Two early Arabic applications of model-theoretic consequence

Wilfrid Hodges
wilfrid.hodges@btinternet.com

DRAFT 29 January 2018

1 Introduction

We will be concerned with formal logic, i.e. with logical relations that apply to all configurations of sentences that share a certain form. For example an inference

(1) Some C is a B.
No B is an A.
Therefore not every C is an A.

is a logical inference because it holds whatever we put for A, B and C.

There are two broad kinds of explanation of what it means to say that a set Φ of formal sentences has the formal sentence θ as a logical consequence. One kind of explanation is along the lines

(2) *There is* a proof that derives θ from Φ.

'Proof' can be understood in various ways here, but a standard notion is that a proof either is or represents a single inference step or a sequence of inference steps. We call a logical consequence along these lines a 'proof-theoretic consequence'. The other kind of explanation is that

(3) *Every* model of Φ is a model of θ.

Again we can understand 'model' in more than one way. The simplest way of taking it is that every interpretation of the formal sentences in $\Phi\cup\{\theta\}$ that makes all the sentences in Φ true makes θ true too. A logical consequence in the sense of this second kind of explanation is called a 'model-theoretic consequence'. The difference between the quantifiers 'There is' and 'Every' is a dividing line between the two kinds of explanation.

Reading any standard history of logic, you would easily get the impression that model-theoretic consequence is a modern notion. In fact it is often credited to a paper of Alfred Tarski [21] in 1936, though some historians point to a broadly similar definition given by Bernard Bolzano in his *Wissenschaftslehre* [4] §147 in 1837. Proof-theoretic consequence by contrast goes all the way back to Aristotle's *Prior Analytics* [2] in the 4th century BC.

So it was quite a surprise to find a fourteenth century Baghdad logician developing formal logic in a purely model-theoretic way, using no formal proofs at all. When we trace back through his logical ancestors, we find that the origin of his model-theoretic ideas is already visible in Aristotle's *Prior Analytics*, though Aristotle's version looks very different from the fourteenth century Baghdad version.

Technically this is an exciting discovery, particularly since the Baghdad logician combined this model theory with the invention of logical diagrams several hundred years earlier than anything similar in the West. But philosophers of logic may be disappointed, because the model-theoretic methods are not accompanied by any philosophical analysis of the different kinds of logical consequence. It seems to me that there is an implied philosophical analysis of a sort in the development of logic between Aristotle's Athens and fourteenth century Baghdad, though the development is scattered through several different texts. Key moments in this development were the elementary textbook of Paul the Persian in the sixth century, and logical experiments of Ibn Sīnā in eleventh century Persia. I will mention the work of Paul and Ibn Sīnā below, but it will take much more work than this paper allows to establish details of the flow of ideas.

2 Some Aristotelian background

The dominant logic, in the mid nineteenth century and for some two thousand years before that, was a form of logic based on Aristotle's categorical logic. We begin by outlining some differences and some similarities between Aristotle's logic and logics that are familiar today.

Aristotle worked with formal sentences: sentences where some of the words are replaced by letters. Starting from a formal sentence, we can get an ordinary sentence (called a 'material sentence' by some traditional logicians) by putting words or phrases in place of the letters. An 'interpretation' is a list of letters which assigns to each letter a word or phrase. If ϕ is a formal sentence and I is an interpretation, then we write $\phi[I]$ for the sentence got from ϕ by putting in place of each letter of ϕ the word or phrase assigned to it by I. For example at *Prior Analytics* 46b12–14 Aristotle gives the interpretation

(4) A: mortal animal. B: footed. C: footless. D: human.

If we write I for this interpretation, and ϕ for the sentence 'Some A is a D', then

(5) $\phi[I]$ is 'Some mortal animal is a human'.

For this definition to make sense, we require that the letters of ϕ are among the letters listed by I, and that the words or phrases assigned by I have a suitable form to be put in place of the letters of ϕ. We will always assume that these two requirements have been taken care of. Aristotle himself usually referred to an interpretation briefly as 'terms'. Some Roman Empire logicians called interpretations 'matter'—so that an ordinary sentence results from combining matter and the 'form' of a formal sentence.

We say that the interpretation I 'verifies' the formal sentence ϕ, and that I is a 'model' of ϕ, if $\phi[I]$ is a true sentence. We say that I 'falsifies' ϕ if $\phi[I]$ is a false sentence. If Φ is a set of formal sentences, we say that I is a 'model' of Φ if I is a model of every sentence in Φ. (The notion of a 'model' is not found explicitly in Aristotle.)

Aristotle is interested in what we can deduce from a pair of formal sentences; we will refer to an ordered pair of formal sentences as a 'premise-pair'. But here is our first major difference between Aristotelian logic and today's logic. For every premise-pair that he considers, there is a small finite number of formal sentences that Aristotle is willing to consider as possible logical consequences of the premise-pair; we will call these the 'candidate conclusions' of the premise-pair, or for brevity just the 'candidates'. The candidates are specified in terms of 'figures'; we will bypass the details. But for example the premise-pair

(6) No C is a B. Some B is an A.

has four candidates:

(7) Every C is an A. No C is an A.
Some C is an A. Not every C is an A.

You should be able to convince yourself that the two premises in (6) have the logical consequence 'Not every A is a C'. But this formal sentence is not one of the four candidates, because its letters are the wrong way round. So for Aristotle's successors (and for Aristotle himself most of the time) it is simply not counted as a consequence of (6).

The four sentence forms in (7) are typical examples of 'categorical' formal sentences; Aristotle's categorical logic is chiefly concerned with sentences of these forms. There are other kinds of formal logic, for example modal logic, but in this paper we will be mainly dealing with categorical logic.

The logical properties of a premise-pair and its candidates don't depend on the choice of letters; we can make a one-to-one replacement of the letters throughout the premise-pair and its candidates without affecting the logic. Two premise-pairs, together with their candidates, are said to have the same 'mood' if one comes from the other by such a one-to-one replacement of letters. For us it will be convenient to choose the letters so that the candidates always have first letter C and second letter A; the letter B can occur in the premises. This is near enough the convention followed by both Aristotle himself and the Arabic logicians.

Aristotle distinguishes two kinds of mood. For the first kind, at least one of the candidates is a logical consequence of the premises. In this case Aristotle says 'there is a syllogism', and he counts the strongest candidate that is a logical consequence as the 'conclusion' of the premise-pair. (A 'strongest' such candidate is a candidate that is a logical consequence of the premises and entails any other candidates that are logical consequences of the premises.) We say that moods or premise-pairs of this kind are 'productive'.

For the second kind of mood, none of the candidates is a logical consequence of the premise-pair. In this case Aristotle says 'there is no syllogism', and we will describe the mood or premise-pair as 'nonproductive'. Ibn Sīnā called nonproductive moods 'sterile', because he accepted a common Arabic fiction that a conclusion is a child whose parents are the two premises.

So for Aristotle and many other traditional logicians, the major problem of formal logic was to determine which premise-pairs are productive and which are non-productive, and in the productive case, to find the conclusion. Because of the restriction to candidates, these problems don't correspond exactly to any problems that today's logicians spend time on. For this reason a modern logician aiming to make sense of the relevant traditional logical texts must be prepared to do some lateral thinking.

In the definitions above, should 'logical consequence' be read as proof-theoretic or model-theoretic? Aristotle doesn't tell us. Some readers of Aristotle claim to know what Aristotle believed about this, but I make no such claim. It seems to me that one of the tasks that Aristotle's logic left for later generations of logicians was to get to a point where they could meaningfully formulate and compare different definitions of logical consequence.

3 Abū al-Barakāt

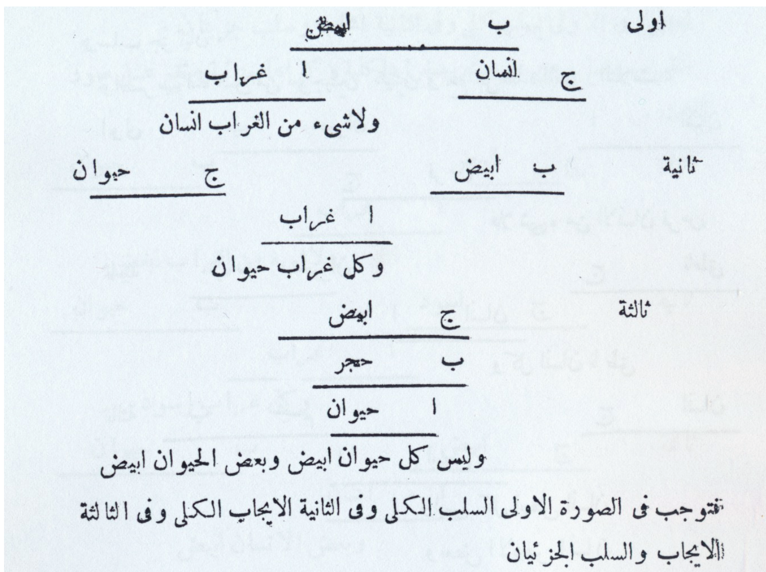
It's time to introduce our Baghdad logician. His name was Abū Barakāt bin Malka al-Baghdādī. He lived from some time in the 1080s to around 1170. He was a Jew, though late in life he converted to Islam. It was only recently realised that he must be the same person as the respected Talmudic scholar Rabbi Baruch ben Melekh (Gil [9] p. 469).

His major philosophical work, written in Arabic, was called *Kitāb al-mu^{c}tabar*, which can be roughly translated as 'A book about some things that I considered'. The book is encyclopedic; in fact Abū al-Barakāt used as a template for it Ibn Sīnā's encyclopedia *Shifā'* ('The Cure', written in the 1020s). The first part of Ibn Sīnā's book was on logic, and accordingly Abū al-Barakāt devotes most of the first three hundred pages of his book to logic. Later he moves on to physics, just as Ibn Sīnā did; the *Kitāb al-mu^{c}tabar* is said to be the first book to contain a statement that bodies fall at a constant rate of acceleration. The book also contains rich material on metaphysics and epistemology.

In the preface to a recent book on Abū al-Barakāt's philosophy, Moshe Pavlov ([18] p. ix) remarks that 'Abū'l-Barakāt's thought is most everywhere novel'. Certainly this is true of his approach to categorical logic. Even his choice of sentence forms was an adjustment of what logicians before him had used. But we will come to that later; our first interest is in how he organised his presentation.

Abū al-Barakāt recognises forty-eight moods of categorical logic. He lists them, and for each one he says whether it is productive or nonproductive, and if it is productive he gives its conclusion. (Towards the end of the list he starts to leave out some cases, evidently reckoning that he has said enough to allow the reader to work out the rest.) Together with each mood he gives evidence for his verdict on the mood. So far this is all standard; Aristotle did much the same in *Prior Analytics*.

But there is a striking novelty. In every case the evidence for the verdict consists of between two and four interpretations that are models of the premises. For nonproductive moods Abū al-Barakāt always gives three interpretations; for productive ones the number varies. He never tells us what we are supposed to do with these interpretations; he simply writes down each interpretation together with a picture of it in a notation of his own devising. The notation will be easier to make sense of when we understand what it is supposed to convey; so we will come back to it when we understand Abū al-Barakāt's interpretations. The obvious question is: how do you show that a premise-pair Φ is productive, or that it isn't, just by giving some interpretations that are models of Φ?



Another novelty is that in all of Abū al-Barakāt's interpretations the nouns or noun phrases assigned to the letters describe nonempty classes. We will describe interpretations with this property as 'nonempty'. Abū al-Barakāt's usage is in sharp contrast to Ibn Sīnā, whose interpretations contained phrases such as 'time or situation when there is a vacuum' (Ibn Sīnā believed that there never is a vacuum), or 'time or situation when something has infinite length' (ditto). Ibn Sīnā clearly intended these phrases to be descriptions of an empty class. Aristotle had authorised the use of examples of this kind when he invoked 'a time when nothing is moving except human beings' (*Prior Analytics* [20] i.14, 34b11f). I will assume provisionally that Abū al-Barakāt's interpretations were intended to be nonempty. A number of things will fall into place around this assumption.

Since Abū al-Barakāt's justifications for his verdicts consist entirely of interpretations, we have to suppose that he is using a notion of logical consequence that can be expressed in terms of interpretations. So we should see whether his examples fit the most obvious such notion, granting that he is using nonempty interpretations:

(8) Given a premise-pair Φ and a candidate θ, we count θ as a logical consequence of Φ if and only if every nonempty model of Φ is also a model of θ.

This defines a form of model-theoretic consequence.

If Abū al-Barakāt is using this notion of logical consequence, what does he need to do?

First suppose his task is to prove that Φ is *nonproductive*, where Φ has four candidates $\theta_1, \ldots, \theta_4$. Then Abū al-Barakāt must produce interpretations $I_1, \ldots, I_4$, not necessarily all distinct, such that each of these interpretations is a model of Φ, and for each i from 1 to 4, I_i falsifies θ_i. As we noted, Abū al-Barakāt in fact gives three interpretations for each nonproductive premise-pair, and we will see below that these three between them do always falsify all four of the candidates.

Second, suppose Abū al-Barakāt has to prove that Φ is *productive*, and to find its conclusion. In this case it seems that he has to consider *every* nonempty model of Φ, and show that all of them are models of some fixed candidate θ. This is one place where a glimmer of background theory peeps out, because Abū al-Barakāt has an Arabic word for 'is fixed or persistent', namely *istamarra*, and he seems to use it in the right places. But a major problem for us it that if Φ has any models at all, there are bound to be indefinitely many models of it, because Abū al-Barakāt has the whole of Arabic available as a source for the words or phrases used in his interpretations. So how can he hope to establish productivity by giving fewer than five interpretations?

We turn to the details.

3.1 The nonproductive case

Abū al-Barakāt's categorical logic has four sentence forms, which we assume are understood as in the formulas below.

(9) 'Every C is an A', $\forall x\, (Cx \rightarrow Ax)$.

(10) 'No C is an A', $\forall x\, (Cx \rightarrow \neg Ax)$.

(11) 'Some C is an A', $\exists x\,(Cx \wedge Ax)$.

(12) 'Not every C is an A', $\exists x\,(Cx \wedge \neg Ax)$.

We note at once that an interpretation verifies (9) if and only if it falsifies (12), and vice versa; and that an interpretation verifies (10) if and only if it falsifies (11), and vice versa. Also every nonempty interpretation that verifies (9) verifies (11) too, and every nonempty interpretation that verifies (10) verifies (12) too, but both these implications fail if we leave out the condition 'nonempty'. No nonempty interpretation verifies both (9) and (10).

After some slight adjustments of Abū al-Barakāt's choice of letters, each of his premise-pairs consists of a sentence whose letters are C and B (not necessarily in that order) and a sentence whose letters are B and A (again not necessarily in that order). The candidates are always the four formal sentences (9)–(12), with the letters C and A in that order.

We can justify Abū al-Barakāt's treatment of nonproductive moods by the following two facts.

Fact 1 *Suppose I is a nonempty interpretation for the letters C and A. Then I falls into exactly one of the following three types:*

Type One verifies 'Every C is an A'.
Type Two verifies 'No C is an A'.
Type Three verifies 'Some C is an A' and 'Not every C is an A'.

Fact 2 *Suppose Φ is a premise-pair in Abū al-Barakāt's categorical logic. Then the following are equivalent:*

(a) Φ is nonproductive.

(b) Φ has nonempty models of all three types.

The 'nonempty' is needed for the first fact, because an interpretation with C empty will verify both 'Every C is an A' and 'No C is an A'. It is also needed for (b) $\Rightarrow$ (a) in the second fact; without it we have the counterexample

(13) Every B is a C. No B is an A.

which by the definition (8) has the conclusion 'Some C is an A', though (13) also has models of all three types if we allow models in which B is empty.

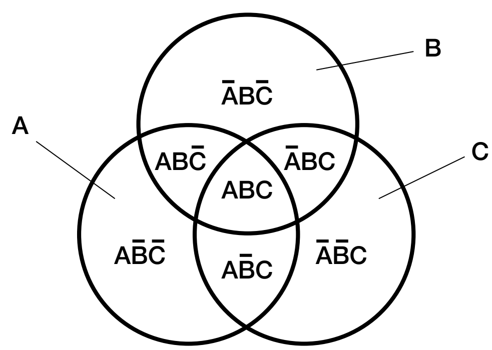
Abū al-Barakāt's procedure, for each of his nonproductive premise-pairs Φ, is to give three models of Φ, the first of Type One, the second of Type Two and the third of Type Three.

How does he find these three models? We don't know, and quite possibly he just uses trial and error. But for a properly algorithmic procedure he would need some way of listing systematically a collection of interpretations that is bound to contain at least one of each type. This problem of systematic listing will come up with the productive premise-pairs too. We turn to these.

3.2 The productive case

The first problem is to reduce the indefinitely large space of possible interpretations to a manageable set that still contains all the required possibilities. We can get a hold on this problem by drawing a picture to represent an arbitrary interpretation. In the picture below, the circle labelled A represents the class described by the noun or phrase assigned to A, and likewise with the other letters. Putting a bar over the top, $\overline{A}$ represents the class of things that are not in the class A; then for example the area labelled $\overline{A}B\overline{C}$ represents the class of things that are in B but not in either A or C.

(14)



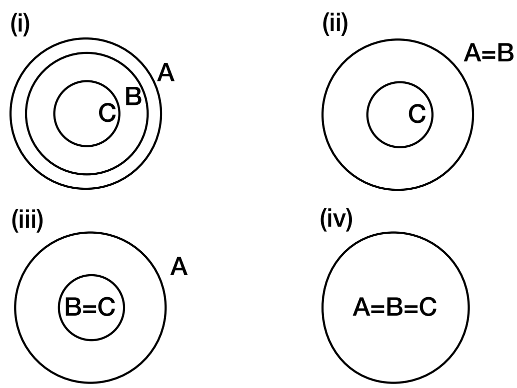
Fact 3 *Suppose the diagram above represents the interpretation I, and ϕ is a formal sentence of Abū al-Barakāt's categorical logic, using letters from among A, B and C. Then the question whether I verifies ϕ is completely determined by knowing which of the seven enclosed labelled areas in the diagram are empty and which are not.*

In the jargon of model theory, we say that two interpretations are 'elementarily equivalent' if they agree on which of the seven labelled areas they make empty. It follows from Fact 3 that when we are listing models of a premise-pair to see which candidates they verify, we need only list enough models to cover the $2^7 = 128$ elementary equivalence classes. This is a radical improvement, but it is still too large a number for it to be feasible to run through the lot every time we want to prove something. We can cut down a little, from 128 to 109, by excluding empty interpretations.

In practice we often get a huge drop in the number of models to consider when we restrict to models of the premise-pair Φ. For example let Φ be the premise-pair

(15) Every C is a B. Every B is an A.

Then everything in C is in B, but there are two possibilities according as C equals B or is a proper subclass of B. There are the same two possibilities for the relationship between B and A. This gives $2 \times 2 = 4$ possible models to consider:



For this premise-pair Abū al-Barakāt gives four interpretations that correspond to these four pictures:

(16)
(i) C = human, B = animal, A = body.
(ii) C = human, B = animal, A = capable of perceiving.
(iii) C = human, B = rational, A = capable of perceiving.
(iv) C = human, B = rational, A = capable of laughing.

(You might query some of these items. But Abū al-Barakāt believed, like Ibn Sīnā, that all and only human beings are rational.) Inspection shows that each of these four models verifies 'Every C is an A', so Abū al-Barakāt duly reports this as the conclusion.

The next productive premise-pair that he considers,

(17) Every C is a B. No B is an A.

needs just two models, and Abū al-Barakāt gives two models that exactly do the job. Again inspection finds the conclusion that Abū al-Barakāt reports, namely 'No C is an A'. But the third one that he considers,

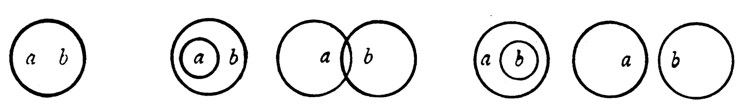
(18) Some C is a B. Every B is an A.

needs sixteen models and Abū al-Barakāt gives only four of them. The fourth needs twelve and again he gives only four of them.

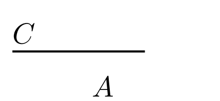
One can sympathise with Abū al-Barakāt not wanting to set out all sixteen possibilities. But is there any particular reason why he gave up at four? Not as far as I can see. *If* he already knew that the premise-pair (18) was productive, he would be entitled to stop as soon as only one of the candidates was *mustamirr*, i.e. persistently verified by every model examined so far. But in fact, if I read his assumptions correctly, he stopped before reaching any model that falsifies 'Not every C is an A', which is not the conclusion. So he probably just reckoned four was enough to show willing and to give the reader the idea of what was needed.

3.3 The pictures

Abū al-Barakāt accompanies each interpretation with a line diagram, which he calls a 'figure' (*shakl*) or a 'representation' (*tamthīl*). So the diagrams represent interpretations up to elementary equivalence; they don't represent sentences. This makes them different from the diagrams of Euler and Venn, and more like those that Gergonne [8] gave in 1816/7:

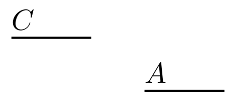
(19)


(I take the pictures from Faris [6]. Gergonne himself didn't draw these diagrams, but he described them very clearly in French prose.) Abū al-Barakāt aims to do the same thing but with lines instead of circles. Thus Gergonne's third diagram, representing that the two classes straddle each other, appears in the *Kitāb al-mu^{c}tabar* as

(20)


Gergonne's fifth diagram, representing that the two classes are disjoint, appears in the *Kitāb al-mu^{c}tabar* as

(21)



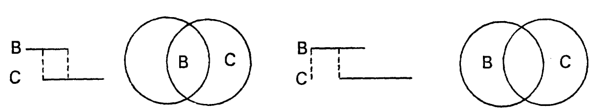
Let me make three preliminary observations about these line diagrams of Abū al-Barakāt.

First, there is no natural way of representing empty classes in Abū al-Barakāt's notation. Since this is the first time that such a notation appears in the literature, and also the first time we meet a logician who seems to be systematically avoiding the use of empty classes in his interpretations, we can reasonably ask if these two facts are related. Did Abū al-Barakāt reach his procedures by experimenting with the diagrams, so that the avoidance of empty classes was built into the procedures from the outset?

Second, Abū al-Barakāt differs from Gergonne in using these diagrams to represent relationships between three classes, not just two. If one were to generalise Gergonne's figures to three classes, one would need 109 figures rather than the five in (19) above. (And for the record, for four classes Gergonne would need 32,297 figures. There is a double exponential involved.) Not all of the 109 diagrams for three classes can be flattened down into line diagrams. For example there is no horizontal line representation of an interpretation where all seven labelled areas of (14) are empty except for $\overline{A}BC$, $A\overline{B}C$ and $AB\overline{C}$. Interpretations with this property are never needed for any of Abū al-Barakāt's productivity or nonproductivity proofs, luckily for him.

Margaret Baron [3] reviews the early history of logic diagrams, and mentions that Leibniz experimented with line diagrams. She cites two examples:

(22)



Some B is C Some B is not C

Both these examples are intended to represent sentences, not elementary equivalence classes of models; this is the reason for the vertical dashed lines. Baron also points out that Ramon Llull already had some kind of circle diagrams in the 13th century; but these were intended for some not entirely clear religious purpose, and their logical content is dubious. Abū al-Barakāt's line diagrams were undoubtedly part of a logical procedure. He takes the history of logic diagrams back to the twelfth century.

Third, the only publications known to me of the relevant passages in *Kitāb al-mu^{c}tabar* are taken from the edition published in Hyderabad in 1938/9. In this edition some of the diagrams fit exactly with the style illustrated by (20) and (21), and the others have at least a statistical tendency in that direction. But quite a few of the diagrams seem to be drawn at random. One possible explanation is that the copyists or printer had no idea what the diagrams were about. This seems the best explanation of the errors in Naṣīr al-Dīn al-Ṭūsī's account of these diagrams in his 13th century Persian text *Asās al-Iqtibās* [22] pp. In this edition there are some gross errors in the printing of the diagrams; also (granting the weakness of my Persian) it is not clear that Ṭūsī understood the diagrams—he seems to think they represent sentences.

So either the copyists and printer made a shipwreck of Abū al-Barakāt's diagrams, or he himself had a remarkably hippie attitude towards his own invention. Sight of a manuscript of the *Kitāb* might clarify the situation, but I have not so far managed to track one down. (I hope to pursue this issue further, together with other questions about Abū al-Barakāt's diagrams, in joint work with Amirouche Moktefi.)

Before we leave Abū al-Barakāt, there are a couple of observations to make on the theory behind his diagrams. First, we have seen that everything fits together neatly under the assumption that he intended all his interpretations to be nonempty. As far as I know, he was the first logician to assume that in categorical logic we are dealing with nonempty classes.

Second, Abū al-Barakāt makes an interesting remark about the productive premise-pair

(23) Not every C is a B. Every A is a B.

He says that this premise-pair needs two models (falsely—it actually needs twelve). He observes (correctly) that all the models verify 'Not every C is an A', so that this is the conclusion. (Again correct; one of his diagrams rules out the stronger candidate 'No C is an A'.) Then, atypically, he compares with Aristotle's procedure. For this premise-pair Aristotle uses a justifiable but slightly troublesome procedure called 'ecthesis', which involves introducing a fourth letter. Abū al-Barakāt sketches Aristotle's proof. Then he comments 'The representation with diagrams explains the conclusion without needing any of that' ([1] 139.14f).

4 Ibn Sīnā

Was Abū al-Barakāt's use of model theory entirely his own idea, or did he build on work of earlier logicians?

We are in speculative territory here. But for the nonproductivity proofs there are strong indications of a similar movement of thought in Ibn Sīnā's logical writings, and particularly in section vi.2 of the book *Qiyās* ([16], 'Syllogism') of his *Shifā'*. The *Shifā'* is the work that Abū al-Barakāt used as his template for the *Kitāb al-mu*^{c}*tabar*, so there is every reason to think that Abū al-Barakāt knew this work well.

This section is one of the richest and most radical of Ibn Sīnā's contributions to formal logic. Ibn Sīnā is introducing a logic of his own invention, extending the hypothetical logic that al-Fārābī had reported from sources in Roman Empire logic. The description below is an oversimplification, cutting through to what is relevant to Abū al-Barakāt's logic. The papers [13] and [14] study this section of *Qiyās* more directly, referring to the logic involved as PL3.

PL3 can be seen as a fragment of boolean algebra, using eight sentence forms:

(24)
$$\begin{array}{cccc} C \subseteq A & C \subseteq \overline{A} & \overline{C} \subseteq A & \overline{C} \subseteq \overline{A} \\ C \not\subseteq A & C \not\subseteq \overline{A} & \overline{C} \not\subseteq A & \overline{C} \not\subseteq \overline{A} \end{array}$$

where $\overline{C}$ is the boolean complement of C. The candidates for a premise-pair are precisely the eight sentences above, with letters C and A in that order. Ibn Sīnā lists premise-pairs and indicates which are productive (and with what conclusion), and which are sterile. He discusses roughly sixty premise-pairs in detail. Some forty of the ones that he discusses are productive. Seventeen are sterile (though he mistakenly takes one of these to be productive).

For the productive premise-pairs, he derives a conclusion using a proof theory based on Aristotle's proof theory for categorical logic. This part of the logic is set out clearly and explicitly; the only mistakes are occasional carelessnesses.

For the sterile premise-pairs, Ibn Sīnā's *language* indicates that he is following the method that Aristotle devised for proving nonproductivity in categorical logic. But this method is invalid for PL3, and Ibn Sīnā's *calculations* can be read as following a different method. This other method is sound, and he carries it through successfully in about two of every three cases. It is also thoroughly model-theoretic. But in view of both IS's misleading descriptions and the cases where his calculations seem not to work, there has to be a doubt whether he really knows what he is doing.

4.1 Aristotle and the method of pseudoconclusions

Let me explain Aristotle's procedure for proving nonproductivity in categorical logic. Aristotle works with four main sentence forms, which are often read as follows:

(25) 'Every C is an A', $(\forall x\,(Cx \to Ax) \wedge \exists x Cx)$.

(26) 'No C is an A', $\forall x\,(Cx \to \neg Ax)$.

(27) 'Some C is an A', $\exists x\,(Cx \wedge Ax)$.

(28) 'Not every C is an A', $(\exists x\,(Cx \wedge \neg Ax) \vee \forall x \neg Cx)$.

The formulas express a reading of Aristotle's sentences that was first made explicit by Al-Fārābī in the 10th century and more fully by Ibn Sīnā in the 11th century. Stephen Read has argued recently [19] that these formulas are faithful to Aristotle's intentions.

Aristotle establishes the productivity of a premise-pair by stating its conclusion, and then either giving a formal proof of the conclusion from the premises, or stating that the conclusion is self-evidently a consequence of the premise-pair so that no proof is needed.

Aristotle proves the nonproductivity of a premise-pair Φ as follows. The candidates are the four formal sentences (25)–(28), with the letters C and A in that order. We write down two interpretations I and J such that

(a) Both I and J are models of Φ.

(b) I verifies 'Every C is an A'.

(c) J verifies 'No C is an A'.

A modern understanding of Aristotle's method is to observe that the given data show that there is no candidate θ such that

(29) Every model of Φ is a model of θ.

The reasoning is as follows. Suppose θ is either 'No C is an A' or 'Not every C is an A'. Then the model I of Φ falsifies θ since I verifies 'Every C is an

A' (which now includes the clause $\exists x Cx$). Alternatively suppose θ is either 'Every C is an A' or 'Some C is an A'. Then the model J of Φ falsifies θ since J verifies 'No C is an A'.

Now (29) is a definition of model-theoretic consequence, the same as our earlier definition (8) except that in (29) we impose no requirement that the models are nonempty. So in this revised sense, Aristotle's data show that no candidate is a model-theoretic consequence of Φ. To complete the argument we need only suppose that whatever Aristotle understands by θ being a syllogistic consequence of Φ, it implies that θ is a model-theoretic consequence of Φ. (Some writers paraphrase this supposition as 'Aristotle's notion of consequence is truth-preserving'.)

However, Aristotle himself never mentions model-theoretic consequence. All he requires is that I verifies the sentence 'Every C is an A' and J verifies the sentence 'No C is an A'. In Aristotle's own terminology these two sentences are 'conclusions' got from I and J; to avoid confusion with logical conclusions I refer to them as (formal) 'pseudoconclusions' for the two interpretations. These notions and the relevant texts are analysed further in [11] and [14].

4.2 Ibn Sīnā and the method of complementary models

Ibn Sīnā claims to show that premise-pairs in PL3 are sterile by giving two models of Φ, one I for which $C \subseteq A$ is a pseudoconclusion and one J for which $C \subseteq \overline{A}$ is a pseudoconclusion. But this won't work. The interpretation I falsifies $C \nsubseteq A$, but not any of the other candidates. Likewise J falsifies just one candidate. Since there are now eight candidates to be falsified, the arithmetic doesn't add up.

But there is an alternative reading of what Ibn Sīnā is doing. Forget about the pseudoconclusions. Each interpretation verifies exactly four of the candidates and falsifies the remaining four. So if Ibn Sīnā can find models I and J such that I verifies exactly the candidates that J falsifies, then every candidate will be falsified by one of them, and hence we will have a direct proof that no candidate is a model-theoretic consequence of Φ. We describe such a pair of interpretations as 'complementary'.

At this point we must go to the sixteen pairs of interpretations that Ibn Sīnā offers in his sixteen sterility proofs, and check whether they are in fact complementary models of the relevant premise-pairs. The results can be found in [14]. In sum, five of the sterility proofs seem to be fatally flawed. Ten can be read as valid sterility proofs along the lines we sketched in the previous paragraph, and the remaining one proof can be fairly easily repaired to form a valid proof. This is a much better result than chance, but also much less than 10 out of 10 for Ibn Sīnā.

There is a twist that should be mentioned. For the sterile premise-pair

(30) $$C \subseteq \overline{B}. \quad B \subseteq A.$$

Ibn Sīnā gives the two interpretations

(31)
I: C = human, B = stone, A = mineral.
J: C = human, B = stone, A = bodily object.

Certainly I and J are models of the premises. But consider the candidate

(32) $$\overline{C} \subseteq \overline{A}, \text{ i.e. } A \subseteq C.$$

If I and J are a complementary pair, then exactly one of them verifies this sentence. It seems to me that this requires that exactly one of these two interpretations has its universe restricted to humans. For most of Ibn Sīnā's rescuable sterility proofs, rescuing them mainly consists of assigning universes to interpretations.

Now Ibn Sīnā never mentions restriction of the universe. But in some cases it is implied, for example when in another logic he restricts the interpretations of certain variables to be times or situations. It is also very natural to restrict the universe when one is taking complements of classes; this is exactly the context in which De Morgan introduced universes into modern logic ([5] p. 2f). The domains of structures in today's model theory owe something to De Morgan's 'universes', though probably the domains of algebraic structures in Grassmann, Dedekind and Weber were a stronger influence.

In sum, Ibn Sīnā's sterility proofs in *Qiyās* vi.2 can mostly be read as applications of model-theoretic consequence in a rather full-blooded sense—the models, with their universes, can be taken as models in a fully modern sense. But this reading attributes to Ibn Sīnā more understanding than he probably had. Nevertheless there is enough sense in Ibn Sīnā's sterility proofs to allow the possibility that they helped to inspire Abū al-Barakāt's work in categorical logic.

There are other things in Ibn Sīnā that might have pointed Abū al-Barakāt in the same direction.

5 Paul the Persian

In the middle of the 6th century Paul the Persian published an elementary introduction to logic [17]. There are strong indications that this treatise is a summary of the introductory logic course taught at the School of Alexandria; this school had a proud history as one of the main philosophical centres of the Roman Empire. (See for example Gutas [10] on Paul the Persian and the curriculum at Alexandria.)

The final sections of Paul's treatise are about categorical syllogistics. He lists the moods, and as he comes to each mood he tells us whether it is productive (and with what conclusion) or nonproductive. So far this is exactly what both Aristotle and Abū al-Barakāt do. But there are major differences, particularly in his treatment of the productive moods. Unlike Aristotle, Paul doesn't accompany the productive moods with either proofs or statements that they are self-evident. He does tell us Aristotle's proofs for the productive moods, but he postpones these to a final section of his treatise, after the main listing of the moods. Unlike Abū al-Barakāt, Paul doesn't offer model-theoretic arguments to support the productive moods.

Paul adds a feature which is very suggestive for the topic of this paper. He presents the productive moods and the nonproductive ones in a way designed to make them seem parallel. For example *every* premise-pair has a conclusion. In the productive case there is just one conclusion, and it is the logical conclusion from the premises; taking a word from the *Prior Analytics*, Paul calls it a 'necessary' conclusion. In the nonproductive case there are two conclusions, and they are what we called pseudoconclusions in the previous section; Paul calls them 'non-necessary' conclusions. Each conclusion is accompanied by an interpretation that is a model of the premises; in the productive case the interpretation is just for illustration, while in the nonproductive case the two interpretations generate the pseudoconclusions.

For any reader able to think below the surface of the text, there is a clear message here. Every premise-pair has a family of candidates that are verified by models of the premise-pair; the difference between the productive and the nonproductive case is that in the former, the same candidate is verified by all models. From this conclusion it's a natural step to try to find the 'necessary' conclusions by listing models, just as Abū al-Barakāt does. But as far as we know, Abū al-Barakāt was the first person to take this step, and we don't know what role Paul the Persian or the Alexandrian tradition played in Abū al-Barakāt's thinking.

Though Paul certainly downgrades the role of formal proof in categorical logic, it's very striking how close he stays to the *Prior Analytics* in both the terminology and the logical content of his treatment of the categorical moods. This fact should give pause to those (for example Hugonnard-Roche [15] p. 272f) who see in Paul's treatment of logic a move away from Aristotle's 'formal' theory and towards a more 'material' understanding. One might equally well label as unaristotelian the accounts of Aristotle's logic that stress his formal proofs and ignore his nonproductivity arguments.

In fact both proof-theoretic consequence and model-theoretic consequence are at least implicit in Aristotle's presentation of logic. Arguably the main difference between the two in his treatment is not that he is more attached to proofs than to models. Rather it is that proofs call on less background theory; everything is there on the page. With interpretations and models, by contrast, one needs to have a good understanding of metatheoretical notions such as truth, interpretation or the definition of $\phi[I]$ in Section 2 above. These metatheoretical notions took a long time to mature. As late as 1906 Gottlob Frege was declaring that proofs of 'the independence of a thought from a group of thoughts' were a move into 'new territory' for which there was not yet any 'mathematical' theory available ([7] p. 425f). Eight hundred years before Frege, Abū al-Barakāt clearly knew what he was doing with interpretations and models, but he had no vocabulary for explaining it.

References

[1] Abū al-Barakāt, *Kitāb al-mu^{c}tabar fī al-ḥikmat al-ilāhiyya*, Byblion, Jbeil Lebanon 2007.

[2] Aristotle, *Prior Analytics Book I*, trans. and commentary by Gisela Striker, Clarendon Press, Oxford 2009.

[3] Margaret E. Baron, 'A note on the historical development of logic diagrams: Leibniz, Euler and Venn', *The Mathematical Gazette* 53 (384) (1969) 113–125.

[4] Bernard Bolzano, *Grundlegung der Logik (Wissenschaftslehre I/II)*, ed. Friedrich Kambartel, Meiner, Hamburg 1963.

[5] Augustus De Morgan, 'On the syllogism: I, On the structure of the syllogism', *Transactions of the Cambridge Philosophical Society* 8 (1846)

379–408; in Augustus De Morgan, *On the Syllogism and Other Logical Writings*, ed. Peter Heath, Routledge, London 1966, pp. 1–21.

[6] J. A. Faris, 'The Gergonne relations', *Journal of Symbolic Logic* 20 (3) (1955) 207–231.

[7] Gottlob Frege, 'Über die Grundlagen der Geometrie', *Jahresbericht der Deutschen Mathematikervereinigung* 15 (1906) 293–309, 377–403, 423–430.

[8] Joseph D. Gergonne, 'Variétés. Essai de dialectique rationnelle', *Annales de Mathématiques Pures et Appliquées* 7 (1816/7) 189–228.

[9] Moshe Gil, *Jews in Islamic Countries in the Middle Ages*, Brill, Leiden 2004.

[10] Dimitri Gutas, 'Paul the Persian on the classification of the parts of Aristotle's philosophy: a milestone between Alexandria and Baghdad', in Dimitri Gutas, *Greek Philosophers in the Arabic Tradition*, Ashgate, Aldershot 2000, pp. 231–267.

[11] Wilfrid Hodges, 'Nonproductivity proofs from Alexander to Abū al-Barakāt: 1. Aristotelian and logical background', draft online.

[12] Wilfrid Hodges, 'Nonproductivity proofs from Alexander to Abū al-Barakāt: 2. Alexander and Paul', draft online.

[13] Wilfrid Hodges, 'Identifying Ibn Sīnā's hypothetical logic: I. Sentence forms', draft online.

[14] Wilfrid Hodges, 'Identifying Ibn Sīnā's hypothetical logic: II. Interpretations', draft online.

[15] Henri Hugonnard-Roche, *La Logique d'Aristote du Grec au Syriaque: Études sur la Transmission des Textes de l'Organon et leur Interprétation Philosophique*, Vrin, Paris 2004.

[16] Ibn Sīnā, *Al-qiyās (Syllogism)*, ed. S. Zayed, Cairo 1964.

[17] Paul the Persian, *Logica*, in Jan Pieter Nicolaas Land ed., *Anecdota Syriaca* Vol. 4, Brill, Lugdunum Batavorum (Katwijk) 1875, pp. 1–30.

[18] Moshe M. Pavlov, *Abu 'l-Barakat al-Baghdadi's Metaphysical Philosophy, The Kitab al-Mu^{c}tabar*, Routledge, Abingdon 2017.

[19] Stephen Read, 'Aristotle and Lukasiewicz on existential import', *Journal of the American Philosophical Association* 1.3 (2015) 535–544.

[20] W. D. Ross, *Aristotle's Prior and Posterior Analytics: A Revised Text with Introduction and Commentary*, Clarendon Press, Oxford 1949.

[21] Alfred Tarski, 'On the concept of logical consequence', translated in Alfred Tarski, *Logic, Semantics, Metamathematics,* J. Corcoran (ed.), Hackett, Indianapolis 1983, pp. 409?-420.

[22] Naṣīr al-Dīn al-Ṭūsī, *Asās al-iqtibās* (Persian), Ferdows, Tehran 2010.