wedge B) = 0$. As for (7-a). This doesn't follow, because it is not guaranteed that $P(C|\neg A) \neq 0$, which makes the conclusion false.[14] Inference (7-b) is immediately verified. There are no other modal syllogisms with only universal statements of the second figure to be checked, and we don't know about Aristotle's intuitions on only 'universal' modal syllogisms of the fourth figure (Cameses$_4$). Because all valid syllogisms of the third figure involve non-universal sentences as well, we predict for all modal syllogisms that only involve universal sentence in accordance with Aristotle's intuition.

[13] Note, though, that we would predict invalidity as well if we interpreted $Aa^{\Box}B$ as being true iff $A \subseteq B$ and $P(\neg A \cap \neg B) \neq 0$ and interpreted AeB as true iff either $P(B|A) = 0$ or $P(A|B) = 0$. Although these interpretation rules would also give us the correct predictions for inferences 1 until 5, the interpretation rule for $Aa^{\Box}B$ would, unfortunately, not give us inference 7.

[14] Alternatively, we could say that AeB is true iff either $P(B|A) = 0$ or $P(A|B) = 0$. That would get those inferences right as well.

As for modal syllogisms with non-universal sentences, we first need to know what makes sentences like $Ai^{\Box}B$ true. In counterfactual terms, it seems natural to propose that $Ai^{\Box}B$ is true iff $\exists x : xaA, \exists D : xa^{\Box}D$ and $P(B_D|\neg B, \neg D) = 1$, where xaA is the singular categorical sentence that (all) x is A, and $xa^{\Box}D$ the singular categorical sentence that (all) x is necessary D. Notice that in non-counterfactual terms, our interpretation of $Ai^{\Box}B$ comes down to the following: $Ai^{\Box}B$ is true iff $\exists x : xaA$ & xaB and $\exists y : yeB$. But we want to account for conversion $Ai^{\Box}B \models Bi^{\Box}A$ as well. Therefore, we will propose a more symmetric definition: $Ai^{\Box}B$ is true iff $\exists x, \exists D$: $xa^{\Box}D$ and (i) xaA and $PS^B_D = 1$ or (ii) xaB and $PS^A_D = 1$. To simplify things, however, we won't make use of property D, but just use singular modal sentences like $xa^{\Box}B$, instead. Notice that this modal sentence just reduces to the conjunction of two non-modal sentences: xaB and $\exists y : yeB$. We will do the same to give the truth conditions of the modal sentence $Ao^{\Box}B$.

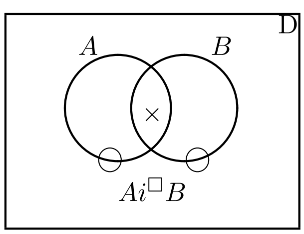
Definition 4 *Truth conditions of non-universal modal sentences.*

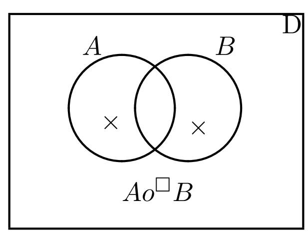
- $Ai^{\Box}B$ *is true iff* $\exists x : xaA$ & xaB *and* ($xa^{\Box}B$ *or* $xa^{\Box}A$)
  iff $\exists x : xaA$ & xaB *and* $\exists y : yeB$ *or* yeA

- $Ao^{\Box}B$ *is true iff* $\exists x : xaA$ & xeB *and* $\exists y : yeA$ & yaB

Notice that we didn't provide the simpler and perhaps more intuitive truth conditions for $Ao^{\Box}B$: $Ao^{\Box}B$ is true iff $\exists x : xaA$ and $xe^{\Box}B$. Our truth conditions are more complicated, because we used y such that yeA instead of $\neg x$. We need these more complicated truth conditions because the simpler truth conditions can't account, for instance, for Aristotle's claimed invalidity of Baroco XLL, at least if we interpret AaB as true iff $P(B|A) = 1$.[15]

Interestingly, also these non-universal modal sentences can be captured in terms of Venn diagrams, if we make one addition: if we have circles ◯ in two areas, then we know that at least one of those areas must be non-empty:



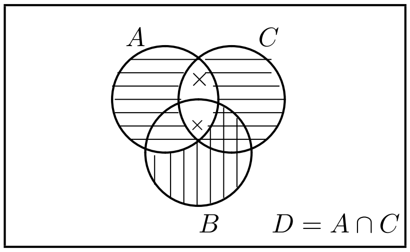


Notice that from the above interpretation rules of $Aa^{\Box}B$ and $Ai^{\Box}B$, inference 7, the conversion inference $Aa^{\Box}B \therefore Bi^{\Box}A$ is immediately predicted to be valid, in accordance with Aristotle's intuitions. Let us now see whether we can account for Aristotles' claims with respect to modal syllogisms involving also non-universal sentences. First, Aristotle claims (8-a) (of the first figure) to be valid, but (8-b) not to be so:

[15] To be clear, the simpler interpretation rule for $Co^{\Box}B$ – which would come down to $P(\neg B|A) \neq 0$ and $P(B) \neq 0$ – is possible, if Baroco XLL were not valid. It is interesting to observe that although Aristotle claims that he found a counterexample to Baroco XLL, several commentators (e.g. Van Rijen, 1989; Patterson, 1995) have argued that he was mistaken. For discussion, see Malink (2013). Alternatively, we could use the simpler and more intuitive interpretation rule for $Co^{\Box}B$, if we would interpret AaB as true iff $A \subseteq B$. This interpretation rule for AaB gives problems at other places, however.

(8) a. $Ba^{\Box}A, CiB \models Ci^{\Box}A$ Darii LXL
b. $BaA, Ci^{\Box}B \not\models Ci^{\Box}A$ Darii XLL

Our interpretation rules of non-universal modal sentences indeed make Darii LXL valid. Moreover, these interpretation rule makes Darii XLL invalid, as desired. The following Venn diagram shows the counterexample to Darii XLL.



Aristotle also claims a distinction between the following syllogisms, also of the first figure:

(9) a. $Be^{\Box}A, CiB \models Co^{\Box}A$ Ferio LXL
b. $BeA, Ci^{\Box}B \not\models Co^{\Box}A$ Ferio XLL

Inference (9-a) follows immediately if we analyse $Co^{\Box}A$ as true iff $\exists x : xaC, \exists D : xa^{\Box}D$ and $P(\neg A_D|B, \neg D) = 1$. There is an easy counterexample to (9-b), again due to the fact that the conclusion $Co^{\Box}A$ demands that there is at least one A, while premise BeA can be true without there being such an A. Notice, though, that we have not analysed $Co^{\Box}A$ as above, but rather as in definition 4. Fortunately, the validity of Ferio LXL and the invalidity of Ferio XLL follows from this interpretation rule as well, as might be checked by a Venn diagram. We leave this to the reader.

Aristotle didn't give his opinion on every possible syllogism which involves sentences with necessity modals. In fact, he limited himself to syllogisms that (i) have a necessity modal in the conclusion, (ii) are of the first three figures and (ii) that are valid without any modal. Still, there are 6 valid syllogisms in each figure, and 3 possible combinations where at least one of the premises has a necessity modal. Of those 54 syllogisms, Aristotle expressed his opinion on 42 of those modal syllogisms.[16] 23 of those syllogisms he counted valid, and the others non-valid. He looked at 14 syllogisms where all categorical sentences involved had a necessity modal, such as Barbara LLL, and he counted all of them as valid. We can check that all such modal syllogisms are valid on our analysis as well. Let us go to one of the more challenging ones to explain: Darii LLL, $Ba^{\Box}A, Ci^{\Box}B \models Ci^{\Box}A$. The first premise means that $PS^A_B = 1$. According to the second premise, $\exists x : xaC, \exists D : xa^{\Box}D$ and $PS^B_D = 1$. Because if $PS^A_B = 1$ and $PS^B_D = 1$, it follows by transitivity that also $PS^A_D = 1$. It follows that thus $\exists x : xaC, \exists D : xa^{\Box}D$ and $PS^A_D = 1$, which means that conclusion $Ci^{\Box}A$ is true.

As for the other 30 modal syllogisms of this type that Aristotle considered, we checked them as well, and our analysis predicts in accordance with Aristotle's intuitions. Thus, our analysis makes predictions exactly in accordance with Aristotle's explicity discussed claims of (in)validity for **every** modal syllogism in which at most the modal $\Box$ occurs (a system also known as 'apodeictic syllogisms')!

Theorem 1 *Using the truth conditions of categorical sentences as given in definitions*

[16] We base ourselves here completely on appendix A of Malink (2013).

2 until 4, all and only all <u>*apodeictic*</u> *syllogisms are predicted to be valid that Aristotle counted as valid.*

We think this result is quite remarkable. What is perhaps even more remarkable is that validity of epodeictic syllogisms can be decided by means of Venn diagrams:

Theorem 2 *Validity of epodeictic modal syllogisms as discussed by Aristotle in his Prior Analytics can be decided by means of Venn diagrams.*

We haven't checked our predictions for *all* 16.384 modal syllogisms, though. In fact, we didn't check any syllogism that involve possibility and contingency modals that Aristotle also discussed. In this paper we did not even propose meanings of such sentences. But the smoothness of our explanation of Aritotle's intuitions concerning apodeictic syllogisms makes one optimistic that we can also account for Aristotle's intuitions on other modal syllogisms.

But there is further ground for optimism. Malink (2013) and Vecchio (2016) have recently shown how to account for most (if not all) of the Aristotle's claims about modal syllogisms making use of *essences*. $Aa^{\Box}B$ is true iff all As are B *in virtue of* what it is to be an A. But that is exactly how we think of our own proposal as well.

5 A challenge: counterexamples to Barbara LXL?

We have shown in the previous section that our causal power analysis can account for why the modal syllogism Barbara LXL, $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$ is valid, although Barbara XLL, $BaC, Aa^{\Box}B \therefore Aa^{\Box}C$, is not. We have seen that this can be shown if we analyse statements like $Ba^{\Box}C = 1$ by $\frac{P(C|B)-P(C|\neg B)}{1-P(C|\neg B)} = 1$ and $AaB = 1$ by $P(B|A) = 1$. We have also seen that the causal notions of causal power and PS_A^C come down to this probabilistic notion under certain circumstances.

Although Aristotle claimed that Barbara LXL is valid, very soon (putative) counterexamples to this modal syllogisms were offered:[17]

(10) a. All litererats necessarily have knowledge, all men are litarate, thus all men necessarily have knowledge.
b. $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$ Barbara LXL

In fact, Aristotle himself provided a (putative) counterexample to Celarent LXL himself.

(11) a. All ill people are necessarily not healthy, all men are ill, thus all men are necessarily not healthy.
b. $Be^{\Box}C, AaB \therefore Ae^{\Box}C$ Celarent LXL

Malink (2013) and Crager (2015) argue that these counterexamples can be explained away if we take seriously Aristotle's analysis of 'genuine predication' from Aristotle's *Categories*. The idea is that terms can denote sets of different *ontological types*: some denote *substances*, while others denote *qualities*. Just as each substance has an essence, this is also the case for each quality. However, denotations of the same type can only stand in a limited number of extensional relations with each other. For instance, for any two substances A and B, it cannot be that $A \cap B \neq \emptyset$ without either

[17] For modern discussion, see van Rijen (1989), Rini (1989), Malink (2013) and Crager (2015).

$A \subset B$ or $B \subset A$. Beyond this *extensional* constraint, there lays a more important *intensional* constraint: if A and B are of the same ontological type, then, if $A \subset B$, then $Aa^{\Box}B$. Malink (2013) and Crager (2015) argue that Aristotle took Barbara LXL and Celarent LXL to be valid because he demanded that in a demonstration with a necessary conclusion, also the seemingly nonmodal premise (in our cases, the minor premise AaB) should be a case of genuine predication.

If Malink (2013) and Crager (2015) are correct, it means that valid modal syllogisms with a necessity modal in the conclusion should, in the end, all be of the form LLL. It also suggests that our explanation in the previous section of the validity of Barbara LXL and Celarent LXL will not be correct, for otherwise the (putative) counterexamples above would likely be genuine counterexamples. If we want to stick to our causal analysis, this suggests that instead of looking at the *extensional* notion $\frac{P(C|B)-P(C|\neg B)}{1-P(C|\neg B)} = 1$ for the analysis of $Ba^{\Box}C$ we should look at the *intensional* counterpart, $\frac{P(C_B)-P(C_{\neg B})}{1-P(C|\neg B)} = 1$, where intervention still plays an important role, and the counterfactual probabiltiy $P(B_A)$ is not reduced to the conditional probability $P(B|A)$. Indeed, on such an intensional analysis Barbara LXL, $Ba^{\Box}C, AaB \therefore Aa^{\Box}C$, would not be valid, because from $\frac{P(C_B)-P(C_{\neg B})}{1-P(C|\neg B)} = 1$ and $P(B|A) = 1$, we cannot conclude that $\frac{P(C_A)-P(C_{\neg A})}{1-P(C|\neg A)} = 1$.

We don't know, though, whether Malink's (2013) and Crager's (2015) interpretation of Aristotle is correct. For one thing, Malink (2013) himself already notes that Aristotle explicitly discusses modal syllogisms that he takes to be valid even though the nonmodal premise does not seem to involve genuine predication. But, of course, if Malink and Crager are not correct, we would have to explain away the above 'putative' counterexamples in another way. In fact, Vecchio (2016, chapter 1) argues that Aristotle himself explained away the (putative) counterexamples to Barbara LXL and the like in a more straightforward way than was suggested by Malink (2013): by demanding that the terms are interpreted in an omnitemporal way, which makes the non-modal premise false. Vecchio (2016, chapter 3) also argues explicitly that Aristotle used syllogisms of the form Barbara LXL in his analysis of scientific demonstrations in the *Posterior Analytics*, just as we suggested in section 3.2. Vecchio argues that Barbara LXL can be used to turn a *nominal* definition, 'Thunder is a noise in the clouds' (of form AaB) to a *real* definition 'Thunder is (necessarily) the extinguishing of fire in the clouds' (of form $Aa^{\Box}C$) via the essential major premise 'A noise in the clouds is (by necessity) the extinguishing of fire in the clouds' (of form $Ba^{\Box}C$).[18] Note that if Vecchio is right, our 'extensional' causal analysis might be on the right track after all.

6 Conclusion and Outlook

In this paper we have shown that Aristotle's intuitions about apodeictic syllogisms as expressed in his *Prior Analytics* can be captured semantically by giving a causal semantics of modal categorical sentences. Moreover, we have seen that this causal semantics can be reduced to an extensional analysis just making use of probabilities, which allowed to check modal syllogisms by simple Venn diagrams. The only

[18]There exists an interesting analogue between this and the way natural kind terms receive their content according to the causal theory of reference: first a set of superficial properties is used to identify a set of things, and later having these superficial properties is explained by some essential properties all the things in the set have in common.

real complication is that whereas for standard syllogisms no domain of discourse was required, we need such a domain now, because for our analysis of modal syllogisms information about the *complement* of the denotations of terms is crucial. (Of course, we need such complications as well, once we allow negative terms to occur in standard syllogisms.) Finally, we have argued that we can motive our causal analysis by Aristotle's analysis of *demonstrative proofs* as worked out in his *Posterior Analytics*.

Of course, we will never know whether our causal analysis fits Aristotle's semantic intuitions on modal syllogisms, because he never clearly stated these intuitions in the first place. But this leaves open the question whether our semantic analysis is plausible in the first case. One reviewers doubted the plausibility of our analysis, suggesting that the difference between AaB and $Aa^{\Box}B$ should not just be that $P(\neg A \wedge \neg B) \neq 0$. More in general one might doubt whether the truth conditions of modal categorical sentences could be described at all by Venn diagrams. We think that there are two points to be made here. *First* of all, our basic idea is that a sentence like $Aa^{\Box}B$ should be analysed causally as saying that $p_{ab} = 1$, or better perhaps that $PS_A^B = 1$. On this causal view, modality statements are really treated in an *intensional* (or even hyperintensional) manner. It is just that by making certain assumptions that $PS_A^B = 1$ holds exactly if $P(B|A) = 1$ and $P(B|\neg A) \neq 1$. Notice that if one of those assumptions is not made, the reduction of the causal notion to the purely probabilistic one would not go through. For instance, one might doubt that for causality we should really demand the *consistency assumption*, saying that if A (or $\neg A$) holds, the truth value of B_A (or $B_{\neg A}$) is the same as the truth value of B. This assumption comes down to the *strong centering* assumption known from conditional logic, and corresponds with the inference $A, B \therefore A \Rightarrow B$. Intuitively, one might argue, this inference should not hold if '$\Rightarrow$' expresses a relation of causal relevance. Indeed, A and B can both be true without there being a causal relation between them. Once the consistency condition is given up, truth conditions of modal categorical sentences could not be reduced to simple probabilistic claims that can be expressed by Venn diagrams. Something similar holds when we give up the exogeneity condition or the monotonicity condition. Importantly, however, we think that our semantics is still appropriate if we disregard the reduction to simple probabilistic claims.[19] *Second*, we don't think it is strange that the complements of the denotations of A and B *should* play a role for the semantic analysis of $Aa^{\Box}B$. Recall that the basic idea of our analysis is that $Aa^{\Box}B$ is true if A has the causal power to make B true. For A to have the causal power to make B true means that A must *make a difference* to the truth B. But if B is a necessary truth, A cannot make such a difference. So, for $Aa^{\Box}B$ to hold, there must be a non-B individual. But obviously, $Aa^{\Box}B \models AaB$, so this non-B individual cannot be an A-individual. Thus, there must be a $\neg A \wedge \neg B$-individual, meaning that $P(\neg A \wedge \neg B) \neq 0$.

Although we are surprised that our semantic analysis captures so many of Aristotle's intuitions, and in particular that this could be done by using Venn diagrams, we don't think that our analysis is, in general, unnatural. There is only one interpretation rule that we feel is really artificial: our interpretation rule for $Ao^{\Box}B$. This interpretation rule is artificial already because it is symmetric. This interpretation rule was given just to get the 'facts' right. These 'facts' are now Aristotle's intuitions, and we noted already in footnote 14 that his intuitions might as well be mistaken on the

[19] There is one real worry we have, though, and that is our semantic analysis of $Ao^{\Box}B$. We fear that our proposed analysis is not exactly natural, for one thing because it entails that $Ao^{\Box}B$ entails $Bo^{\Box}A$.

crucial modal syllogism (Baroco XLL) that forced us to our artificial interpretation rule.

As mentioned above, in his *Prior Analytics* Aristotle also discussed inferences concerning *possibility* and *contingency* modals. Of course, for the standard possibility modal, a natural analysis suggests itself:

(12)
$$Aa^{\Diamond}B \equiv \neg(Ao^{\Box}B) \qquad Ae^{\Diamond}B \equiv \neg(Ai^{\Box}B)$$
$$Ai^{\Diamond}B \equiv \neg(Ae^{\Box}B) \qquad Ao^{\Diamond}B \equiv \neg(Aa^{\Box}B)$$

We think, however, that to provide a semantic account of possibility statements we need to give up the assumption that we made in section 3.1: that statements like C_A have a truth value in $\{0, 1\}$. We hypothesise that such statements have to have a value in $[0, 1]$, instead, thought of as the *chance* of C after an intervention to make A true. But it remains to be seen whether such an analysis gives rise to predictions that accord with Aristotl's intuitions. It is even less clear whether we can account for Aristotle's claims involving the contingency modal, Δ, a task that is perhaps the most challenging. Striker (1985) argues, though, that sentences like $Aa^{\Delta}B$ should be interpreted basically as generic sentences, where B applies *by nature*, or *for the most part*, to A. Interestingly, this suggestion would be much in line with van Rooij & Schulz's (2020) analysis of generic sentences, according to which sentences of the form 'As are B' are interpreted as having high causal power, i.e. $p_{ab} \approx 1$. But it is more natural to interpret $Aa^{\Delta}B$ as $\forall x \in A : \neg\exists D : xa^{\Box}D$ and $(Da^{\Box}B$ or $De^{\Box}B)$ and $Ai^{\Delta}B$ as $\neg\exists x \in A : \exists D : xa^{\Box}D$ and $(Da^{\Box}B$ or $De^{\Box}B)$ to account for Aristotle's claims that $Aa^{\Delta}B$ is equivalent with $Ae^{\Delta}B$ and $Ai^{\Delta}B$ with $Ao^{\Delta}B$, and that not only $Ai^{\Delta}B$ is equivalent with $Bi^{\Delta}A$, but also that $Ao^{\Delta}B$ is equivalent with $Bo^{\Delta}A$. We don't know whether with this interpretation we can account for all of Aristotle's intuitions w.r.t. modal syllogisms involving Δ.

The bulk of this paper is about modal syllogisms, involving sentences that are either true or false. As mentioned in section 3, however, our approach was motivated by the *quantitative* causal analysis of conditionals and generic sentences of van Rooij & Schulz (2019, 2020). It is well-known that Adams (1965, 1966) developed a well-behaving probabilistic entailment relation $\models^p$ based on the assumption that the assertability of conditional $A \Rightarrow C$ 'goes with' the corresponding conditional probability, $P(C|A)$. This logic can be axiomatised and is now known as the basic non-monotonic logic: system **P**. A question that is still open is whether a similarly well-behaved logic can be developed that is based on the assumption of van Rooij & Schulz (2019, 2020) that conditionals and generic sentences express relations of causal relevance. The causal relevance of A for B is measured by Cheng's notion of the *causal power* of A to produce B, or (better perhaps) by Pearl's notion of the 'probability of causal sufficiency'. Because the values of these measures can be anywhere between -1 and 1, this open question is difficult to handle. The question would be easier to handle, however, when we care only whether these causal powers have values 1 or 0. Then the question becomes whether it is possible to develop a logic for conditionals that express such *qualitative* causal relevance relations. But notice that on our causal semantics of Aristotelian modal sentences we have limited ourselves to qualitative causal relevance relations. This suggests that Aristotle's system of modal syllogisms, or something very close to it, can actually be viewed as the qualitative logic that deals with causal conditionals!

A Appendix

A.1 Table of modal syllogisms with necessity modals

Table 1: Conversion rules for necessity modality

Form of conversion rule	Validness
From $Aa^{\Box}B$ to $Bi^{\Box}A$	valid
From $Ai^{\Box}B$ to $Bi^{\Box}A$	valid
From $Ae^{\Box}B$ to $Be^{\Box}A$	valid

Table 2: Apodeictic syllogistic of first figure discussed by Aristotle

Name of syllogisms	Form of syllogism	Validness
Barbara LLL	From $Ba^{\Box}C$, $Aa^{\Box}B$ to $Aa^{\Box}C$	valid
Barbara LXL	From $Ba^{\Box}C$, AaC to $Aa^{\Box}C$	valid
Barbara XLL	From BaC, $Aa^{\Box}C$ to $Aa^{\Box}C$	invalid
Celarent LLL	From $Be^{\Box}C$, $Aa^{\Box}B$ to $Ae^{\Box}C$	valid
Celartent LXL	From $Be^{\Box}C$, AaC to $Ae^{\Box}C$	valid
Celarent XLL	From BeC, $Aa^{\Box}C$ to $Ae^{\Box}C$	invalid
Darii LLL	From $Ba^{\Box}C$, $Ai^{\Box}B$ to $Ai^{\Box}C$	valid
Darii LXL	From $Ba^{\Box}C$, AiB to $Ai^{\Box}C$	valid
Darii XLL	From BaC, $Ai^{\Box}B$ to $Ai^{\Box}C$	invalid
Ferio LLL	From $Be^{\Box}C$, $Ai^{\Box}B$ to $Ao^{\Box}C$	valid
Ferio LXL	From $Be^{\Box}C$, AiB to $Ao^{\Box}C$	valid
Ferio XLL	From BeC, $Ai^{\Box}B$ to $Ao^{\Box}C$	invalid

Table 3: Apodeictic syllogistic of second figure discussed by Aristotle

Name of syllogisms	Form of syllogism	Validness
Cesare LLL	From $Ce^{\Box}B$, $Aa^{\Box}B$ to $Ae^{\Box}C$	valid
Cesare LXL	From $Ce^{\Box}B$, AaB to $Ae^{\Box}C$	valid
Cesare XLL	From CeB, $Aa^{\Box}B$ to $Ae^{\Box}C$	invalid
Camestres LLL	From $Ca^{\Box}B$, $Ae^{\Box}B$ to $Ae^{\Box}C$	valid
Camestres LXL	From $Ca^{\Box}B$, AeB to $Ae^{\Box}C$	invalid
Camestres XLL	From CaB, $Ae^{\Box}B$ to $Ae^{\Box}C$	valid
Festino LLL	From $Ce^{\Box}B$, $Ai^{\Box}B$ to $Ao^{\Box}C$	valid
Festino LXL	From $Ce^{\Box}B$, AiB to $Ao^{\Box}C$	valid
Festino XLL	From CeB, $Ai^{\Box}B$ to $Ao^{\Box}C$	invalid
Baroco LLL	From $Ca^{\Box}B$, $Ao^{\Box}B$ to $Ao^{\Box}C$	valid
Baroco LXL	From $Ca^{\Box}A$, AoB to $Ao^{\Box}C$	invalid
Baroco XLL	From CaB, $Ao^{\Box}B$ to $Ao^{\Box}C$	invalid

Table 4: Apodeictic syllogistic of third figure discussed by Aristotle

Name of syllogisms	Form of syllogism	Validness
Darapti LLL	From $Ba^{\Box}C$, $Ba^{\Box}A$ to $Ai^{\Box}C$	valid
Darapti LXL	From $Ba^{\Box}C$, BaA to $Ai^{\Box}C$	valid
Darapti XLL	From BaC, $Ba^{\Box}A$ to $Ai^{\Box}C$	valid
Felapton LLL	From $Be^{\Box}C$, $Ba^{\Box}A$ to $Ao^{\Box}C$	valid
Felapton LXL	From $Be^{\Box}C$, BaA to $Ao^{\Box}C$	valid
Felapton XLL	From BeC, $Ba^{\Box}A$ to $Ao^{\Box}C$	invalid
Disamis LLL	From $Bi^{\Box}C$, $Ba^{\Box}A$ to $Ai^{\Box}C$	valid
Disamis LXL	From $Bi^{\Box}C$, BaA to $Ai^{\Box}C$	invalid
Disamis XLL	From BiC, $Ba^{\Box}A$ to $Ai^{\Box}C$	valid
Datisi LLL	From $Ba^{\Box}C$, $Bi^{\Box}A$ to $Ai^{\Box}C$	valid
Datisi LXL	From $Ba^{\Box}C$, BiA to $Ai^{\Box}C$	valid
Datisi XLL	From BaC, $Bi^{\Box}A$ to $Ai^{\Box}C$	invalid
Bocardo LLL	From $Bo^{\Box}A$, $Ba^{\Box}A$ to $Ao^{\Box}C$	valid
Bocardo LXL	From $Bo^{\Box}A$, BaA to $Ao^{\Box}C$	invalid
Bocardo XLL	From BoA, $Ba^{\Box}A$ to $Ao^{\Box}C$	invalid
Ferison LLL	From $Be^{\Box}C$, $Bi^{\Box}A$ to $Ao^{\Box}C$	valid
Ferison LXL	From $Be^{\Box}C$, BiA to $Ao^{\Box}C$	valid
Ferison XLL	From BeC, $Bi^{\Box}A$ to $Ao^{\Box}C$	invalid

References

Aristotle (1973), *Aristotle in Twenty-three Volumes*, Harvard University Press.

Adams, E.W. (1965), 'A logic of conditionals', *Inquiry*, **8**: 166-197.

Adams, E. (1966), 'Probability and the Logic of Conditionals', *Studies in Logic and the Foundations of Mathematics*, Vol 43, 265-316.

[0] Benthem, J. van (1986a), *Essays in Logical Semantics*, Foris publications, D. Reidel, Dordrecht.

Bird, A. (2007), *Nature's Metaphysics Laws and Properties*, Oxford University Press, Oxford.

Cartwright, N. (1989), *Nature's Capacities and their Measurement*, Oxford University Press, Oxford.

Cheng, P. (1997), 'From covariation to causation: A causal power theory', *Psychological Review*, **104**: 367-405.

Crager, A. (2015), *Metalogic in Aristotle's Epistemology*, PhD Dissertation, Princeton University.

Danks, D. (2014), *Unifying the Mind. Cognitive Representations as Graphical Models*, MIT Press.

Douven, I. (2008), 'The evidential support theory of conditionals', *Synthese*, **164**: 19-44.

[0] Eijck, J. van (1985), 'Generalized quantifiers and traditional logic', in J. van Benthem and A. ter Meulen (eds.), *Generalized Quantifiers and Natural Language*, Foris, Dordrecht, pp. 1-19.

Harré, R. and E. Madden (1975), *Causal Powers: A theory of natural necessity*, Basic Blackwell, Oxford.

Hintikka, J. (1973), *Time and Necessity: Aristotle's Theory of Modality*, Oxford.

[1] Leibniz, G. (1966), 'A Mathematics of Reason', in Parkinson (ed.), *Leibniz. Logical Papers*, pp. 95-104.

Lukasiewicz, J (1957), *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*, 2d edition, Clarendon Press, Oxford.

Malink, M. (2013), *Aristotle's Modal Syllogistic*, Harvard University Press.

McCall, S. (1963): *Aristotle's Modal Syllogisms*, Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam.

Oaksford, M. & N. Chater (2007), *Bayesian Rationality*, Oxford University Press.

Patterson, R. (1995), *Aristotle's Modal Logic*, Cambridge University Press.

Patzig, G., (1968), *Aristotle's Theory of the Syllogism*, translated by J. Barnes, Reidel, Dordrecht.

Pearl, J. (2000), *Causality: Models, Reasoning and Inference*, Cambridge University Press.

Rescher, N. (1964), 'Aristotle's theory of modal syllogisms and its interpretation', In M. Bunge (ed.), *The Critical Approach to Science and Philosophy*, pp. 152-77.

Rijen, J. van (1989), *Aspects of Aristotle's Logic of Modalities*, Synthese Historical Library, Vol 35.

Rini, A. (1998), 'Is there a modal syllogistic?', *Notre Dame Journal of Formal Logic*, **39**: 554-572.

Rooij, R. van & K. Schulz (2019), 'Conditionals, causality and conditional probability', *Journal of Logic, Language and Information*, **28**: 55-71.

Rooij, R. van & K. Schulz (2020), 'A causal power analysis of generics', *Topoi*.

Rott, H (2019), 'Difference-making conditionals and the relevant Ramsey test', *The Review of Symbolic Logic*, 1-32.

Shoemaker, S. (1980), 'Causality and Poperties', in Van Inwagen (ed.), *Time and Cause*, D. Reidel, pp. 109-135.

Spohn, W. (2013), 'A ranking-theoretical approach to conditionals', *Cognitive Science*, 37: 1074-1106.

Striker, G. (1985), 'Notwendigkeit mit Lücken', *Neue Hefte für Philososophie*, **24**: 146-164.

Thom, P. (1991), 'The two Barbaras', *Hist. Phil. Logic*, 12: 135-149.

[1] Thomason, S. K. (1993), 'Semantic analysis of the modal syllogistic', *Journal of Philosophical Logic*, 22: 111-128.

[1] Uckelman, S. L., & Johnston, S. (2010), 'A simple semantics for Aristotelian apodeictic syllogistics'. In L. Beklemishev, V. Goranko, & V. Shehtman (Eds.), *Advances in modal logic Vol. 8*, pp. 428-443, College Publications.

Vecchio, D. J. (2016), *Essence and Necessity, and the Aristotelian Modal Syllogistic: A Historical and Analytical Study*, PhD dissertation, Marquette University.