Detecting the Logical Content: Burley's 'Purity of Logic'

Wilfrid Hodges

(to appear)

In one of his inaugural lectures Dov Gabbay demonstrated God's providence as follows—I quote from memory. God had seen fit to send his revelation to mankind in a language with no verb tenses. (In classical Hebrew the verb forms are distinguished as perfective, imperfective, passive, reduplicative, etc., but not as past, present or future.) So there was a problem to explain how facts about a temporal world can be expressed in a language with no tenses. Logicians could not only solve this problem; they could also apply for research grants for solving it. Thus God ensured that logicians would not go short of spare cash.

As a hardened atheist, what can I rescue from this story? Actually, quite a lot. In the history of logic we study texts that are written in a language generally quite different from our own, on subjects that supposedly have something to do with what we understand as logic. The links are often hard to make.

One thing that historians of logic do is to describe the deductive practices of earlier thinkers. In this sense Netz's [13] close study of Euclid's procedures is an important contribution to the history of Greek logic. Here I try to make a small step in the same direction with a medieval logician. But *reconstructing practices is not enough.* Netz remarks [13, p. 216]:

> One of the most impressive features of Greek mathematics is its being practically mistake-free.

The point is obvious but still worth making: in mathematics some things are right and some are wrong, and we can often tell which. For example any procedure, using any system of concepts and representations, is just wrong if it yields the conclusion that π is exactly $22/7$. This makes a sharp division between the history of mathematics and the history of philosophy.

Logic comes somewhere between the two, and another purpose of this essay is to test the question: Are logician X's procedures correct for what he was aiming to do? To answer the question we need to be able to discern the aims behind the procedures, and then we need to be able to tie these aims to something objective. Of course the next generation's notion of what is objective in logic may differ from ours, just as it sometimes does in mathematics. So this kind of analysis of the past needs to be redone in each generation.

To make things as concrete as possible, I chose a particular text, 'On the Purity of the Art of Logic' (*De Puritate Artis Logicae Tractatus Longior*), PL for short, written by Walter Burley in the late 1320s. Theophilus Boehner [3] has edited the Latin original, and Paul Vincent Spade gives a translation [4]. Spade helpfully numbers the paragraphs (1) to (1053), and my references follow this numbering.

I chose Burley because he is the beneficiary of the preceding century and a half's work by terminist logicians, but he is mercifully free from the lapses of common sense and the metaphysical irrelevances that disfigure much of later Western medieval logic.

This essay was written in a hurry against a publication deadline, so that some details will certainly be wrong. I intend to put on my website at `www.maths.qmul.ac.uk/~wilfrid` the full reference list mentioned in Section 1, together with corrections as I find them.

1 The book

The backbone of Burley's PL consists of a large number of inferences expressed by Latin sentences; each inference is labelled either good or not good. Burley calls the inferences 'consequences' (*consequentiae*). Listing them I found 128 consequences labelled 'good' and 178 labelled 'not good', making 306 in all; below I refer to this as the *reference list*. The number is not exact, for several reasons. Some consequences are parts of Burley's own argumentation, not inferences that he is discussing. Some consequences are expressed in irregular ways. Some consequences are broken up into more than one consequence during their discussion. In some consequences Burley uses a letter as an abbreviation for a certain piece of Latin text; I included these. But I left out all consequences that contain schematic letters, since these are strictly not consequences but rules defining classes of consequences.

Burley never mentions that the sentences in his consequences are in

Latin. (Neither of the words *Latinus* or *lingua* appears in PL.) Nevertheless we need to mention it. Some of Burley's rules rely on facts or alleged facts about Latin that don't hold for other languages. These facts are of three kinds: (a) the use of word order for determining scope, (b) the use of noun cases for determining subjects, (c) the possible antecedents of anaphoric pronouns.

(a) Burley's rule of thumb for scopes is that an expression comes at the start of its scope. Thus (107)

> A negation doesn't include in its scope what precedes it.
> . . . *negatio non habet dominium supra praecedens.*

This is false in Burley's own first language of English, where we negate a finite verb by putting 'not' *after* it. Thus for example ([1] p. 143) from about 1300:

> Bot al men can noht, I-wis,
> Understand Latin and Frankis.

Here the 'noht' follows not only the verb 'can' but also the quantifier 'al' which clearly lies within its semantic scope. (John Marenbon suggested to me that Burley might have dismissed English as a low status language.)

Translating *dominium* as 'scope' may convey the impression that Burley has a technical term here. But in fact this is the one occurrence of the word *dominium* in PL, and mostly we have to infer Burley's view of scope from his practice. For example at (98) he distinguishes between

> All day someone is indoors here.
> *tota die aliquis homo est hic intus.*
>
> Someone is all day indoors here.
> *aliquis homo tota die est hic intus.*

He reads the sentences in the way that we would describe by saying that the earlier quantifier has wider scope.

(b) Among the many logical 'rules' that Burley lists, one (662) refers explicitly to nominative case. The rule says that 'Only an A is a B' is equivalent to 'Every B is an A' when the subject A is in the nominative.

His explanation (643) is more puzzling than the rule it explains. The explanation is that without the restriction 'nominative' (*rectus*) the rule would give us that 'Any man's is a donkey' (*cuiuslibet hominis est asinus*) is equivalent to 'Only a donkey is man's', which he says is wrong because 'Any man's is a donkey' would be true if each man has both a donkey and a bull.

Clearly he is reading ‘Any man’s is a donkey’ as meaning ‘Each man has a donkey’; so the logical predicate is not ‘donkey’ but ‘has a donkey’, and the correct converse would be ‘Only a thing having a donkey is a man’.

One might guess that he is using the genitive case of ‘man’ (*hominis*) as a heuristic to warn us that the logical subject is not necessarily the grammatical subject; at (169ff) he uses the case endings in this way. But of course this has only an indirect connection with the rules for ‘only’. Moreover it doesn’t show that his rule needs ‘nominative’, since the word in the genitive is not in the proposition with ‘only’. His reasoning here is too loose to allow any definitive correction.

At (701) he notes correctly that Latin nominalises sentences by putting the subject into the accusative (and the verb into the infinitive). Thus a sentence with another sentence as the subject can have what appears to be a subject in the accusative. English works differently; Burley’s *Deum esse Deum* comes into English as ‘that God is God’ or ‘for God to be God’.

(c) At (123) Burley states that a reflexive pronoun can have its antecedent either in the same clause or in an another clause. This is so far from the behaviour of English reflexives that I was surprised to find that Burley is right. There are classical examples, for instance Plautus [16] *Poenulus* I.ii:

> The lady can make a lump of flint fall in love with herself [sic].
> *Illa mulier lapidem silicem subigere ut se amet potest.*

Similar things are reported in Korean and Chechen. (Spade’s explanatory example ‘Socrates looked in the mirror and he saw himself’ [4] p. 113 is off the point; the antecedent of ‘himself’ is ‘he’ in the same clause.)

One sometimes hears it said that the later medieval Western logicians adopted a rigid form of Latin that was intended to serve as a formal language. I found nothing whatever in PL to support this view. Burley never once suggests that his readings are anything other than the correct readings of normal Latin. He does acknowledge that some of his readings disagree with linguistic usage (*usus loquendi*), but in such cases he insists that his reading is correct literally (*de virtute sermonis*) (for example (191), (192), (731), (741)).

In any case it’s hard to see what place a formal language would have in Burley’s scheme of things. It’s even harder to see how he could have saddled himself with a formal language as confusing as his rather stilted Latin. For example at (673) he reads the sentence

> *Tantum sciens grammaticam est homo.*

as false in the case where everybody knows both grammar and logic. This is mysterious. Spade ([4] p. 226 note 250) may have it right when he translates as

> Someone who knows nothing but grammar is a man.

But this includes *tantum* within the scope of *sciens*, contrary to the order of the words. If this is how to invent a formal language, it puts Burley in the same class as the man who allegedly first invented the vacuum cleaner except that he made it blow instead of suck.

2 Consequences

A consequence consists of one or more Latin sentences (called the 'antecedents', *antecedentes* or 'premises', *praemissae*), followed by 'Therefore' (*ergo*), followed by one Latin sentence (called the 'consequent', *consequens* or 'conclusion' *conclusio*). The parts of the consequence can be run together as a single sentence when convenient.

Among the 128 good consequences in the reference list, Burley phrases 65 in the style:

> It follows: A therefore B.
> *Sequitur* A *ergo* B.

Sometimes he adds 'well', apparently not meaning anything different:

> It follows well: A therefore B.
> *Sequitur bene:* A *ergo* B. (12 times)

Sometimes he says that the conclusion follows:

> If A, it follows that B.
> *Si* A, *sequitur quod* B. (1 time)
>
> With reference to A, B follows.
> *Ad* A *sequitur* B. (1 time)
>
> When A, it follows that B.
> *Cum* A, *sequitur quod* B. (1 time)
>
> From this: A, B follows.
> *Ex ista* A *sequitur* B. (1 time)
>
> From these premises: A, this conclusion follows well: B.
> *Ex istis praemissis* A *ista conclusio bene sequitur* B. (1 time)

(In the last two, ‘this’ plays the role of quotation marks, which Burley doesn’t have.) When the antecedents have been stated, he sometimes says:

> The conclusion follows well: B.
> *Bene sequitur conclusio* B. (8 times)

Sometimes he says about a consequence:

> It is a good consequence.
> *Est consequentia bona.* (5 times)
>
> The consequence holds.
> *Tenet consequentia.* (3 times)

A syllogism is a particular kind of consequence; sometimes Burley uses the same language as above but with ‘syllogism’ in place of ‘consequence’. For example we have *Syllogismus est bonus* 4 times.

Apart from the restriction to syllogisms, all the locutions above seem to be stylistic variants. When any of them are true of a consequence, I will say that the consequence is *good*. Burley describes the remaining consequences as ‘not good’, but I will shorten that to *bad*.

Compare the three items:

(i) If A then B. (*Si* A, B.)

(ii) With respect to A, B follows. (*Ad* A *sequitur* B.)

(iii) A, therefore B. (A *ergo* B.)

Here item (i) is not a consequence but a conditional. Burley is clear that conditionals and consequences are not the same thing. However, he notes that for each conditional (i) there is a corresponding sentence (ii), and he says (353f) that the sentence follows from the conditional and vice versa. Also since (ii) and (iii) seem to be stylistic variants, it appears that Burley takes all three forms as equivalent for purposes of reasoning. In (353) he describes the relationship between (i) and (ii) as being that (i) performs an act which is meant by (ii). This is less revealing than it might be, because in (706) he describes the relationship between (ii) and (iii) in exactly the same terms. (In Latin the distinction is between *actus significatus* and *actus exercitus*; see Nuchelmans [14] for the history of this distinction. I think the word *actus* here refers to the action performed by ‘If’ in relating two clauses, not to a speech act.)

In fact Burley's text shows some slippage between conditionals and consequences. For example at (308) he applies the word *sequitur* to a conditional, though the conditional is restated as a consequence a few lines later. At (311) he makes a statement about conditionals, but his example to illustrate it is a consequence. At (319) he speaks of an 'inference' being made in a conditional. At (69) and (153), *si* is answered by *ergo*.

3 'Follows'

Burley's text doesn't distinguish between primitive and derived notions. But he uses the word 'follows' (*sequitur*) hundreds of times, with no attempt to reduce it to any more basic notion. In particular he doesn't paraphrase it in such terms as 'If the premises are true then the conclusion must be true'. (He does state this as a necessary condition for a consequence to be good (258), but he never suggests it is sufficient.)

In fact a good deal of what he says about truth seems to presuppose some notion of following. For example he uses the goodness or badness of certain consequences in order to explain the circumstances in which a sentence is true. Section 14 below will report some bad consequences used this way. Sometimes Burley uses good consequences of the form

> 'If we suppose that' (*posito quod, supposito quod*) A; then (*tunc*) the sentence B is true.

to clarify the conditions under which a sentence is true. (Thus at (557), (593), (595), (643), (886); likewise at (98), (163), (507), (535), (594), (673), (735), (767), (776), (947) without *tunc*.) In one place (735) he says 'If we suppose that' A, 'it follows that' (*sequitur*) B is true. So this talk of 'supposing' is a way of expressing conditionals or consequences; but Burley seems to use it only in cases where the main interest lies in B, while A appears only as an example of circumstances in which B is true.

Another explanation of truth apparently in terms of consequence is Burley's theory of descent. At (82) he explains that the term 'man' supposits determinately (*supponit determinate*) in 'Any man runs', because from 'Any man runs' there follow 'Socrates runs' and 'Plato runs' and so on. We must come back to this later; but for the moment we note that the notion of suppositing determinately is defined in terms of the notion of certain things following. So if (as is sometimes claimed) the notion of suppositing determinately is intended to explain the conditions under which certain sentences are true, then for Burley the notion of 'following' is prior to the notion of being 'true under certain conditions'.

For 'prior' perhaps we should read 'not posterior'. There is little indication that Burley has his definitions in a hierarchy. Holism rules.

In any case we have to suppose that Burley expects his readers to come ready equipped with the notion of 'follows'. A corollary is that at least in simple cases, he must expect readers to be able to see for themselves that a consequence is good.

Many times through the book, Burley gives a rule for good consequences and then a consequence that is an instance of the rule. He doesn't tell us how to read the text, but two possible ways suggest themselves. One way is to read it forwards: we learn the rule, then we see how to use it to construct good consequences. In a classroom the teacher might invite the class to construct other instances of the rule, using the given instance as a template. Elementary mathematical texts often use a similar format today.

The second way of reading is backwards, and this is more interesting. We can suppose that the abler readers, already having the notion of 'follows', will be able to see for themselves that the example consequences are good. They can then generalise from these examples to see how the rules arise.

This way lies danger. Can we really infer a general rule from study of a single instance? That question looms large over the rest of this paper, because Burley often proceeds as if the answer is yes. To understand what he is doing, we will need to look at his procedures for proving universally quantified statements (and we will do this in Section 9).

Burley is not above justifying rules, or indeed anything else, by barefaced appeals to authority. For example at (345) he answers the objection that Boethius didn't mention a certain rule by retorting that it's in Aristotle; but this could be *ad hominem* against some old fogey. At (284) he says that a certain method of argument is correct because Aristotle used it. At (510) and (753) he says that something is his own opinion (*mihi videtur*, *ego dico tibi*), but calls in Aristotle to back him up. Generally his uses of Aristotle are benign—they give credit or aim to pull things together. For example at (32) and (957) he uses notions that he credits to Aristotle; at (763) he mentions that a method he uses is also in Aristotle.

In (228) he claims to report Aristotle's views on the meaning of the Latin word *est*. Historically this is bizarre. He may only mean that Aristotle's discussion of the Greek *estí* transfers to the Latin word; but one would have been happier if he had said so. (Al-Fārābī would have done.)

4 Propositions

For Burley every true proposition is true because of (*secundum quod*) something, known as its 'cause of truth' (*causa veritatis*). Whether or not this notion is prior to that of a good consequence, we can usefully treat it next, if only to introduce Burley's notions of sentence structure.

Just as an event can have several possible causes, a true proposition can have several possible causes of truth; any one of them is enough to make the proposition true. Burley confuses actual causes of truth with possible causes of truth. For example at (868) he says that the proposition

> It's not the case that every man except Socrates is running.

has two causes of truth, namely that somebody other than Socrates is not running, and that Socrates himself is running. Since the truth of either of these conditions would make the displayed sentence true, and these are the only things that could make it true, the sentence is equivalent to (*valet*) the disjunction

> Either some man other than Socrates is [not] running, or Socrates is running.

In this example it's clear that the second of the 'causes of truth' is not in fact true. In what follows I will speak of 'possible causes of truth' when the sentences in question are not assumed to be true. (Wittgenstein at 5.101 of the *Tractatus* [23] avoids Burley's potential muddle by defining explicitly: 'Diejenigen Wahrheitsmöglichkeiten seiner Wahrheitsargumente, welche den Satz bewahrheiten, will ich seine Wahrheitsgründe nennen.')

Some propositions are true because of two things. For example according to Burley (651) the proposition 'Only Socrates is running' is made true by the truth of 'Socrates is running' and 'Nothing other than Socrates is running'. Likewise 'Socrates and Plato are running' is made true by the truth of 'Socrates is running' and 'Plato is running'. In such cases Burley normally describes the analysing sentences as 'exponents' (*exponentes*). Just once (868) he gets in a muddle and calls them causes of truth. The reason is that he is discussing negations of conjunctions; such a negation is equivalent to a disjunction, so it has two possible causes of truth, namely the negations of the exponents of the conjunction, a point he makes again at (964).

For future reference, I note that the discussion at (651) and at (757f) seems to imply that for Burley the sentence ‘Every A is a B’ has exponent ‘There is an A’, but ‘Nothing other than a B is an A’ doesn’t. If this is right, then presumably for Burley ‘Every A is a B’ and ‘No A is a non-B’ are not equivalent. (But he should have told us explicitly.)

We can catalogue propositions according to their possible causes of truth. Burley talks of propositional combinations as having two propositions and a ‘principal’ (*principale*), namely the connective that joins them (328), (506); the two joined propositions are the ‘principal parts’ (*partes principales* (522)). Presumably we find the possible causes of truth by looking first at the principal, though Burley doesn’t spell this out. For example a sentence ‘S or T’ has two possible causes of truth, namely S and T. Probably the possible causes of truth of ‘S and T’ are the propositions formed by conjoining a possible cause of truth of S with a possible cause of truth of T. Probably similar ideas work for all propositional combinations. Modal propositions ‘It is necessary that . . . ’ etc. have the modal operator as principal (327).

There remain two kinds of proposition, namely atomic propositions like ‘Socrates runs’, and propositions that are compound but neither modal nor propositional compounds. Burley treats these two types together. In each case he looks for two ‘terms’, normally noun phrases, which are respectively subject and predicate. The sentence expresses something about how many or which of the objects described by the subject term have the property expressed by the predicate term. The things described by the subject term are called the *supposita*. The required connection between the supposita and the property expressed by the predicate is determined by the other words of the proposition (the *syncategoremata*), such as ‘is’ (the principal in such sentences) or ‘not’, or ‘every’ or ‘some’ (when they are attached to the subject). Propositions of this general form are called ‘categoricals’ (*categoricae*). Propositions superficially of this form but with other pieces attached are treated separately. Examples are exceptives like ‘Some man besides Socrates can laugh’, and exclusives like ‘Only a man is a donkey’.

For example a necessary condition for a universally quantified affirmative categorical to be true is (165) that the predicate ‘is in’ all the things contained under the subject (i.e. the supposita); I guess it becomes necessary and sufficient if we add ‘there is at least one suppositum’. This possible cause of truth is at meta-level since it talks about the subject and predicate; the possible causes of truth of a disjunction by contrast were at the same conceptual level as the disjunction itself. This may be one of the reasons for the theory of 'descent to supposita', which essentially resolves a universal affirmative into a family of exponents at the same conceptual level (though Burley doesn't put it this way). In descent we remove the quantifier and replace the subject term by a name of one of its supposita; the resulting proposition is called a 'singular' (*singularis*) of the categorical. For example one of the singulars of 'Every man is running' is 'Socrates is running'. If the categorical is true then so are all its singulars. The converse holds too; Burley doesn't state this explicitly in PL, though he refers to it at (143), (163f), (175), (179) and (183).

Existentially quantified categoricals also allow descent, but this time each of the singulars is a possible cause of truth, so that the categorical is equivalent to their disjunction. Thus for example from 'Some man runs' we reach 'Socrates runs or Plato runs or etc.' (82). I will call this 'disjunctive descent' to distinguish it from the simple descent associated with universal quantifiers.

Burley generalises the idea of descent to other terms besides the subject, and he often talks of whether it's possible to deduce the singulars (*contingit descendere*), either separately or their disjunction. The possible answers are described as forms of 'supposition' (*suppositio*, strictly *suppositio personalis*, though this refinement is irrelevant to the present essay). Burley uses these forms of supposition to classify occurrences of terms in propositions. The forms of supposition are defined in terms of whether certain inferences hold. For example (100) the statement that 'man' in 'Every man runs' has 'confused and distributive supposition' means that from this sentence we can deduce that Socrates runs, that Plato runs etc. (Again I ignore some subtleties about mobile and absolute supposition.)

When we apply descent to terms that aren't the subject, we generally have to forget about dropping the attached quantifier, because there isn't one. Burley passes over this in silence. There is a more serious problem when we apply descent to the subject term but the determiner is something more complicated than a simple quantifier. When we drop the quantifier, the rest of the determiner may make no sense attached to a proper name. For example at (101) and (873) Burley happily descends from 'Every man except Socrates is running' to 'Plato except Socrates is running'. Here it looks as if the notion of descent has overrun its usefulness. (Burley can't handle complicated determiners anyway. See at (732) his inability to make sense of the sentence 'Only three men are running', *tantum tres homines currunt*.)

Altogether this system of possible causes of truth, exponents and descent might look like the beginnings of a Tarski-Montague-style truth definition for medieval Latin. But there are some major differences. One is Burley's lack of any systematic ideas on how sentences are built up. Another is that in spite of appearances, possible causes of truth are not compositional—it's not true that each piece makes its own separate contribution to the possible cause of truth. For example (118):

> For the sentences 'A man is running' and 'He is debating' to be true [in 'A man is running and he is debating'], it's required that 'A man is running' should be verified for one of the supposita of 'man', and that the second part should be verified for the same suppositum.
> *Ad hoc enim quod istae sint verae: 'Homo currit', et: 'Ille disputat', oportet quod ista: 'Homo currit', verificetur pro aliquo supposito hominis, et quod secunda pars verificetur pro eodem supposito.*

Or for another example (207f), in a present tense sentence the supposita of the subject term are limited to things that the term describes now; but in a past tense sentence the supposita can include things that the term used to describe. Burley warns of this at (4) when he says that the supposition of a term is a property that it has in relation to another term in the proposition.

5 When are consequences good?

Just as true sentences are true because of some *causa veritatis*, so good consequences are good because of some rule. Burley says (341):

> Every good consequence holds through some place, i.e. maximal proposition. A maximal proposition is simply a rule through which a consequence holds.
> *Omnis consequentia bona tenet per aliquem locum qui est propositio maxima. Nam propositio maxima non est nisi regula, per quam consequentia tenet.*

As Burley says, rules (in the relevant sense) are propositions, so they are statements that are either true ((422) *regulae verae*) or false ((301) *regulam falsam*). (But see (309) *regulam bonam*.) Burley gives many examples of such rules; they are general statements saying that all consequences with certain features are good.

Burley says in many places—I counted 54—that some consequence holds 'through' (*per*) a certain rule. I found no evidence that he means any more than that the rule is true and the consequence is an instance of it.

For example at (300) he says that 'it's clear' (*patet*) that a particular consequence is through a certain rule; since we are given no information of any kind about the context in which the consequence is being used, the rule must be something recognisable from the form of the consequence itself.

Another piece of evidence in the same direction is this. In a few places Burley gives a consequence schema, with letters in place of expressions, and says that it holds 'through a rule'. Thus at (433):

> The first mood is when both premises have affirmative antecedents. For example 'If A then B. If C then not B. Therefore if C then not A.' This is argued by the following rule: Whatever entails the opposite of the consequent entails the opposite of the antecedent.
> *Primus modus est, quando antecedens utriusque praemissae est propositio affirmativa. Verbi gratia: 'Si* A *est,* B *est; si* C *est,* B *non est; ergo si* C *est,* A *non est'. Et arguitur per hanc regulam: Quidquid antecedit ad oppositum consequentis, antecedit ad oppositum antecedentis* . . .

(At (441) he explains that 'A *est*' stands for a proposition and 'A *non est*' stands for its negation.) The 'mood' here should be read as a rule, saying that every instance of the schema is a good consequence. This is exactly the content of the 'rule' through which Burley says the schema holds. The 'mood' and the 'rule' say the same thing, and the difference between them is that the 'mood' is written as a schema whereas the 'rule' is written as a statement about consequences. (There are similar examples at (434), (443f), (464), (465), (467), (468), (469). In some of these cases Burley refers to the schemas not as moods but as syllogisms.)

In four places ((263), (293), (296), (321)) Burley talks of consequences being 'based on' (*fundatae supra*) some rule. Any impressions that he might mean that the rule is the reason why the consequence is good are dispelled by (263), where he says that the rule in question is false and all consequences based on it are fallacious.

So the passage in (341) probably says rather less than one might have guessed. It doesn't say that for every good consequence there is a rule which is the *reason why* the consequence holds. It doesn't say that our *knowledge* of good consequences is always derived from general principles of reason. There are good Aristotelian authorities for both these views, but they are not in PL. Perhaps this is part of what Burley meant by entitling his book 'The Purity of Logic'.

One possible way to read (341) is as a methodological statement. It's a task of logic to classify good consequences; given a good consequence, a logician should always seek to catalogue it with other good consequences. But a phrase in (289) suggests a more explicit reading of (341):

> But these rules are enough for making syllogisms in conditional hypotheticals by skill.
> *Sed istae regulae sufficiunt ad artificialiter syllogizandum in hypotheticis conditionalibus.*

To do a thing *artificialiter* is to do it by using learned skills like those of a craftsman (*artifex*). Logicians do seek to catalogue good consequences by the rules that they obey; but they have a specific practical reason for doing this. If you learn the rules, you can use them to construct good arguments to a professional standard. Without these rules you have to rely on such common sense as you have.

Burley spells this out more fully in his late Commentary on the Ars Vetus ([5] p. 2):

> And one should know that logic is useful as a faculty or power of distinguishing by skill between true and false in the separate branches of knowledge. For in these branches of knowledge one distinguishes true from false by reasoning from premises known to be true, to conclusions that follow from them, and logic teaches this kind of reasoning, so that it is through logic that in every branch of knowledge, true is distinguished from false by skill. Nor can one have any knowledge by skill except through logic. . . . One should know that there are two ways to have logic, namely by instinct and by skill. A person who uses a syllogism without knowing that he is doing so has instinctive logic, but he doesn't have logic as a skill since he doesn't know that he is using syllogisms and has logic. . . . But a person who uses a syllogism and knows that he is doing so has logic as a skill, since he knows the nature of syllogisms.
> *Et est sciendum quod utilitas logicae est facultas seu potestas discernendi artificialiter verum a falso in singulis scientiis. Nam in singulis scientiis distinguit verum a falso per discursum factum a praemissis notis ad conclusiones sequentes ex illis, et talem discursum docet logica, et ideo per logicam distinguitur artificaliter verum a falso in omni scientia. Nec potest aliqua scientia artificialiter haberi sine logica. . . . Sciendum quod logica potest haberi dupliciter, scilicet usualiter vel*

artificialiter. Utens enim syllogismo nesciens se sillogizare habet logicam usualem sed non habet artificialem quia nescit se syllogizare nec habere logicam. ... Sed utens syllogismo sciens se syllogizare habet logicam artificialem quia novit naturam syllogismorum.

6 Types of rule of good consequences

Burley refers to some consequences as 'syllogisms' (*syllogismi*); this word occurs 232 times in PL. He certainly doesn't mean just the familiar 'syllogisms' of Aristotle; he calls these 'categorical syllogisms' (*syllogismi categorici*). Other kinds of 'syllogism' are catalogued according to the kinds of proposition that occur in them. Thus we meet hypothetical syllogisms (249), syllogisms in causals (614), syllogisms of exclusives (931), syllogising in hypothetical conditionals (289), syllogising in disjunctives (545), syllogising in exceptives (931), syllogising in reduplicatives (935).

Burley describes the types of syllogism sometimes with consequence schemas ('moods') and sometimes with rules. Sometimes (as we saw above) he uses both. I think he is consistent in restricting the word 'syllogism' to actual consequences; a mood is not a syllogism but a way of forming syllogisms. Also syllogisms are good or not good, just as consequences are good or not good. This is slightly confused by his usage 'a syllogism is made' (*fit syllogismus*, e.g. at (496)), by which he always means that a *good* syllogism is made.

From the many examples that he gives, it seems that for Burley a syllogism is always a consequence with two premises and one conclusion. (At (283) he says explicitly that syllogisms, unlike some 'non-syllogistic consequences', have two premises.) In most of his examples the premises and conclusion of a syllogism all have a similar form, for example they might all be exclusives, though some of these exclusives might be negated. (A small exception is at (483ff), where he considers syllogisms that have one conditional premise and one categorical.) In most cases three terms are involved, and each proposition involves two of them. This allows Burley to catalogue most syllogisms into three 'figures' copying Aristotle's classification: in first figure the term appearing in both premises is in different positions in the two premises, in second figure it's in second position in both premises, and in third figure it's in first position in both. In hypothetical syllogisms there are three sub-propositions instead of three terms, and in reduplicatives the propositions each have three terms—though sometimes two are the same. (When one premise is conditional and the other is categorical, the classification into figures breaks down, so that at (486) he says 'This syllogism is not made in any figure'.)

What does Burley mean to convey when he describes a particular consequence or rule as a 'syllogism'? I don't think PL supports anything stronger than this: *A syllogism is a consequence that belongs in a class of consequences that can be handled systematically in a way analogous to Aristotle's treatment of categorical syllogisms.* But this leaves it open what he counts as analogous to Aristotle's treatment of categorical syllogisms.

At (767), commenting on propositions with certain features F, he says:

> But if [the premises have features F], no conclusion follows syllogistically and formally. This is clear from counterexamples in the terms.
> *Sed si in utraque praemissa . . . , nulla conclusio sequitur syllogistice formaliter. Quod patet per instantias in terminis.*

This passage is problematic and we will come back to it in Sections 16 and 17. But it seems clear that he is discussing syllogisms with premises of a certain form. The fact that he limits himself here to *syllogisms* allows him to assume that in each case there is only a small number of possible conclusions, so that he can run through the possibilities. For each possible conclusion, he shows that it doesn't follow 'formally', by giving a 'counterexample in terms'; i.e. he shows that for each possible form of syllogistic conclusion, there is a bad consequence where the premises and the conclusion have the specified forms. The role of syllogisms here is to allow Burley to restrict himself to arguments of a certain type; a modern analogue would be to ask what follows 'by first-order arguments' from premises of a certain form.

Burley says 'syllogistically and formally'. I've said what I think he means by 'syllogistically'; what about 'formally'? Burley discusses this notion in several places, not always connected with syllogisms. The idea is that a consequence C is good 'formally' if all consequences derived from C by replacing the terms of C in certain ways are good too; to say that a consequence is formally good is to say that it is one of a class of good consequences that are related to each other in some way by substitutions. The problem is to pin down what kind of substitutions.

The reader can consult (363f), (380), (387)–(389), (767) and (995) and make some guesses. (Unfortunately (380) contains some textual corruption. Brunellus is a donkey, and the manuscripts disagree about where the arguments are about braying and where they are about laughing. Boehner's ([3]

p. 84) preferred reading makes Burley say in consecutive sentences that the consequence

> Brunellus can laugh. Therefore some man can laugh.
> *Brunellus est risibilis, ergo homo est risibilis.*

is not good, and that it is good. Spade ([4] p. 171) removes this anomaly but still saddles Burley with the view that this consequence is good not because of any relationship between Brunellus and man, but because of the relationship between 'Brunellus', 'man' and 'can laugh'.)

For what it's worth, my impression is that Burley describes a good consequence as 'formally good' when it remains good under substitution for terms (the same substitution for each occurrence of the same term, of course), provided that all *conceptual* truths of the forms 'All A's are B's', 'No A's are B's', and their negations, are preserved by the substitution. For example 'All men are animals' is a conceptual truth, but 'All men can laugh' is just an 'accidental' truth. (See particularly (387).)

Burley allows a good consequence to depend on conceptual relationships of these kinds even when the relationships aren't spelled out as premises. I am guessing that when he says 'syllogism', part of what he means is that any such relationships, if they are needed, *are* stated explicitly as premises. So to say that a consequence holds 'syllogistically and formally' means that it holds under all systematic substitutions for terms, provided only that the substitutions take conceptually true premises to conceptually true premises. This is not the same notion as 'logically true' in the sense of Tarski [22], a notion that never appears in PL.

Because of their restricted forms, syllogisms have a better chance of being 'obviously good' than some other arguments. Burley speaks of some syllogisms as 'obvious (in itself)' (*(de se) evidens*) or 'perfect' (*perfectus*), which seems to mean something similar. See (170), (172), (213), (236), (415), (499).

However, at (212), (213) and (236) he speaks of a syllogism being 'perfect and regulated' (*perfectus et regulatus*), and 'regulated' here means that the syllogism is derivable from a more basic rule of argument. We know this because at (212) and (236) he tells us what that more basic rule is, namely 'dici de omni'. This is important, because it means that even the most self-evident syllogism rules need not be the most basic argument steps. One of my main aims in this paper is to describe what (for Burley) those most basic steps are. Burley makes no attempt to list them.

7 Shallow and deep rules

It will be helpful to make a distinction between shallow and deep rules. A *shallow* rule is one that depends on only a bounded amount of unpacking of the premises and conclusion (the bound depending on the rule, not the consequence). All other rules are *deep*.

For example Burley (409) gives the rule:

> If A is the case, B is the case. If B is the case, C is not the case.
> Therefore if A is the case, C is not the case.
> *Si* A *est*, B *est; si* B *est* C *non est; ergo si* A *est*, C *non est*.

To see that this rule applies to a particular consequence, we need to analyse the premises and conclusion down to the forms

> If (A is the case), (B is the case).
> If (B is the case), (C is not the case).
> If (A is the case), (C is not the case).

and then match up the bracketed parts. So the rule is shallow. (Some applications of it might hide their structure, so that they need to be paraphrased before the analysis. This is a separate issue.)

Most rules of propositional logic are shallow. One deep rule is the Replacement rule saying that if a proposition p is logically equivalent to q, then we can replace any occurrence of p by an occurrence of q inside any proposition, and the resulting proposition will be logically equivalent to the original proposition. Burley uses a (non-propositional) rule of this type at (667), though he doesn't spell out the rule he is using.

Here is a more exotic propositional example (I believe from Leśniewski):

> If the conclusion of a consequence uses no connectives except 'if and only if', and each atomic sentence in the conclusion occurs an even number of times, then the consequence is good.

Since the connective 'if and only if' can occur any number of times in the conclusion, the atomic sentences can lie at any depth in the analysis. So the rule is deep.

More familiar examples of deep rules are the quantifier rules of first order logic. For example:

> Let $\phi(x)$ be a formula with just x free, and c a constant. Then the sequent $\phi(c) \vdash \exists x\phi(x)$ is valid.

This rule is deep because the occurrences of c may lie arbitrarily deep in the syntactic structure of ϕ. But note that a shallow version of this rule copes with most ordinary language applications. Take dirty face sentences:

> Socrates believes that everybody else believes that he believes that everybody else believes that ... that his face is dirty. Therefore someone believes that everybody else believes that he believes that everybody else believes that ... that his face is dirty.

Here the anaphoric pronouns 'he' and 'his' take over the role of the free variables. We can quantify the sentence by putting 'someone' in place of 'Socrates', and for this we need only analyse the sentence as far down as its subject.

In PL Burley explicitly states about eighty rules of good consequences, usually labelled 'rule' or 'way of forming syllogisms'. All of these rules are shallow. In view of the previous paragraph, this is not very surprising. However, there are various pieces of indirect evidence for the use of at least potentially deep rules in Latin logic even before Burley.

My first witness is John of Salisbury, whose Metalogicon [20] dates from the mid twelfth century. Describing a certain logician whom he names by a pseudonym, John says ([20] p. 10f, 829ab):

> So he needed a calculus whenever he had to dispute, so as to be able to recognise affirmative force and negative force. For in many cases two negations have the force of an affirmation, and likewise an odd number of negations creates a negative force. ... So in order to tell whether he was dealing with an odd or even number, he found it a prudent policy to take to debates a handful of beans and peas that he would call on.
> *... ita ut calculo opus esset, quotiens fuerat disputandum; alioquin vis affirmationis et negationis erat incognita. Nam plerumque vim affirmationis habet geminata negatio; itemque vis negatoria ab impari numero convalescit; ... Ut ergo pari loco an impari versetur deprehendi queat, ad disceptationes collectam fabam et pisam deferre, que conveniebatur, consilio prudenti consueverat.*

(So perhaps bean = odd, pea = even. There is a mild pun on *calculus*, which means both algorithm and pebble. The beginning of Chaucer's Miller's Tale relates that the clerk took with him on his travels a set of algorithm stones, *augrim-stones*.)

The implication of this passage is that logicians of this period were familiar with deeply nested negations and had some calculus of positive and negative occurrences. But we have to be cautious, because the nesting need not have been within single propositions. In a debate one might say 'I claim that the following argument, assuming the truth of your last response, refutes the contradictory of my previous statement'. Here there are two nested negations in one statement, but still each statement in the debate might have no nesting deeper than two.

My second witness is the De Probationibus Terminorum [2] of Richard Billingham, composed possibly in the 1340s. Billingham writes ([2] p. 51):

> From inferior to its superior with subject held fixed, and with any expression with the force of negation put later than the inferior and superior, the consequence is good.
> *ab inferiori ad suum superius cum constantia subiecti et cum dictione habente vim negationis postposita inferior et superiori, tenet consequentia.*

In Section 11 we will see how one can read this as a sound rule, at least for suitably regimented languages. The rule is deep because nothing is said about how deep the inferior and superior may lie in the structure of the proposition. The caution this time is that although he doesn't say so, Billingham may have intended the rule to be used only to replace the entire predicate in subject-predicate sentences, and this is shallow. His mention of the subject suggests this.

Billingham's rule has antecedents as far back as the Abbreviatio Montana ([18] p. 86) from the late twelfth century. The Abbreviatio certainly has in mind replacement of subject or predicate in subject-predicate sentences, since it spells out the relevant sentence structures in detail.

8 Burley's quantifier rules

No doubt there is scope for someone to write a dissertation reducing Burley's stated rules to valid sequents in some appropriate formal calculus; I haven't pursued the idea. His propositional rules are standard, and they include a form of reductio ad absurdum (499). Most of these rules are explicitly stated.

Burley shows no awareness of any logical rules for sentences not built up by quantifiers or propositional connectives. At (371) he says that 'singulars' are 'equivalent to particulars' (*aequivalens particulari*) and claims that 'Nothing follows from particulars'. Counterexamples are easy to find. For example, taking as first premise one of the 'singulars' that Burley is discussing here,

> Socrates is running. Socrates is the husband of Xanthippe. Therefore the husband of Xanthippe is running.

Why does Burley think such consequences are bad? We aren't told. Perhaps he means only that no such consequence is a syllogism in any sense of 'syllogism' that he allows. (See Example 3 in Section 17 below.) But the net effect is that if we have to deal with a consequence where the premises explicitly or implicitly involve neither quantifiers nor sentence connectives, nothing in PL will help us.

The lack of such rules wouldn't necessarily hinder a medieval mathematician. For example the first of Euclid's Common Notions is the universal statement that things equal to the same thing are also equal to each other. Appeals to transitivity of identity could be seen as applications of this universally quantified law.

After rules for propositions, and the nonexistent rules for singulars, there come rules for quantifiers. Burley has names for some of the procedures involved; but the names describe a kind of move, not the conditions under which the move is valid. In this section I take some of the simpler cases. Universal quantification and monotonicity need closer treatment and will come in the sections that follow.

8.1 Dici de omni

This rule appears at (213), (236) and (983). It also appears at (210), (212) and again at (236) under the name *dici de omni vel de nullo*. The name *dici de omni* means 'saying about every'.

At (213) and (236) Burley applies this name to argument steps of the form

> From 'All A's are B's' and 'c is an A' infer 'c is a B'.

Burley correctly notes at (213) that if A is a class whose membership varies from one situation to another, then the inference only holds where the premise 'c is an A' holds in the same situation that is intended in the premise 'All A's are B's'. To illustrate he offers

> Everything white was black. Socrates was white. Therefore Socrates was black.
> *Omne album fuit nigrum, Sortes fuit albus, ergo Sortes fuit niger.*

The inference holds only if 'Everything white' means everything that was white, not everything that is white. (Of course Burley's details are dead wrong here. The relevant class of white things depends on the context in which the inference is used, and the possibilities are a great deal more complicated than just 'what was white' and 'what is white'. But the medieval Latins were serially blind to questions of context; we just have to live with that.)

At (981) Burley gives a bad consequence that appears to be of the *dici de omni* form. The consequence fails when we realise that the term in the 'c' position needs analysis, and after analysis the inference no longer has the *dici de omni* form. The term in question is 'Socrates by virtue of the fact that he is an animate substance' (*Sortes inquantum est substantia animata*).

At (983) Burley seems to apply the name *dici de omni* to a different form of inference:

> Every man has perceptions by virtue of the fact that he is an animal. Everything that can laugh is a man. Therefore everything that can laugh has perceptions by virtue of the fact that it is an animal.
> *Omnis homo est sensibilis inquantum animal, omne risibile est homo, ergo omne risibile est sensibile inquantum animal.*

There is not enough here to tell whether he intends a different form of *dici de omni*, or whether he means that the argument holds by *dici de omni* together with other appropriate quantifier rules.

8.2 Descent

Descent allows us to infer, from a proposition S containing a term T, any sentence S′ got from S by replacing T by the name of an object described by T in the context of the sentence S. If T is the subject term, then we have to drop any quantifier expression attached to the term when we replace it by the name.

We discussed descent briefly in Section 4. A term T that allows descent as above is said to have confused and distributive supposition in S. So any statement of the form "Under such-and-such conditions a term has confused and distributive supposition" is in fact a rule of good consequences. The main rule of this kind is that in universally quantified categorical propositions the subject term has confused and distributive supposition. This statement is equivalent to the rule of *dici de omni* that we have just discussed.

> From 'Every A is a B' and 'c is an A' there follows 'c is a B'.

Another case (87) is a categorical proposition where the subject term carries a negative universal quantifier such as 'none of' or 'neither of'; in such a proposition the predicate has confused and distributive supposition. Thus:

> From 'No A is a B' and 'c is a B' there follows 'No A is c'.

In Section 10 we will see how to derive this by a monotonicity argument.

8.3 Existential generalisation

Burley certainly knows the move made in inferences such as

> Socrates is running. Therefore a man is running.

But he seems to have no name for it. At (85) he describes it rather clumsily as 'The proposition is inferred from any one of the supposita of the term', i.e. the term 'man' in the case above. Note that since this is the subject term, English demands an explicit existential quantifier. (In Latin one can add *quidam*, but it's not required.) A term that allows this inference rule is said to have 'determinate' *determinata*) supposition if it also allows disjunctive descent, and 'simply confused' (*confusa tantum*) supposition if it doesn't.

Just as with descent, any rule saying that certain terms have one of these kinds of supposition yields rules of good consequence. For example when Burley tells us at (164) that in 'Each of them said something true', 'something true' has simply confused supposition, he is (among other things) licensing the inference

> Each of them said the two times table. Therefore each of them said something true.

This example is quite interesting because the term 'something true' is a proper part of the predicate. So any justification of Burley's claim would need something tending towards a deep rule. But he gives none. Again we will see (in Sections 10f) how he could have done.

9 Universal generalisation

From the reference list, 25 of Burley's good consequences have universally quantified conclusions. In all of these the conclusion begins with a sign of quantification ('every' *omnis*, 'no' *nullus*, 'nothing' *nihil*, 'only' *tantum*). Sometimes Burley takes 'Every A is a B' as implying that 'There is an A', which he proves separately.

In some of these twenty-five consequences ((127), (543), (645), (667), (764), (921), (985)) Burley states the conclusion without any indication of how it's derived, except perhaps the general rule which it illustrates. In (750) his argument is unclear to me.

The rest of the twenty-five fall into two classes. First there are those inferences where the conclusion is derived by propositional rules from some other proposition or propositions that already contain either it or its exponents or a paraphrase of it ((642), (816), (818), (833), (895)). This includes those cases where he derives the negation of the premise from the negation of the consequent, as at (280), (486), (638) with a variation at (996). It also includes the cases where he derives the conclusion by transitivity of entailment, as at (305) and (990); (988) is a slight variation of this.

Second there are those cases where a premise already contains a corresponding quantifier, and Burley carries out some manipulation of the proposition while holding the quantifier fixed. In (667) he replaces the phrase 'man' inside a universally quantified sentence by the phrase 'distinct from any non-man'. At (85) he uses existential generalisation inside the scope of a universal quantifier. The remaining cases (648), (648) and (996) are basically monotonicity arguments. We will come to monotonicity arguments in the next section.

In some of these cases, Burley proves the consequence by invoking a categorical syllogism and then deriving the conclusion by propositional rules or paraphrase. One could ask how he proves the categorical syllogisms, but PL is largely silent on this.

Now all the proofs in these two classes have an interesting feature in common. The universal quantifier never appears or disappears; we can trace it from the conclusion, through intermediate steps, back into the premises. Frege in his Begriffsschrift of 1893 officially kept to this style too: he had no rules for adding universal quantifiers. But in practice Frege found the restriction intolerable, and he introduced a convention that allows one to drop universal quantifiers in favour of latin letters ([7] §17). Then we can compute using the latin letters, and restore the quantifiers later. Today we find it more natural to regard his rules for adding and removing latin letters as quantifier rules of his system.

One can distinguish at least three ways of reaching a universally quantified conclusion without maintaining the quantifier from premises to conclusion:

Full Enumeration: To prove that all A's are B's, list all the A's and check that each of them is a B.

Sample: To prove that all A's are B's, look at a suitable sample of A's and check that all A's in the sample are B's. (This method can lead to false results, for example if through lack of imagination you miss an important sort of A.)

Universal Generalisation: The method is to introduce a letter, say a, and proceed to make deductions using the formal proposition that a is an A, but no other assumptions using the letter a. If you succeed in deducing that a is a B, then it follows that all a's are B's. (This is the standard mathematicians' method, used by geometers since at least the fourth century BC.)

There is no evidence in PL that Burley understands any of Full Enumeration, Sample or Universal Generalisation well enough to use them reliably. There is some evidence that he doesn't.

For example at (672) he gives what I think must be intended as an example of Full Enumeration:

> And when it is proved: Everything distinct from any non-man is an animal, therefore everything distinct from any donkey is an animal; for the singulars of the antecedent imply the singulars of the consequent. For it follows: This thing differing from any non-man is an animal, therefore this thing distinct from any donkey is an animal, and likewise with the rest.
> *Et cum probatur: Omne differens a non-homine est animal, ergo omne differens ab asino est animal; nam singulares antecedentis inferunt singulares consequentis; sequitur enim: Hoc differens a non-homine est animal, ergo hoc differens ab asino est animal, et sic de aliis.*

The 'and likewise with the rest' makes it clear that Burley is talking about an enumeration, not a formal argument as in Universal Generalisation. But for this he should be enumerating the items in the range of the universal quantifier of the consequent, namely the things distinct from any donkey;

he has gone to the wrong quantifier. A second problem is that it's not plausible to list all things distinct from any non-man (or from any donkey). The consequence is in fact bad, but Burley never gets round to explaining why. Is that perhaps because he didn't have the matter clear in his own mind? If he really had a clear understanding of the moves involved, one suspects he would have chosen a less confusing example to illustrate them.

There is a rule for making deductions from existentially quantified propositions 'Some A is a B':

Existential Instantiation: The method is to introduce a letter, say a, and proceed to make deductions using the formal proposition that a is an A and a B, but no other assumptions using the letter a. If you succeed in deducing a conclusion not mentioning a, then the conclusion follows already from 'Some A is a B'.

This rule is a kind of dual to Universal Generalisation.

Burley shows at (150) that he has come across some version of Existential Instantiation. But he thinks that the method works by taking one of the A's and naming it a, an operation that he calls 'signing' (*signat*). Thus he considers an argument which starts 'For every magnitude there is a smaller magnitude; let A be this smaller magnitude'. He rightly objects to this argument: he says that it is not legitimate to replace the term 'a smaller magnitude' by a sign naming a particular magnitude (*non licet ponere aliquod suppositum eius*). But a is not a sign naming a particular magnitude. If the problem were what Burley says it is, then we would be inhibited from making deductions from 'There is a grain of sand on the Siberian coast' until we'd identified and named a grain of sand on the Siberian coast.

Burley's misunderstandings of Universal Generalisation and Existential Instantiation come together at (761f). Here he gives a proof of a universal 'Every B is an A' from an exclusive. He turns it around so as to derive the negation of the exclusive from the sentence 'Some B is not an A'. He correctly reasons: Let c be a B that is not an A, etc. We would suppose that he had correctly understood Existential Instantiation if he hadn't gone on to add:

> And the same goes for any other singular.
> *Et eodem modo est de qualibet alia singulari.*

There are two mistakes here. First, he thinks that a singular has been mentioned; it hasn't. Second, he seems to think that part of the argument is to generalise from one singular to all singulars. To me this suggests he has confused Existential Instantiation with Universal Generalisation, and then Universal Generalisation in turn with Full Enumeration.

John Peckham's *Perspectiva Communis*, written a few decades before Burley's book, proves a number of universally quantified geometrical statements. His arguments in general are certainly thin, fitting his intention to write a popular book. Also some of his physical or physiological views are indefensible, and there are a few slips in the geometry. But his geometrical demonstrations of universally quantified statements are normally squeaky clean examples of Universal Generalisation. I found nothing corresponding to the logical misunderstandings in Burley. True, Peckham introduces many of these proofs with 'For example' (*verbi gratia*); but this seems to be a turn of phrase, meaning 'Here's a way of seeing why this is true'. Did Peckham understand universal quantification better than Burley, or was it a feature of the age to fail to transfer to logic what they understood in mathematics?

I haven't mentioned Sample yet. We will see later that Burley's lack of fluency with basic techniques comes home to roost when he has to prove statements about all consequences with a certain form. When he can't deduce these statements from already known general laws, he is reduced to giving examples. Thus when he says that consequences with a certain form are bad, and gives a single example, we can't tell from the form of his argument whether he is meaning to show (a) that not all consequences with that form are good, by giving a counterexample (*instantia*) or (b) that all consequences with that form are bad, by giving an example that he wants us to recognise as typical. We will see that there are some fairly severe problems of interpretation, but (b) seems to be part of the story. And (b) is a case of Sample.

10 Ab inferiori ad superius

Burley refers many times to a move called 'from lower to higher' (*ab inferiori ad superius*). Though he talks of it as a way of arguing, perhaps the best way to think of it is as a class of consequences. The modern name is 'upwards monotonicity', and a modern description might go as follows (making allowance for some looseness about the grammar).

Write $S(T)$ for a sentence in which we mark an occurrence of a noun phrase T. Let T' be another noun phrase, and write $S(T'/T)$ for the sentence got by replacing T by T' at the marked occurrence.

We say that the marked occurrence of T in $S(T)$ is *upwards monotone* if for all noun phrases T' the consequence

Every T is a T'. $S(T)$. Therefore $S(T'/T)$.

is good. Burley describes a consequence of this form, whether or not it's good, as 'from lower to higher'.

We say that the marked occurrence of T in $S(T)$ is *downwards monotone* if for all noun phrases T' the consequence

Every T is a T'. $S(T'/T)$. Therefore $S(T)$.

is good. Burley describes a consequence of this form, whether or not it's good, as 'from higher to lower'. (Recall that for Burley 'Every T is a T'' implies there is a T, so downwards monotonicity doesn't go to empty terms. This will cause some technical nuisances in Section 17.)

A number of Burley's remarks on 'from lower to higher' could be paraphrased by saying that occurrences of noun phrases in certain situations are (or are not) upwards monotone. For example at (102) he shows that in 'No animal except one of these is a man', 'man' is not upwards monotone. At (304) he shows that in 'If a man is running, a thing that can laugh is running' the word 'man' is not upwards monotone. At (305) he remarks that in 'If a man is running, an animal is running', 'animal' is upwards monotone. At (648f) we read that 'man' is upwards monotone in 'Only a man is running', but not in 'Only something that can laugh is a man'. An example at (927) has two occurrences of the relevant term: in 'Whatever is true is true at this instant', 'true' is upwards monotone at its first occurrence. At (981) there is an example involving 'by virtue of'.

Burley knows that when we build up compound sentences, terms that are upwards monotone in the component sentences may stay upwards monotone in the compound, and that we can sometimes show this by climbing up through the construction. For example at (568) he correctly notes that in a disjunction 'P or Q', an upwards monotone occurrence in one of the disjuncts remains upwards monotone in the disjunction (though he doesn't prove this in detail). In the same paragraph he shows he knows that negating a sentence reverses the monotonicities of occurrences of terms in it.

An argument at (302) is revealing:

> From the same rule, viz., that whatever follows from the consequent [follows from the antecedent], it's clear that in a conditional whose antecedent is a particular or indefinite proposition,

> the subject of the antecedent has confused and distributive supposition with respect to the consequent, so that ... there follows a conditional in which the subject of the antecedent is inferior to the subject of the first conditional. For example it follows: If an animal is running then a substance is running, therefore if a man is running, a substance is running.
> *Ex eadem regula, scilicet quidquid sequitur ad consequens in etc., patet quod in conditionali, cuius antecedens est propositio particularis vel indefinita, subiectum antecedentis supponit confuse et distributive respectu consequentis, ita quod ... ad talem conditionalem, cuius antecedens est propositio particularis vel indefinita, sequitur conditionalis, in cuius antecedente subiicitur aliquod inferius ad subiectum primae conditionalis. Verbi gratia, sequitur enim: Si animal currit, substantia currit, ergo si homo currit, substantia currit.*

Here Burley is showing that 'animal' is downwards monotone in 'If an animal is running then a substance is running'. The argument he uses is the transitivity of 'following'. We can reconstruct as follows. First, 'man' in 'A man is running' is obviously upwards monotone. It follows that the consequence

> If a man is running then an animal is running.

is good. But this together with 'If an animal is running then a substance is running' yields the stated conclusion. Here Burley moves the monotonicity one step deeper in a compound sentence by using an argument rule directly related to the principal of the compound sentence. It's clear that this could be iterated any number of times, and so we would discover monotone occurrences at arbitrary depth inside sentences.

Burley also pushes the matter forward in another direction: we can build up compound *terms*. The appropriate definition now is where $P(T)$ is a term and $P(T'/T)$ the result of substituting. We say that the marked occurrence of T in $P(T)$ is *upwards monotone* if for all noun phrases T' the consequence

> Every T is a T'. Therefore every $P(T)$ is a $P(T'/T)$.

is good. We say that the marked occurrence of T in $P(T)$ is *downwards monotone* if for all noun phrases T' the consequence

> Every T' is a T. Therefore every $P(T)$ is a $P(T'/T)$.

is good. Burley has no technical term to cover these, but he discusses the phenomenon. For example at (671) he notes that in ‘distinct from any donkey’ the word ‘donkey’ is downwards monotone. At (194) he notes that a certain argument would work if the word ‘going’ in ‘going to Rome’ was upwards monotone; to show that it isn’t, he cites ‘existing to Rome’. (This example raises important issues not connected with monotonicity, of course.)

Burley is also aware of constructions that block monotonicity. He knows that sentences about knowledge provide examples: ‘man’ is not upwards or downwards monotone in ‘You know whether a man is running’ (383), (385). At (1029) he knows that something subtler than ‘Every T is a T'’ may be needed in temporal contexts. But rather than analyse what really is needed, he opts for an idle solution and requires that ‘Every T is a T'’ is a *necessary* truth.

Burley tends to connect monotonicity and distribution. For example at (302) he explains the downwards monotonicity of ‘animal’ in ‘If an animal is running, a substance is running’ as a case of confused and distributive supposition. At (304) he describes a case of upwards monotonicity as ‘ascent’. At (168) he refers to a case of failure of upwards monotonicity as a matter of distribution. At (648f) he uses confused and distributive supposition as the reason why a certain term isn’t upwards monotone.

In fact downwards monotonicity and descent to singulars are quite different phenomena, and Burley loses information by confusing them. Descent to singulars replaces a noun phrase by a proper noun, and deletes any quantifier attached to the noun phrase. Downwards monotonicity replaces a noun phrase by another noun phrase, and leaves the quantifier in place. (At (382) and (384) Burley discusses a move that replaces a noun phrase by a noun phrase *and* removes the quantifier. This is neither descent nor monotonicity; Burley makes it a matter of distribution.)

Downwards monotonicity implies descent to singulars, but not the other way round. To derive descent from downwards monotonicity, the simplest route is to use a noun phrase that is true of just the one individual. For example ‘man’ is downwards monotone in ‘Every man is running’. We deduce that Dov Gabbay is running by carrying out the consequence

> Every person who is Dov is a man. Every man is running. Therefore every person who is Dov is running.

Since Dov is the only person who is Dov, the conclusion says simply that Dov is running. The same device works with ‘There is’ (but clearly not with quantifiers implying the existence of more than one thing, like ‘For at least two’).

In a kind of dual operation, upwards monotonicity implies ascent from singulars. For example the position T in ‘Some T writes books’ is upwards monotone, and this licenses the consequence

> Every person who is Dov is a logician. Some person who is Dov writes books. Therefore some logician writes books.

In other words, Dov is a logician, Dov writes books, so some logician writes books.

I think the main cause of the conflation of descent with distribution is that confused and distributive supposition was taken—long before Burley—to be the characteristic effect of universal quantifiers. (In PL see (87), (97), (111), (377), (383), (384), (752).) But universal quantifiers have both these properties: they generate downwards monotonicity and they allow descent to singulars. Since the former property implies the latter and not vice versa, it would have been more sensible if the medievals had taken the former as the characteristic property throughout, instead of shifting tacitly between them as Burley does.

11 A calculus of monotonicity

Burley’s treatment of monotonicity almost amounts to a calculus. In this section let me set it out more systematically. Since we are talking about natural language sentences, the word DEFAULT should be up in neon lights throughout—or maybe John of Salisbury’s judicious *plerumque*.

First consider simple affirmative categoricals. In ‘Every A is a B’, A is downwards monotone and B is upwards monotone. In ‘Some A is a B’, both terms are upwards monotone.

More complicated subject-predicate sentences are probably best treated as generalised quantifiers on two or more terms. For example in ‘Ignoring A’s, all B’s are C’s’, A and C are upwards monotone while B is downwards. In ‘Only A’s are B’s’ (read as Burley reads *tantum*), A is upwards monotone and B downwards. In ‘At least five A’s are B’s’, A and B are both upwards monotone.

Negating a sentence reverses all monotonicities in it.

Conjunction with ‘and’ and disjunction with ‘or’ preserve all monotonicities. This holds even where there are anaphoric pronouns, but of course we should avoid stupidities like trying to assign upwards monotonicity to a pronoun.

In 'If P then Q', monotonicities are reversed in P and preserved in Q. We should add that if conditionals are read intentionally, as Burley usually reads them, then the licensing 'Every T is a T'' should be true necessarily.

One can add further clauses in a similar spirit, for example to cope with monotonicity of noun phrases inside other noun phrases. The language covered by these constructions, with 'If ... then' interpreted so as to allow unrestricted monotonicities, can reasonably be called Monotone Latin.

Monotone Latin excludes constructions that block monotonicity. For example the constructions 'X knows whether P' or 'X is pleased that P' block monotonicities in P. The generalised quantifier 'A's are B's by virtue of the fact that they are C's' is arguably upwards monotone at B and downwards at A, but C is neither upwards nor downwards monotone.

This calculus yields a high proportion of the good consequences discussed by Burley, and infinitely many more not discussed by him. The rule of deducing $S(T'/T)$ from $S(T)$ and 'Every T is a T'', where T is upwards monotone in $S(T)$, is of course a deep rule. (Jan van Eijck [6] implements in Haskell a calculus which proves all good categorical syllogisms by two rules: monotonicity and symmetry.)

I think Burley and his colleagues could reasonably claim credit for the calculus of monotonicity. Granted, today's style demands greater rigour. Victor Sánchez Valencia in his PhD thesis [19] builds a calculus of monotonicity based on the λ-calculus, with a fragment of natural language to illustrate. In one direction Burley goes further: he includes temporal and modal phenomena.

In fact I think it's fair to say that Burley understands the theory behind the calculus of monotonicity better than he understands how to apply it. He is not always reliable in recognising upwards and downwards monotonicity. One can extend the notion to occurrences of sentences (so that for example in 'If P then Q', the occurrence of P is downwards monotone and the occurrence of Q is upwards). Burley certainly has this idea; it was in the background in his discussion at (302). But for example at (452) he considers second figure conditional syllogisms with one premise an affirmed conditional and the other a negated one, and he says that in this case the proposition common to both premises is affirmed in both or denied in both. This is exactly what he should *not* be saying; the crucial fact that he misses here is that in these syllogisms one of the occurrences is upwards monotone and the other is downwards. In Section 17 below, we prove on general principle that if both had the same monotonicity, there would be no valid syllogisms of this type.

In [21] Spade describes what is essentially a fragment of Monotone Latin, and observes that in all categoricals in his language, every term allows either descent to singulars or ascent from singulars (his Theorem 2 on page 199). He sketches an argument in terms of quantifiers. In fact a stronger statement follows immediately from the construction of the calculus: Throughout Monotone Latin, every occurrence of a term is either upwards or downwards monotone.

12 Invalid consequences

Of the 306 consequences in the reference list, 178 are not good. This is 58% of the total. A rough count on a modern textbook (Kalish and Montague [10], in some ways a modern counterpart of Burley) found 49 valid English arguments and 29 invalid; here 37% of the total are invalid.

The relative importance of invalid arguments in medieval Latin logic is an acknowledged fact. De Rijk ([17] p. 22) says:

> ... the doctrine of fallacy forms the basis of terminist logic. For this logic developed as a result of the fact that ... the proposition was beginning to be subjected to a strictly linguistic analysis. The first impulse to this was given by the discovery of Aristotle's *Sophistici Elenchi* and especially by the circumstance that scholars made themselves familiar with this work.

Looking at De Rijk's evidence, I query only his phrase 'strictly linguistic'. One of the chief morals of the *Sophistici Elenchi* was that you can't distinguish valid from invalid arguments by strictly syntactic criteria; you have to look at the meanings. So far as there are valid argument forms, these forms are at least partly semantic. The terminists set out to describe the features of meaning that count towards validity. Burley inherited their work and many of their attitudes. For example he inherited their shallow grasp of syntax, and to this extent both he and they were anti-linguists. We have seen examples of this.

Sixteenth and seventeenth century thinkers were apt to complain that scholastic logicians had refined the art of proving invalidity to a point where they refused to accept refutation even by valid arguments. Thus Locke ([12] §189) in 1690:

> Is there any thing more inconsistent with Civil Conversation, and the End of all Debate, than not to take an Answer, though never so full and satisfactory, but still to go on with the Dispute as long as equivocal Sounds can furnish (a *medius terminus*) a Term to wrangle with on the one Side, or a Distinction on the other . . . ?

There is an uncomfortable amount of truth in the charge.

For example Paul of Venice ([15] p. 20f) in the early fifteenth century argues that the following consequence is not good:

> There is no chimera. Therefore it is the case that there is no chimera.
> *Nulla chimaera est. Igitur ita est quod nulla chimaera est.*

He has two arguments. The first is that 'an affirmative proposition without any modal term does not follow from a non-pregnant negative proposition'. The second is that it's imaginable (*imaginabile*) that nothing exists or is the case (*nihil nec aliqualiter esset*), in which case the premise is true and the conclusion is false.

The second argument is an absurd appeal to introspection; happily Burley is free of this kind of nonsense. The first argument is an attempt to batter the reader with technology; but it makes no sense, because the manifest goodness of the consequence invalidates any theory that claims it's bad. *Eppur si muove.* This is highly relevant to Burley, because he makes several attempts at general rules guaranteeing the invalidity of consequences. As we go, we must ask how Burley hopes to justify these rules.

13 Showing that a consequence is bad

Burley's 'first general and principal rule of consequences' (258) is that

> if in some possible circumstances it is possible that there is a time when the antecedents are true and the consequent is false, then the consequence is not good.
> . . . *si aliquo casu possibili posito possit antecedens aliquando esse verum sine consequente, tunc non fuit consequentia bona.*

There are two nested possibles and one temporal operator here; I don't know whether Burley intended this or he just wrote carelessly. In fact for all his applications of the rule a simpler form suffices:

> If in some possible circumstances the antecedents are true and the consequent is false, then the consequence is not good.

He never claims the converse.

At least, I think he never means to claim the converse. At (499) he wants to show that a consequence is good. His method is to deduce a contradiction from the premises and the negation of the conclusion, and then appeal to *per impossibile*. But at the start of this argument he says 'If we suppose that the conclusion doesn't follow, let the negation of the conclusion be assumed'. This suggests that he thinks that badness of a consequence allows us to assume the premises and the negation of the conclusion. But a closer look shows that the words 'If we suppose that the conclusion doesn't follow' (*si datur, quod conclusio non sequitur*) are not a part of the argument at all, and it would have been clearer if he had left them out.

For 47 of the bad consequences in the reference list, Burley claims that the premises are true and the conclusion false. For a further 4 in the list he claims that this is the case under a posit. There are a large number of other cases where he leaves it to us to see that a given consequence is bad, and for most of these cases the true-premise false-conclusion test works. The test is logically sound.

More puzzling are the other arguments that Burley seems to use in order to show that certain consequences are bad. In later sections we will look at the systematic methods that he uses. Here I note one case where he uses a special argument. Burley is mistaken, but in an interesting way.

Burley argues at (239):

> Some are predicates that determinately include nonexistence, for example to be dead, to be decomposed and so on. And when it's argued from a proposition in which such a predicate is predicated, to simple existence, this is a fallacy of relatively / absolutely. Thus it doesn't follow: Caesar is dead, therefore Caesar exists.
>
> *. . . quaedam sunt praedicata, quae determinate includunt nonesse, sicut esse mortuum, esse corruptum et sic de aliis. Et quando arguitur a propositione, in qua praedicatur tale praedicatum, ad esse simpliciter, est fallacia secundum quid et simpliciter. Et ideo non sequitur: Caesar est mortuus, ergo Caesar est.*

Here's a counterexample to Burley's claim:

> The current Pope is dead. Therefore the current Pope exists.

If there is such a person as the current Pope, then he must exist. So the consequence is perfectly sound, and not an example of any fallacy.

Burley has made the following mistake. It's correct that if P is a predicate of the kind that he describes, then from 'A is P' (under reasonable assumptions on the form of A) we can correctly deduce 'A doesn't exist'. But it in no way follows that we can't also deduce 'A exists'. Burley has confused 'We can infer not-q' with 'We can't infer q'.

This confusion is still very common in the subcultures of logic, for example among those many people who kindly send me their refutations of Cantor's diagonal argument [8]. In the most sophisticated example to reach me, my correspondent quoted an inference rule from Mostowski and restated it as a non-inference rule in the Burley fashion.

Burley's mistake would rule out arguments *per impossibile*. Not all logicians accept arguments *per impossibile*, but we have seen that Burley did.

14 Bad consequences and false rules

Why does Burley give examples of bad consequences?

For many of his examples the reason is clear. First, there are consequences that are given to illustrate some point about the meanings of words in the consequence.

For example at (56) he clarifies the use of the aristotelian notion of *perfectio* with the help of some inferences involving it; (69) does the same for the notion of 'in the first instance' (*primo*). A consequence in (194) is to explain the meaning of 'being' (*ens*), and (242)–(248) perform the same service for 'is' (*est*), and likewise (690) and (694) for the notion 'one' (*unum*). At (118) and (184) he illustrates the behaviour of anaphoric pronouns in inferences, and at (126) he illustrates reflexive pronouns in the same way. An inference in (630) makes a point about truth conditions of sentences about the past, and one at (377) illuminates truth conditions for statements involving 'knows that'. Inferences at (921) and (924) illustrate truth conditions for sentences containing 'unless' (*nisi*), and (884) illustrates the use of numerical expressions. At (360) he explains the difference between 'following from A or (from) B' and 'following from A-or-B'.

Second, there are examples that illustrate or prove that some rule is false. Recall that a rule of good consequences is 'true' if and only if every consequence obeying the rule is good. So a false rule is one that is obeyed by at least one bad consequence. Burley's word for a bad consequence proving that a rule is false is *instantia*.

For example, if our account of formal consequence in Section 6 is correct, then to say that a consequence is not formally good is to say that a certain class of consequences contains at least one bad consequence. There are examples at (102), (160), (364) and (380). I suspect that (162) is another example, though Burley doesn't mention formal consequence here. He claims that the consequence 'It's impossible that a person who is standing is sitting; Socrates is standing; therefore it's impossible that Socrates is sitting' is bad because 'the premises don't have any terms in common, as is clear if the propositions are unpacked' (*praemissae non communicant in aliquo termino, ut patet, si istae propositiones resolvantur*). He presumably means that the first premise should be read as '(The proposition that a person who is standing is sitting) is impossible', so that its terms are 'the proposition that a person who is standing is sitting' and 'impossible'. This suggests that he has in mind a class of inferences which is closed under some kind of replacement of entire terms.

Sometimes Burley uses examples to show that there is no syllogism of a certain form. We saw that the notion of a syllogism doesn't have a sharp definition, so there is an element of vagueness about this kind of argument. Examples are at (420) and (497). Also (421) and (445) are in the middle of discussions of syllogisms and should probably be understood this way.

Sometimes Burley simply wants to point out that some rules that some people might think are good are in fact bad. Probably his examples call on his teaching experience.

For example at (277) he gives an example to show that while the goodness of 'Not Q, therefore not P' ensures the goodness of 'P, therefore Q', the implication fails if we replace 'Not Q' by a sentence that is merely inconsistent with Q. At (658) he points out that the equivalence of 'a is not a B' and 'It's not true that a is a B' (where a is a proper name) breaks down if we replace 'a' by 'Only a'. Other examples are at (300), (334), (659) and (878).

Some of the rules that Burley refutes are very silly. For example at (417) he seems to be attacking the rule 'If P then Q; if R then S; therefore if P then S' (though his example puts in place of the second S a sentence obviously implied by S). At (382) and (383) I was unable to see any remotely plausible rule that fits his description.

15 Fallacies

There are a number of false rules that Burley describes as 'fallacies' (*fallaciae*) and gives names to:

Figure of Speech (*figurae dictionis*) (93), (94), (95), (145), (147), (149), (153), (155), (157), (164), (1030).

Consequent (*consequentis*) (168), (241), (250), (263), (376), (420), (421), (514), (652), (656), (671), (927).

Relative and Absolute (*secundum quid et simpliciter*) (239), (241), (242), (245), (684), (902).

Varying the Common Part (*accidentis ex variatione medii*) (310), (608), (610), (613)

Ambiguity (*aequivocationis*) (229)

For each fallacy Burley indicates a feature of a consequence that would make it an instance of the fallacy. These features need not be definitions of the fallacies; the feature that he indicates for Figure of Speech is one that particularly interests him, but the name itself suggests a wider class of features.

Now there are three things that we might say about a consequence C in connection with a particular fallacy Fa and corresponding feature Fe:

(I) C has feature Fe.

(II) C commits fallacy Fa.

(III) C is a bad consequence.

Burley is consistent about the relationships between (I), (II) and (III). They entail each other as follows:

$$\text{(I)} \Rightarrow \text{(II)} \Rightarrow \text{(III)}.$$

We can deduce this from the way he describes fallacious consequences. For example he says

> The consequence C is bad and a fallacy, because (usually *quia*, sometimes *nam*) it has the feature.
> (93), (155), (164), (168), (229), (376), (671), (927).

> The consequence C is bad because it commits the fallacy.
> (239), (657).
>
> The consequence C commits the fallacy because (usually *quoniam*, *quia*) it has the feature.
> (153), (1957), (613). (94) is similar.
>
> Every consequence with the feature commits the fallacy.
> (95), (245), (263). (420), (421), (514) and (611) are similar.
>
> Every consequence with the feature commits the fallacy and is bad.
> (241), (310), (652).

A comparison with the list at the head of this section confirms that these comments are spread fairly evenly across the types of fallacy.

One implication of all this is that there are various features, each of which guarantees the badness of every consequence that has it. Such features are rare on the ground in modern logic, but Burley seems to find them all over the place. The obvious candidates are that the premises are true and the conclusion false, or (if we want a formally recognisable property) that the premises are propositional tautologies and the conclusion is an explicit contradiction. But Burley's features go way beyond these. What has he found that we are missing?

In the case of Figure of Speech I think I know the answer. The feature that constitutes this fallacy is that we infer

> For every A there is a B such that . . . . Therefore there is a B such that for every A, ...

For example at (93)

> Twice (i.e. on each of two occasions) you ate some bread. Therefore there is some bread that you ate twice.

Now certainly not every consequence with this feature is bad. For example

> Everybody here has heard of someone who is Alfred Hitchcock. Therefore there is someone who is Alfred Hitchcock, and whom everybody here has heard of.

From Burley's detailed discussions of this fallacy, it seems clear that he would have ruled this out as a counterexample, because in this case no word 'imports multiplicity'. He appears to say at (89) that 'every', 'each', 'three times' etc. import multiplicity, as if we could check it from the word alone. But we already know from Section 4 that his semantic analyses are not compositional, and it would have been open to him to say that in this particular example 'Everybody' fails to import multiplicity because of the fact that the description 'is Alfred Hitchcock' can only hold of one thing.

If this is right, then it's part of the definition of the feature constituting Figure of Speech that the relevant consequence is bad. Since the implication from (I) to (III) goes by way of (II), it would follow that it's part of the definition of Figure of Speech that the consequence is bad; so the implication from (II) to (III) is part of the definition of (II).

Unfortunately this explanation is implausible for the fallacy of Consequent. First, Burley claims at (263) that a certain two rules (*regulae*) 'always produce a fallacy of the Consequent'. This claim is obtuse if the only reason that these rules produce a fallacy is that good consequences don't count as instances of them. Moreover at (300f), discussing a bad consequent, he says that it is clear (*patet*) that the consequence is argued by one of the false rules of (263), and the implication is that we can see that the conference is fallacious by seeing that it has a certain syntactic form. (I think this knocks out one possible soft-option escape route: namely that Burley means only that if the only reason we have for accepting a consequence is that it's an instance of a particular rule, then we don't have any good reason for accepting it.)

The 'false' rule that Burley says produces the fallacy at (300) is

> If q follows from p and r follows from p, then r follows from q.

Burley himself at (254), just a few pages earlier, has said

> From an impossibility anything follows.
> *Ex impossibili sequitur quodlibet.*

Thus if p and q are impossible, then q follows from p and r follows from p, so by the false rule, r follows from q; but in this case r really does follow from q, contrary to Burley's statement that the false rule always produces bad consequences. Another way to get good consequences out of the false rule is to start from any good consequence 'q therefore r' and take p arbitrarily.

There is no evidence in PL that Burley is aware of the contradiction between (254) and (263); so it's an idle question how he would have resolved it. However, this is one place where we can call on some logical facts. In some sense of 'follow', (254) is certainly correct. So the 'false' rule doesn't in fact always create bad consequences, and Burley is mistaken to say that it does.

But note that the two counterexamples above have an interesting feature in common. Going by way of (254), we finish with a consequence 'r follows from q' which is good regardless of what r is. Starting with a good consequence 'q therefore r', we satisfy the false rule regardless of what p is. So in both cases the counterexample depends on a redundancy. Similar redundancies will be important for us in Section 17.

With Varying the Common Part the situation is different yet again. Burley gives as an example the following ((310) slightly simplified):

> If it's no time then it's not day. If it's not day and it's some time, then it's night. Therefore if it's no time then it's night.

The fault in the argument is that 'it's not day' in the first premise is matched against 'it's not day and it's some time' in the second, and they are different. His point is that there is no good rule of the form 'If P then Q. If R then S. Therefore if P then S.' Now manifestly not every example of this rule is bad. For example the consequence 'If it's no time then it's not day but it is some time. If it's not day, then it's night. Therefore if it's no time then it's night.' is good and has this form. Examples are so easy to find that Burley can't conceivably mean that this form has any tendency to create bad consequences. His examples, here and elsewhere ((172), (608)), fit the following pattern:

> If we take a true syllogistic rule, and in one of the premises we replace the middle term by a new term, then the resulting rule is no longer true.

If this is what Burley means, then he is quite correct, as one easily checks.

We have here a *class of rules* in which every rule is false. This is a different matter altogether from a *class of consequences* in which every consequence is bad. Burley has done a disservice by lumping together the two kinds of class under the common name of 'fallacy'. But classes of false rules are an important notion in PL, and we devote our final two sections to them.

16 Classes of false rules

At (789) Burley makes the following very revealing remarks:

> But if each premise is exclusive affirmative and the principal word is negated in each, no conclusion follows by rules of syllogisms. For if any conclusion followed by rules of syllogisms, a negative conclusion would follow since each premise is negative. But no negative does follow. For it doesn't follow: Only an intelligent being is not a non-animal, only a non-man is not a non-animal, therefore only a non-man is not intelligent. Neither does it follow that a non-man is not intelligent, since the premises are true and the conclusion false. But if any negative conclusion followed, one of these would follow.
> *Si vero utraque praemissa sit exclusiva affirmativa et verbum principale negetur in utraque, nulla conclusio sequitur per regulas syllogismorum. Quia si aliqua conclusio sequeretur per regulas syllogismorum, conclusio negativa sequeretur, cum utraque praemissa sit negativa; sed nulla negativa sequitur. Non enim sequitur: Tantum intelligibile non est non-animal, tantum non-homo non est non-animal, ergo tantum non-homo non est intelligibilis. Nec etiam sequitur, quod non-homo non est intelligibilis, quia praemissae sunt verae et conclusio falsa. Si tamen aliqua conclusio negativa sequeretur, altera istarum sequeretur.* (I follow Spade's text against Boehner's here; otherwise the example wouldn't fit Burley's description.)

Here Burley is talking about a certain class of second-figure rules of syllogistic type. He argues that every such rule is false, for the following reasons. (i) Since the premises are negative, the conclusion must be negative. (ii) No negative conclusion follows by rules of syllogisms.

How does Burley justify this argument? He presents no case at all for (i). It happens to be an established fact for categorical syllogisms, and for this case I don't have a neat proof either. But the syllogisms under discussion here are not categorical. As a general rule of argument (i) is a non-starter:

> Catullus never fails to delight me. Therefore Catullus delights me.
>
> 2 is nothing other than $1 + 1$. Therefore 2 is $1 + 1$.

And so on. It's curious that at (419), (769) and (820) Burley states a more general rule that 'Nothing follows from negatives'. If he had had the confidence to use this rule here it would have made (ii) unnecessary. I don't think we can absolve Burley of assuming that a rule which works for categorical syllogisms works for all other syllogisms. (The rule does have a folksy kind of plausibility: you can't get something for nothing.)

His proof of (ii) is that (iii) any negative conclusion would have to have one or other of two given forms; but (iv) for each of these forms he has counterexamples. Again he offers no argument for (iii). It certainly isn't true that the two forms exhaust the possible forms that he has been considering in this section of PL. He should have tried the two terms in either order, allowing at least the subject term to be negated, and this yields four forms.

Whether Burley is right about (iv) is a matter of interpretation. If not being a non-animal is the same as being an animal, then the existential exponents of the two premises say that there is an animal, and it follows that some intelligent being is a non-man, not a form that Burley bothers to consider. In this case Burley's statement is wrong. On the other hand if—as is probable—the 'not a non-animal' formulation is meant to cancel the existential exponents, then Burley's statement is correct and we will prove it in Example One of Section 17.

As a sample of Burley's reasoning style, (789) is comparatively mild. Elsewhere he explicitly claims to prove a general rule from a single instance. Thus for example at (769) he says

> But if the exclusion is negated in each premise, no conclusion follows, since nothing follows from negatives. And this is clear from a counterexample in the terms.
> *Si vero in utraque praemissa negetur exclusio, nulla conclusio sequitur, quia ex negativis nihil sequitur. Et patet per instantiam in terminis.*

He makes the same claim for other general rules at (814), (836) and (838). (At (838) he leaves it to the reader to find the counterexample.) At (767) and (799) he claims that 'counterexamples' make the truth of certain general statements clear.

We are in a situation we discussed in Section 9: Burley is aiming to prove a universally quantified statement. Unless he can find a way of converting some established universally quantified fact into the form required, he is stuck with considering examples, the method we called Sample. For exclusives he does have the method of translating into categoricals and then checking the rules for categorical syllogisms. But for instance 'Only A's are not B's' translates into 'All non-B's are A's', which is not a classical form of categorical. Yet he does allow exclusive sentences with negated predicate terms to occur in syllogisms, as at (787) 'Only an intelligent being is not a non-animal'.

A number of Burley's results, though correct, depend crucially on the existential exponents. For example at (797) he claims that there is no exclusive syllogism in second figure where one premise is negated exclusive and the other is categorical. Thus by implication he rejects the syllogism

> Not only A's are not B's. Every C is a B. Therefore not only C's are not A's.

This syllogism has counterexamples, but in all of them there are no non-B's. The fact that Burley's discussion never mentions this is witness to the fact that his methods are insensitive to existential assumptions.

Understandably he does make mistakes. At (767) he claims that from exclusive premises in first figure, where one premise is affirmative and the other is negative, no conclusion follows syllogistically. He misses the following good syllogism mood:

> Not only not-B's are A's. Only C's are B's. Therefore not only not-C's are A's.

Probably he was checking against categorical syllogisms, and the 'not-C's' in the subject of the conclusion threw him. At (827) he claims that if the premise containing the predicate term of the conclusion is a negated exclusive, then there is no good syllogism in third figure. A counterexample is the mood

> Not only a non-A is a B. Every A is an C. Therefore not only a non-C is a B.

I can't account for his oversight here, except that it's quite late in his discussion of exclusives and his sampling method does require an undue amount of concentration.

In fact the interesting thing is that Burley makes as few mistakes as he does. Clearly he has intuitions that are sounder than his methods. We can't profitably guess how those intuitions went; but we can at least report logical facts that yield most of the conclusions that he wanted, using tools closely related to his.

17 The Medieval Interpolation Theorem

In [9] I showed how the Lyndon interpolation theorem would have come to the rescue of medieval logicians if they had known it. That theorem is quite sophisticated, and it's wholly unrealistic to imagine any medieval thinking in those terms. So here let me rephrase the basic result in a way that bypasses most of the machinery of Lyndon's theorem. I came on it by first using Hintikka sets to prove Lyndon's theorem, and then stripping down to bare essentials. Obviously nobody in the middle ages knew the result, but I believe the account below is entirely in terms that they could have understood.

For simplicity I will ignore the difference between the medieval notion of formal inference and the modern notion of an inference rule valid under all substitutions for its function or relation symbols (two occurrences of the same symbol being replaced by the same symbol in both places). Today we know that for a wide range of argument forms, if there is a counterexample at all then there is one in the natural numbers, where the relationships all hold for purely conceptual reasons. (This is due to Kurt Gödel in 1930 for first-order logic.) So I will speak of a set of sentences as *formally consistent* when there is some replacement of terms which turns it into a set of sentences that can be simultaneously true.

We consider two sets of sentences, Φ and Ψ. We say that the pair Φ, Ψ is *formally consistent* if the set consisting of all the sentences in either Φ or Ψ is formally consistent. If T is a term, we say that the pair Φ, Ψ is *formally consistent under variation of terms* if some pair $\Phi(T'/T)$, Ψ is formally consistent, where $\Phi(T'/T)$ is the result of replacing all occurrences of T in sentences of Φ by occurrences of a new term T' not present in sentences of either Φ or Ψ. (Since we are talking of formal consistency, it would be equivalent to say 'every pair' instead of 'some pair'.)

Theorem 1 (Medieval Interpolation Theorem) *Suppose Φ and Ψ are sets of sentences of Monotone Latin, and T is a noun phrase which occurs in sentences of Φ and Ψ only with upward monotonicity, or only with downward monotonicity. Then if the pair Φ, Ψ is not formally consistent, it is not formally consistent under variation of T either.*

Sketch proof The proof goes by recursion on the construction of sentences of Monotone Latin. Since the syntax of Monotone Latin is not completely determined, the proof has to be a bit vague. With a formal language there would be no difficulty in tightening it up to a rigorous argument.

First suppose that Φ and Ψ consist of simple categoricals with singular subjects, for example 'Socrates is running' or 'Brussels is not a village'. Here the terms are the predicates; they are upwards monotone in affirmative categoricals and downwards in negated ones. So if T is everywhere upwards monotone, all categoricals using it are affirmative, and clearly these can't lead to a formal contradiction. Hence in this case, if Φ, Ψ is formally inconsistent, this must be entirely because of categoricals in which T doesn't occur. So the formal inconsistency will still be there if we vary the term T in Φ. Essentially the same argument applies when T is everywhere downwards monotone.

The remaining cases consider more complex sentences in terms of their exponents or possible causes of truth.

For example suppose Φ contains a sentence 'P or Q', which has two possible causes of truth, namely P and Q. If Φ, Ψ is formally consistent under variation of T, then there is some possible situation in which all the sentences of $\Phi(T'/T)$ and Ψ are true. Such a situation must make at least one of the sentences P, Q true, say P; and we note that every term that occurs in P also appears in 'P or Q' *with the same monotonicity*. So if Φ' is Φ with 'P or Q' replaced by P, then Φ', Ψ is still formally consistent under variation of T. But the sentences in Φ', Ψ are simpler than those in Φ, Ψ, so we can assume that the theorem has been proved for Φ', Ψ, and hence Φ', Ψ is formally consistent. Now P was a possible cause of truth of 'P or Q', and *this remains the case when T is replaced by T' throughout* P *and* Q. Hence Φ, Ψ is formally consistent too.

Other cases are the same in principle. The hardest are those involving quantifiers, and here I grant that some familiarity with modern methods would help. Suppose Φ contains a sentence S beginning with a universal quantifier: 'No A's are B's'. If Φ, Ψ is formally consistent under variation of T, then there is some possible situation in which all the sentences of $\Phi(T'/T)$ and Ψ are true. This situation makes true all the sentences got from S by descent to singulars of A in the form 'Either c is not an A or c is not a B', and then replacing T by T' in them. The term T' (or T) has the same monotonicities in these sentences as it did in the sentence 'No A's are B's'. Replacing that sentence in Φ by the new sentences (before T has been replaced by T' in them) gives a new pair Φ', Ψ which is still formally consistent under variation of T. But the sentences in this pair are less complex than those in Φ, Ψ, so we can assume that the theorem holds for Φ', Ψ, and it tells us that Φ', Ψ is formally consistent. It follows that Φ, Ψ was formally consistent too. □

To apply the result in Burley's context, we need to draw out the existential assumptions explicitly. I don't know exactly what they are, but for example it seems we have the following monotonicities:

- Some A$\uparrow$ is a B$\uparrow$.
- No A$\downarrow$ is a B$\downarrow$. (Negation swaps the monotonicities.)
- Every A is a B $\equiv$ No A$\downarrow$ is a non-B$\uparrow$, and there is an A$\uparrow$.
- Only A's are B's $\equiv$ No non-A$\uparrow$ is a B$\downarrow$, and there is a B$\uparrow$.
- Not only A's are B's $\equiv$ Some non-A$\downarrow$ is a B$\uparrow$, or there is no B$\downarrow$.

Here are a few applications of the theorem. In each case we show that if a certain sort of consequence was formally valid, then it would remain formally valid if one of the terms was replaced, at one occurrence and not the other, by a new term. It's normally clear at once that none of the resulting consequences could be formally valid.

Example One. We saw in Section 16 that at (789) Burley claims that no syllogistic conclusion follows from a pair of premises of the form

Only an A is not a C. Only a B is not a C.

We also saw that Burley's claim is false if we read the premises as implying that something is not a C. So suppose we drop that implication. There remains:

No non-A$\uparrow$ is not a C$\uparrow$. No non-B$\uparrow$ is not a C$\uparrow$.

Suppose a conclusion P follows syllogistically. Let Φ consist of the first premise, and let Ψ consist of the second premise together with the contradictory negation of P. Since we are talking of syllogisms, the term C occurs nowhere in P, and hence it occurs with only upward monotonicity in both Φ and Ψ. So by the Medieval Monotonicity Theorem, if P follows formally from the two premises, then it already follows formally from the two premises

No non-A is not a D. No non-B is not a C.

But clearly the two premises don't interact, and nothing follows except what follows separately from each premise.

Example Two. This easy example uses monotonicity of sentences. At (423) Burley maintains that there is no syllogism with premises

If P$\downarrow$ then R$\uparrow$. If Q$\downarrow$ then R$\uparrow$.

or the same with ‘not-R’ in place of R. He says this is obvious, but gives a reason; the reason is obscure to me, but it seems to be an appeal to an analogous fact about categorical syllogisms. However, R is upwards monotone in both occurrences, so that by the Medieval Interpolation Theorem any syllogistic consequence would follow also from ‘If P then S’, ‘If Q then R’, which is absurd.

Example Three. At (372) and elsewhere, Burley quotes an old rule that nothing follows from two particular (i.e. existentially quantified) premises. In categorical syllogisms there are two possible forms of particular proposition, ‘Some A$\uparrow$ is a B$\uparrow$’ and ‘Some A$\uparrow$ is not a B$\downarrow$’. In both of these, A is upwards monotone. The Medieval Interpolation Theorem allows us to deduce quickly that there is no valid syllogism of this type in third figure, where both premises have the same subject. If we are not in third figure, then at least one term appearing in the conclusion must be the subject of a premise, and so must be upwards monotone in the premise. In second figure both terms of the conclusion must therefore be upwards monotone in premises, and so the Medieval Interpolation Theorem yields that they must both be upwards monotone in the conclusion. There is only one such consequence, namely

> Some A is a C. Some B is a C. Therefore some A is a B.

We find counterexamples at once. There remains the first figure; here there are two possibilities not ruled out by the Medieval Interpolation Theorem, namely

> Some A is not a B. Some B is a C. Therefore some A is a C.
> Some A is not a B. Some B is not a C. Therefore some A is not a C.

Both have obvious counterexamples. So the rule is correct for categorical syllogisms.

But note that in both the two first-figure examples there is a nontrivial conclusion about A and C that we could have derived from the same premises. For the first it’s ‘Some A is distinct from some C’. One could say that the reason why the Medieval Interpolation Theorem failed to rule out these two cases is that they are not intrinsically invalid, it’s just that the syllogistic calculus is too limited in the types of proposition that it allows. There is an obvious moral, namely that one should *never* take for granted that a rule which works for categorical syllogisms works for any other class of consequences. (Burley is particularly open to the charge of playing fast and loose here, since at (667) he is perfectly willing to allow consequences containing ‘Some A is distinct from all C’.)

Example Four. We saw that in (767) Burley claimed wrongly that nothing follows syllogistically from

Not only not-B’s$\uparrow$ are A’s. Only C’s are B’s.

Even ignoring the existential implications, B occurs with different monotonicities in the two premises, and this makes it likely that something will follow validly. To see whether any proposition P relating A and C does follow, it’s reasonable to start with consequences that don’t depend on the existential implications. Ignoring them, A has upwards monotonicity in the first premise and C has upwards in the second. So the Medieval Interpolation Theorem advises us to look for a conclusion where A and C both have upwards monotonicity. This greatly simplifies the set of examples that we need to search through. To find the counterexample that Burley missed took about a minute by this route.

References

[1] Albert C. Baugh and Thomas Cable, *A History of the English Language*, Routledge and Kegan Paul, London 1978.

[2] Richard Billingham, *De Probationibus Terminorum (first recension)*, in L. M. De Rijk, *Some 14th Century Tracts on the Probationes Terminorum*, Ingenium, Nijmegen 1982, pp. 45–76.

[3] Walter Burleigh, *De Puritate Artis Logicae Tractatus Longior* with a revised edition of the Tractatus Brevior, ed. Philotheus Boehner, The Franciscan Institute, St. Bonaventure NY 1955.

[4] Walter Burley, *On the Purity of the Art of Logic: The Shorter and the Longer Treatises*, translated by Paul Vincent Spade, Sheridan Books, Chelsea Michigan 2000.

[5] G. Burlaeus, *Super Artem Veterem* (reprint of edition printed in Venice 1497), Minerva, Frankfurt 1967.

[6] Jan van Eijk, ‘Syllogistics = Monotonicity + Symmetry + Existential Import’, preprint May 2005.

[7] Gottlob Frege, *Grundgesetze der Arithmetik Begriffsschriftliche Abgeleitet, I*, Pohle, Jena 1893.

[8] Wilfrid Hodges, 'An editor recalls some hopeless papers', *Bulletin of Symbolic Logic* etc.

[9] Wilfrid Hodges, 'The laws of distribution for syllogisms', *Notre Dame Journal of Formal Logic* 39 (1998) 221–230.

[10] Donald Kalish and Richard Montague, *Logic: Techniques of Formal Reasoning*, Harcourt, Brace and World Inc., New York 1964.

[11] David C. Lindberg, *John Pecham and the Science of Optics*, University of Wisconsin Press, Madison Wisconsin 1970.

[12] John Locke, *Some Thoughts concerning Education*, Clay, London 1884.

[13] Reviel Netz, *The Shaping of Deduction in Greek Mathematics*, Cambridge University Press, Cambridge 1999.

[14] Gabriel Nuchelmans, 'The distinction of *actus exercitus / actus significatus* in medieval semantics', in Gabriel Nuchelmans, *Studies on the History of Logic and Semantics, 12th–17th Centuries*, Variorum Reprints, Aldershot, Hampshire 1996, ch. vi.

[15] Paul of Venice, *Logica Magna ii fasc. 6, Tractatus de Veritate et Falsitate Propositionis et Tractatus de Significato Propositionis*, ed. Francesco del Punta, British Academy, Oxford 1978.

[16] Titus Maccius Plautus, *Comoediae*, ed. W. M. Lindsay, vol. ii, Clarendon Press, Oxford 1910?.

[17] L. M. De Rijk, *Logica Modernorum i: On the Twelfth Century Theories of Fallacy*, Van Gorcum, Assen 1962.

[18] L. M. De Rijk, *Logica Modernorum ii Part Two: The Origin and Early Development of the Theory of Supposition*, Van Gorcum, Assen 1967.

[19] Víctor Manuel Sánchez Valencia, *Studies on Natural Logic and Categorial Grammar*, PhD dissertation, University of Amsterdam 1991.

[20] John of Salisbury, *Metalogicon*, ed. Clemens C. I. Webb, Clarendon Press, Oxford 1929.

[21] Paul Vincent Spade, 'The logic of the categorical: The medieval theory of descent and ascent', in *Meaning and Inference in Medieval Philosophy*, ed. Normann Kretzmann, Kluwer, Dordrecht 1988, pp. 187–224.

[22] Alfred Tarski, 'O pojęciu wynikania logicznego', *Przegląd Filozoficzny* 39 (1936) 58–68; English translation 'On the concept of logical consequence', pp. 409–420 of Alfred Tarski, *Logic, Semantics, Metamathematics: papers from 1923 to 1938*, ed. John Corcoran, Hackett Publishing Company, Indianapolis, Indiana 1983.

[23] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, Routledge and Kegan Paul, London 1961.