Can Negation be Defined in Terms of Incompatibility?

Nils Kurbis

1 Abstract

Every theory needs primitives. A primitive is a term that is not defined any further, but is used to define others. Thus primitives should be terms that can be expected to be understood by everyone.

Negation is a a very fundamental concept. Everyone understands it. No one has problems grasping it. It is a perfect choice for a primitive. Nonetheless, there have been attempts to define it in terms of allegedly more fundamental concepts.

The motivation behind such attempts is to provide a principled basis on which to settle the debate between rival logicians concerning the correct properties of negation. Most prominently, the debate between classicists and intuitionists is largely one about the laws governing negation. If negation is chosen as a primitive, no principled decision can be made.

I shall investigate how successful such attempts are. In particular, I shall argue that defining negation in terms of incompatibility fails, because the latter notion is conceptually rather more demanding notion than negation. Besides, the approach fails to decide between classicists and intuitionists. As a matter of fact different incompatibility-theorists come to different conclusions concerning which logic is the right one. Thus quite apart from the conceptual difficulties involved in choosing incompatibility as primitive, in the light of the debate between classicists and intuitionists, the approach does not fare any better than choosing negation as primitive.

2 On Primitives

Choice of primitives is an important issue in the philosophy of logic, and it is one that I think is not paid as much attention to as it deserves. What is a primitive? Every theory needs primitives. They are terms that are not defined any further, but in terms of which others are defined. Accordingly one should use terms as primitive that can be expected to be understood by everyone, of which we can expect everyone to have an intuitive, pre-theoretical understanding.

So why not *not*? If all this is the case, why would anyone ever want to define negation, rather than take it as a primitive? Everyone understands negation. No one has problems grasping it. It is a prime candidate and a perfect choice for a primitive.

Moreover, the most straightforward way that comes to mind to define negation, namely in terms of truth and falsity by '$\neg A$ is true iff A is false', doesn't actually evade use of negation. Something needs to be said about the relation between truth and falsity, and this makes use of negation, e.g. 'If A is true, then A is not false'. So it seems that we don't get around negation.

Despite the fact that negation seems to be a most obvious choice for a primitive, philosophers have suggested that negation should be defined in terms of other concepts. Dummett suggests to define negation in terms of rules of inference. It has become fashionable to propose definitions of negation in terms of incompatibility.[1] Price and Rumfitt suggest to define negation in terms of assertion and denial, and they, too, make use of a primitive notion of incompatibility between speech acts.

3 Disputes over Logical Laws

The motivation behind attempts at defining negation in terms of some allegedly more fundamental notion is the aim to settle disputes between certain rival schools of logicians. Even though negation is such a simple notion, as a matter of fact intuitions have diverged concerning which logical laws hold

[1] Cf., e.g., Christopher Peacocke ('Understanding the Logical Constants. A Realist's Account', *Proceedings of the British Academy* 73 (1987), 153-200), Huw Price ('Why 'Not'?', *Mind* 99 (1990), 221-238), Neil Tennant ('Negation, Absurdity and Contrariety', in *What is Negation?* edd. Dov Gabbay and Heinrich Wansing, (Dortrecht: Kluwer, 1999) 199-222) and to some extend, following them, Ian Rumfitt ('"Yes" and "No"', *Mind* 109 (2000), 781-823).

for it. It has been debated whether $A \vee \neg A$ is a logical law, whether from $A \& \neg A$ everything follows, and whether there is an understanding of negation on which the negation of a sentence containing 'presupposition failures' is defective in the same way the sentence itself. A more arcane dispute is the one whether contradictions can be true.

These disputes cannot be settled if negation is a primitive. A primitive is formalised on the basis of intuition, reflection and conceptual investigation. But as these diverge, each rival camp will start of with a different negation. There is then no basis for deciding the issue between them. What is missing is a common ground on which to debate which formalisation is 'the correct one'.

Defining negation in terms of something else promises to provide a basis on which a principled decision can be made concerning which is the right way of formalising negation. From the perspective of the *justification of deduction*, finding such a basis for settling disputes over very fundamental logical laws is the Holy Grail of Logic. It is a criterion for the success of a theory of the justification of deduction if it can settle disputes over logical laws.

4 Why Dummett Fails

Dummett's *proof-theoretic justification of deduction* was designed to settle this debate by formulating neutral requirements for definitions of the logical constants. Primitives of the theory are a 'thin' notion of truth and rules of inference.

I have shown that this project fails, because *negation* cannot be defined in this way. [For details see Section 4 (esp. 4.2.1) of 'What is wrong with classical negation?] Negation must enter the theory as an additional primitive. As a consequence, Dummett's theory cannot settle the debate between classicists and intuitionists.

5 Incompatibility

The most prominent way of remedying Dummett's theory is to define negation in terms of some notion of incompatibility—be it that its incompatibility between facts (Tennant), propositions (Brandom) or speech acts (Price, Rumfitt).

A superficial look at the question whether negation can be defined in terms of incompatibility may elicit an obvious response. Interestingly, however, there are two contradictory such responses: 'obviously yes' and 'obviously no'. The 'obviously no' camp would point out that 'incompatibility' is a negative notion; thus the definition is circular—this is in fact, I think, Russell's reaction to the proposal. The 'obviously yes' camp would point out that there are several ways of defining $\neg A$ that use some notion of incompatibilty, for instance the Sheffer Stroke 'not both p and q'.

At a more reflected level, what the 'obviously yes' camp needs to address is the question what the theoretical advantages of defining negation in terms of some notion of incompatibility are. The 'obviously no' camp needs to address the point that no circularity arises as a primitive notion of incompatibility, although undoubtedly a negative notion, is not analysed any further as 'not compatible'.

The fundamental observation behind defining negation in terms of incompatibility is that there seems to be something incompatible about 'a is red' and 'a is green', without this relying on negation. It seems to be straightforward how to define negation on this basis, without any circularity: if p implies that a is red and that a is green, this should suffice for $\neg p$ to be true.

5.1 Some Unsuccessful Definitions

There are certain accounts of incompatibility that may be ruled out. For instance, if 'p is incompatible with q' amounts to 'one of them is the negation of the other', then the approach is either circular or a dispute over whether negation can be defined in terms of incompatibility is merely verbal. Such a notion of incompatibility is simply not different enough from the notion of negation to make the project worth while: negation is merely sold under a new heading.

For similar reasons, incompatibility should also not amount to something like 'not both', only expressed without the not. 'Not both', or rather 'neither-nor', may be viewed as a generalised negation, which applies to a number of sentences rather than only to one.

That 'not both' is not a suitable notion of incompatibility can also be seen by considering that on our intuitive understanding of 'incompatibility', no contingent or logically true sentence p is incompatible with itself. Quite to the contrary: 'p is incompatible with p' suggests itself as a definition of

'p is a contradiction'. If p is contingent or logically true, 'p is incompatible with p' should be a contradiction, but 'not both p and p' is not.

What is probably the most obvious way of explaining the notion of incompatibility can also be ruled out, namely to explain 'p is incompatible with q' as 'If p is true, then q is false, and if q is true, then p is false'. One reason has already been given, namely that an approach which appeals to truth and falsity is unlikely to succeed without an appeal to negation, as something has to be said about the relation between these two notions. Furthermore, if a definition of negation in terms of incompatibility helps itself to the notions of truth and falsity, one might as well define negation right away through the equivalence '$\neg p$ is true if and only if p is false'. The notion of incompatibility would appear to be superfluous, as all the work could be done by the notions of truth and falsity.

To sum up, if a definition of negation in terms of incompatibility is proposed, then there should be a genuine difference between negation and incompatibility and the notion of incompatibility should do some real work. Otherwise the dispute is merely verbal or there are no theoretical benefits to be gained from employing the notion of incompatibility as a primitive.

5.2 Tennant's Incompatibility

5.2.1 Outline

Huw Price has put the general idea behind defining negation in terms of incompatibility very neatly: 'it is appropriate to deny a proposition p (or assert $\neg p$) when there is some proposition q such that one believes that q and takes p and q to be incompatible'.[2] Neil Tennant proposes a revision of Dummett's theory in this direction. He suggests to view $\perp$ as a 'structural punctuation marker'[3], which registers 'metaphysico-semantical fact[s] of absurdity'[4], such as 'a is red and a is green' or 'a is here and a is over there simultaneously'. $\perp$ is subject to the rule

$$\frac{A_1 \qquad \ldots \qquad A_n}{\perp} \tag{1}$$

'where by this we are to understand that A_1 to A_n are not jointly assertible,

[2]Huw Price, loc. cit., p.231.
[3]Tennant, loc. cit., p.199
[4]Ibid. p.202

that they are, that is, mutually inconsistent'[5]. According to Tennant, any speaker of a language grasps that certain atomic sentences are incompatible with each other. The notion of inconsistency 'arises by virtue of what the sentences mean and *various ways that we understand the world simply cannot be.*'[6]

Tennant goes on to give a proof-theoretic definition of negation in terms of introduction and elimination rules for it:

$$\cfrac{\begin{array}{c}\overline{A}^{\,i}\\ \Xi\\ \bot\end{array}}{\neg A}\,i \qquad\qquad \frac{\neg A \qquad A}{\bot} \tag{2}$$

As $\bot$ can only be arrived at if mutually incompatible sentences have been derived first, the introduction rule for $\neg$ captures the thought that $\neg A$ is true just in case A entails mutually incompatible sentences. The elimination rule is chosen because it is harmonious with the introduction rule.[7]

Tennant's rules are of course to be understood as holding for an *interpreted* language, not a formal calculus. Theorems of the form $\vdash A_1 \supset . A_2 \supset . \ldots A_{n-1} \supset \neg A_n$ can be deduced, which are not true on all interpretations of the formal language, but only on those which interpret $A_1 \ldots A_{n-1}$ as 'mutually inconsistent'.

5.2.2 Problems

According to Tennant, $\bot$ this is not a proposition at all: it is a 'punctuation mark'—one could as well use a blank space. Hence it is also not something which is always to be interpreted as being false. This has the strange consequence that interpreting $A_1 \ldots A_n$ as sentences which may be true together cannot result in the rule becoming unsound. This, of course, is merely a rhetorical point, just as insisting on calling $\bot$ a punctuation mark rather than a proposition is mere rhetoric. Certainly nothing in the rules Tennant has formulated dictates this interpretation. What is more serious is that the use of empty spaces may well be counterproductive in Tennant's framework,

[5]Ibid. p.217

[6]Ibid. p.217

[7]Tennant puts certain restrictions on these rules to fit his intuitionistic relevant logic, which need not concern us here. The three rules do not suffice to prove *ex falso quodlibet*. This could be remedied by adding a principle *ex adversis propositionibus quodlibet sequitur*.

as the validity of rules would then have to be explained with reference to notions of truth and falsity (cf. also 'What is wrong with classical negation?', section 4.3.3).

On a less *ad hominem* note, what turns out to be a substantial problem for Tennant's approach is an attempt to express in the object language that sentences are incompatible. So far, it is not possible to express this in Tennant's object language, as it only has interpreted sentence letters and the logical constants $\neg, \bot, \supset, \vee, \&, \forall, \exists$. Thus the language is incomplete, as obviously, we are able to say that 'a is red' and 'a is green' are incompatible. Let's use I^n as an n-place predicate of propositions, where $I^n p_1 \ldots p_n$ is to be interpreted as '$p_1 \ldots p_n$ are incompatible'.

Having extended the expressive power of the language in this way has, initially at least, the advantage of enabling us to give rules for negation that avoid the detour through $\bot$. Let's restrict consideration to $n = 2$, and write Ipq. Modifying Tennant's introduction rule for negation in the extended framework yields the following:

$$\dfrac{Iq_1q_2 \qquad \begin{matrix}\overline{\;p\;}^{\,i} \\ \Xi_1 \\ q_1\end{matrix} \qquad \begin{matrix}\overline{\;p\;}^{\,i} \\ \Xi_2 \\ q_2\end{matrix}}{\neg p}\,i \tag{3}$$

$\neg p$ may be inferred p entails: $q_1 \ldots q_n$ and $I^n q_1 \ldots q_n$.

This rules capture the fundamental idea behind the definition of negation in terms of incompatibility. But using this rule alone to govern I results in too weak a logic of I. Given our intuitive understanding of incompatibility, we should have $Ip\neg p$, i.e. 'p and $\neg p$ are incompatible', as a theorem. However, given only (3), this is not possible. Suppose you add the connectives I and the rule (3) to classical logic formalised in $\neg$ and $\supset$. It is easily shown that $Ip\neg p$ is not derivable: interpret Ipq as being true if p and q are both false, and false otherwise. This interpretation, together with the standard interpretation of the connectives $\neg$ and $\supset$, every assignment of truth-values to the atomic propositions satisfies all rules and axioms of the calculus, but no assignment satisfies $Ip\neg p$. Thus even the full force of classical logic does not suffice to derive $Ip\neg p$ as a theorem, hence it is not derivable in Tennant's much weaker logic.

To capture the notion of incompatibility more adequately, further rules governing I must be added. But which rules? Obviously it would be counterproductive to add $Ip\neg p$ as an axiom, as that would mean to characterise incompatibility with reference to negation. The rules we add must not use negation, if the approach of defining negation in terms of incompatibility is not to be thwarted.

Given Tennant's proof-theoretic outlook, the obvious first step towards more rules for I would be to try to formalise rules harmonious to (3). This meets with some difficulties, which are closely connected to the problem of formulation harmonious elimination rules for negation this is the introduction rule for negation:

$$\dfrac{\begin{matrix} \overline{A} \\ \Pi \\ \neg B \end{matrix} \qquad \begin{matrix} \overline{A} \\ \Pi \\ B \end{matrix}}{\neg A}$$

The harmonious elimination rule would be *ex contradictione quodlibet*:

$$\dfrac{\neg A \qquad A}{B}$$

But this rules leads to maximal formulas which cannot be removed from deductions in such a way that no negation rule is used in the transformation.

The remedy that can be used in the case of negation also works for the case of I. We need to use $\perp$. It has a straightforward introduction rule, which captures in one rule exactly the spirit, if not the letter, of Tennant's rule (1):

$$\dfrac{p \qquad q \qquad Ipq}{\perp} \tag{4}$$

$\perp$ may be derived if two sentences have been derived which are incompatible. Negation can then be defined by Tennant's rules (2).

Applying the principle of harmony to (3) yields the following further rule governing I:

$$\dfrac{\begin{matrix} \underbrace{\overline{p}^i \quad \overline{q}^i} \\ \Pi \\ \perp \end{matrix}}{Ipq}\, i \tag{5}$$

This rules is an introduction rule for I.

Adding this rule does indeed yield $Ip\neg p$ as a theorem. However, it also yields something more, namely $\neg p \vdash Ipp$. This is quite unacceptable, at least for atomic p, given the intended interpretation of I, as noted before, as any contingent proposition is compatible with itself.

Hence rule is too strong for the intended interpretation of I as incompatibility. But on Tennant's proof-theoretic approach, he cannot easily evade the point that (5) is the additional rule governing I, as this is required by (4) and the principle of harmony. The connective governed by the rules (4) and (5) is of course the Sheffer function 'not both, p and q'. This is as close as we can get towards a notion of incompatibility in classical and intuitionistic logic. But it is not close enough. It does not capture many intuitions about incompatibility correctly. Hence following up Neil Tennant's notion of incompatibility thus does not lead to a convincing notion at all. In fact, given the difficulties surrounding formalising satisfactory rules for possibility in the proof-theoretic framework one can suspect that it is equally problematic in this framework to formalise the notion of incompatibility, which of course is also a modal notion.

There is thus a lack of fit between Tennant's proof-theoretic approach and his appeal to a primitive notion of incompatibility. There are no theoretical advantages to be had from this choice, rather than choosing negation. In fact, it seems positively harmful, as the notion of incompatibility is not one that can be adequately expressed in Tennant's own framework.

5.3 Brandom's Incompatibility

Robert Brandom attempts to give a semantics with the notion of incompatibility as as the primitive which not only covers propositional logic, but also modal operators. According to Brandom, '*incompatibility* can be thought of as a sort of conceptual vector product of a *negative* and a *modal* component. It is *non*-com*possibility*.'[8] It would of course be a blatant circularity to claim that incompatibility is defined as non-compossibility, and then to claim that negation can be defined in terms of this notion. So Brandom's remarks must be understood as merely heuristic, to get us on the right track of what notion of incompatibility he has in mind.

Brandom's heuristic procedure does, however, reveal that incompatibility is a more complicated notion than negation, and thus is not as good a choice

[8]Locke Lecture 5, p.16

for a primitive than negation. Brandom needs to appeal to the notions of conjunction, negation and possibility to get us on the right track of what he means by 'incompatible', because we have fairly good understanding of the former notions, but not really of the latter. In fact, Brandom himself characterises incompatibility in different ways which do not match up. In fact, two different ways of characterising incompatibility occur in one and the same passage: 'to say that one way things could be is *incompatible* with another is to say that it is not *possible* that the second obtain if the first does—that if the first does, it is *necessary* that the second does not.'[9] Thus p is incompatible with q is on the one hand said to be equivalent to $\neg\Diamond(p \supset q)$ and on the other hand to $\Box(p \supset \neg q)$, i.e. $\neg\Diamond(p \& q)$.

This may of course have just been a slip of the pen. But the equivocation might also have a deeper reason. If the first reading is adopted, it would indeed be a contradiction to say that p is incompatible with p, which is desirable given our intuitive understanding of this notion, as $\Diamond(p \supset p)$ is a logical truth (at least in **D** and hence in **S5**). On the second reading, incompatibility is non-compossibility. But this notion doesn't quite match up with our intuitive understanding of incompatibility, at least not if the possibility used here is the one of **S5**, which, Brandom argues, is the modal logic that turns out to be validated by his incompatibility semantics. To see this, let's have a look at compossibility and compatibility. We should expect them to be the same concepts, on Brandom's account of incompatibility as non-compossibility. The problem is that every contingent or logically true sentence should be compatible with itself:[10] if a sentence is not compatible with itself, that would suggest that it is a contradiction. So 'p is compatible with p' is logically true, if p is such a sentence. However, this shows that $\Diamond(p \& q)$ cannot correctly be interpreted as compatibility, for $\Diamond(p \& p)$ is not a logical truth, at least not in Brandom's modal logic. Compatibility thus is not compossibility, at least not in the most obvious sense.

The notion of incompatibility is not one that is easily pinned down: it seems close to non-compossibility, but as compossibility doesn't seem to be the same as compatibility, it isn't clear how close it is. Incompatibility is thus not a good primitive: our intuitive, pre-theoretic understanding of it is not firm enough.

That our intuitions leave us behind when considering properties of Bran-

[9]Ibid p.10f

[10]Brandom agrees, I've asked him.

dom's incompatibility is not surprising if one takes into account the object language connective expressing this notion. Brandom seeks to employ incompatibility as the sole primitive of the semantic theory. Thus what corresponds to it in the object language is a connective in terms of which all connectives of **S5** can be defined—it is the modal version of the Sheffer Stroke. The former is rather more complicated than the latter: $*$, to be interpreted as 'is incompatible with', suffices to define all other operators of **S5**, where $p * q$ is equivalent to $\neg\Diamond(p\&q) \vee (\Diamond(p\&q)\&\Diamond(p\&\neg q)\&\neg(p\&\neg q)) \vee (\Diamond(p\&q)\&\neg\Diamond(p\&\neg q)\&(p\&q))$. This connective is arguably not one of which we have an immediate, pre-theoretical understanding. In particular, it is not Brandom's 'non-compossibility'. Besides, it is worth noting that $*$ does not adequately capture an intuitive notion of incompatibility adequately: it is not logically true that $p * p$, which is equivalent to $\neg\Diamond p \vee (\Diamond p\&p)$, at least not in **S5**. In fact, $p*p$ can be used as the definition of $\neg p$, so another reason why $*$ does not express our intuitive notion of incompatibility.

In conclude that Brandom's notion of incompatibility is not a suitable primitive. It is not clear what he has in mind when he speaks about incompatibility. Whenever he is explicit, it does not match up with other plausible requirements on a notion of incompatibility.

It is also worth noting the two incompatibility theorist Brandom and Tennant must have different notions of incompatibility in mind, despite the fact that their heuristic explanations of this notion are virtually identical: Tennant claims that a logic based on this notion is intuitionist (or, more precisely, the negation of his idiosyncratic intuitionist relevant negation), but Brandom argues that negation turns out to be classical. It is plausible to surmise that this is due to differing heuristic explanations of the notion of incompatibility. Tennant favours a 'verificationist' notion of truth, whereas Brandom favours a pragmatist one, which then means that 'p and q cannot be true together' has different properties on each reading. It is thus questionable whether choosing the notion of incompatibility, rather than, say, negation, as a primitive succeeds in providing a neutral basis for settling the debate between classicists and intuitionists. The problem is that, because we haven't got a strong enough pre-theoretic understanding of incompatibility, we need to resort to heuristic readings, which then smuggles illegitimate presuppositions into the theory.

As mentioned earlier, it is a criterion of success for a theory aiming at a justification of deduction that disputes over logical laws can be settled on its basis. However, as a matter of fact choosing incompatibility as a primitive fails to solve the question whether negation is classical or intuitionist, as different incompatibility theories come to different conclusions about what kind of negation turns out to be definable in terms of incompatibility. Thus much of the motivation for choosing this primitive, rather than negation, has been lost.

6 Concluding Reflections on Incompatibility

There is something that the pairs 'a is red' and 'a is green', and 'a is here' and 'a is over there' have in common, and we can call this relation 'incompatibility'. It is not difficult to give a general explanation of what incompatibility consists in: two sentences are incompatible, if they cannot be true together, or alternatively, if each entails the negation of the other. These are general characterisations of incompatibility, which make no reference to the specific content of the sentences which stand in this relation. Neither of them, however, is what theorists have in mind who propose to define negation in terms of incompatibility, as they are talking about a notion of incompatibility not explained any further in terms of truth, falsity and negation. Their notion of incompatibility is intimately tied to the specific *content* of sentences, rather than to general features of classes of sentences, such as truth, falsity or entailing negations of other sentences. In fact, the whole point seems to be that the notion is one tied intimately to the content of sentences, rather than being one that could be explained in a formal manner.

The last paragraph leads me to suspect that incoherent requirements need to be imposed on the notion of incompatibility. On the one hand, it is a notion tied to the particular content of sentences, on the other it needs to be a notion that applies across the board independently of the content of sentences, like a logical constant.

Now some pairs of sentences don't exhibit this incompatibility, even though they may be said to exclude each other. One needs the *right kind of exclusiveness*: it would not suffice for logic that one can derive, say, 'Beetroots are revolting' and 'Scotch is disgusting'. There is a sense in which these two sentences exclude each other and cannot be true together – obviously the second is false and the first true – but that would merely result in a logic for my personal prejudice. That is to say, only *certain* atomic sentences which may be said to exclude each other could be used in a definition of negation in terms of incompatibility. 'a is red' and 'a is green' seem to exclude each other in the right way, but 'Scotch is disgusting' and 'Beetroot are revolting' do not, because of their respective meanings. Hence the reasons why 'a is red' and 'a is green' constitute the right kind of exclusiveness is a matter of their *particular content*. If we characterise two atomic sentences as excluding each other this can only be because of their content. However, in order for the notion of exclusiveness to be of use in a definition of *negation*, rather than merely some indication that we find certain sentences unacceptable, there needs to be a *general* method of determining for *any* two atomic sentences whether or not they exclude each other in the desired way. We need to have a way of telling when we have arrived at two sentences which exclude each other in the right way. A general method is mandatory because the negation to be defined should cover any possible extension of the language by new atomic sentences: for any atomic sentences we may add to the language, it needs to be determined which pairs exclude each other. But this is precisely to say that the method needs to *abstract* from the content of atomic sentences. Hence the desired method for determining whether two atomic sentences exclude each other in the right way has to be general and independent of the content of the atomic sentences and at the same time cannot be general, but due to its nature must be particular and tied to the content of the atomic sentences. Hence there is no such method of characterising the right kind of exclusiveness of atomic sentences.

The only way I can see of reconciling these two opposing requirements is to say that, for instance, the reason why 'a is green and a is red' constitute the kind of exclusiveness is that what is green cannot be red and conversely, if something is red, it is not green, hence if something is red as well as green, it is green as well as not green. But this makes use of negation.