The logic of knowledge

BENJ HELLIE

PHLB20 Lecture Notes 1

1 Truth and falsity

A true claim

Snow is cold

A false claim

Snow is hot

We say that truth and falsity are 'truth-values'. Saying 'the truth-value of P is *truth* (or *falsity*)' is a fancy way of saying 'it is true (or false) that P'.

2 Negation

- $\neg P \quad = \quad$ it is not the case that P

Double negation elimination

$$\neg\neg P = P$$

Why? Hint: think of negation as flipping the truth-value of a claim it is applied to.

3 Conjunction and disjunction

- $P \wedge Q \quad = \quad P$ and Q
- $P \vee Q \quad = \quad P$ or Q

  We always use 'inclusive or', aka 'and/or', so that $P \vee Q$ is *compatible with* $P \wedge Q$: sometimes, both are true.

DeMorgan's Law

$$P \wedge Q = \neg(\neg P \vee \neg Q),$$

$$P \vee Q = \neg(\neg P \wedge \neg Q)$$

Try to show that one of these implies the other. Hint: set

$$P = \neg S$$

and

$$Q = \neg T$$

and then use double negation elimination.

What happens to DeMorgan's law if we use 'exclusive or', where $P \vee Q$ is incompatible with $P \wedge Q$?

4 Conditionals

- $P \supset Q \quad = \quad$ if P, Q
- $P \subset Q \quad = \quad$ only if P, Q

  Terminology: in $P \supset Q$, we say that P is the 'antecedent' and Q is the 'consequent'.

Contrapositives

$$P \supset Q = \neg P \subset \neg Q$$

Converses

$$P \supset Q \neq P \subset Q$$

When I say that converses are not equal to one another, I mean that sometimes $P \supset Q$ is true and $P \subset Q$ is false, or vice versa.

But sometimes both are true. When that happens, we say that P and Q are 'equivalent'. We then use the following symbolism:

- $P \equiv Q \quad = \quad P$ just if Q

  Sometimes people write 'iff' or 'if and only if' instead of 'just if'. When we say 'P if and only if Q' we mean that very literally: we mean 'P if Q, and P only if Q'. We mean that the arrow runs in both directions.

5 Technical aside

In philosophy, we define the *material conditional*:

- $P \supset Q \quad = \quad \neg P \vee Q$.

We do this for the following reason. Things go more smoothly if we can say claims like this are always true:

someone is in Toronto $\supset$ they are in Canada

By this we mean (roughly) that every claim on the following list is true:

Aaron is in Toronto $\supset$ he is in Canada; Amelia is in Toronto $\supset$ she is in Canada; ...; where we plug in each person's name.

Suppose that Aaron is in Toronto. Then he poses no problem for the claim assuming that he is also in Canada.

Suppose that Amelia isn't in Toronto. Then she poses no problem for the claim.

So there are two ways for the claim to be true: the Aaron way, where the consequent is true; and the Amelia way, where the antecedent is false.

If neither of these circumstances is met, the claim is false. If Zoran *is* in Toronto but *isn't* in Canada, the claim would be false (of course that doesn't happen).

So the claim $P \supset Q$ is true exactly when either P is false or Q is true; false when neither of these is met (namely, both P is true and Q is false). So it is equivalent to $\neg P \vee Q$.

Now the material conditional is sort of weird: it is true whenever the antecedent is false. So 'Martians teach us the secret to world peace $\supset$ we don't know the secret to world peace' is true! Of course 'if Martians taught us the secret to world peace, we wouldn't know the secret to world peace' is false. So $\supset$ and 'if' don't mean exactly the same thing. For the most part, fortunately, this is an issue we don't need to worry about.

6 Knowledge and belief

- $\mathsf{K}_S P$ = S knows that P
- $\mathsf{B}_S P$ = S believes that P

Knowledge entails truth

$$\mathsf{K}_S P \supset P$$

'One cannot know, what is not so'. If lupini beans are yellow, no one knows that they are green.

Knowledge entails belief

$$\mathsf{K}_S P \supset \mathsf{B}_S P$$

A person's knowledge is stored away in their opinions, in their view of the world, in how they think things are. If Roland has no opinion on the question of whether there are an odd number of trees, Roland certainly does not know that there are.

It follows that ...

Knowledge entails true belief

$$\mathsf{K}_S P \supset (\mathsf{B}_S P \wedge P)$$

The converses are all false:

True belief does not entail knowledge

$$\neg(\mathsf{K}_S P \subset (\mathsf{B}_S P \wedge P))$$

If someone 'gets lucky', we don't want to say they know. Frank is super-confident that his lottery ticket will win (not because it's fixed). And lo and behold it does! But we don't think that Frank knew his lottery ticket would win. That is something no one could have known.

Truth does not entail knowledge

$$\neg(\mathsf{K}_S P \subset P)$$

Let P be any fact about which S is ignorant (say, whatever the cheapest place for tzatziki in North York is). Here P is true but S does not know that P.

Belief does not entail knowledge

$$\neg(\mathsf{K}_S P \subset \mathsf{B}_S P)$$

Let P be some claim that S is mistaken about (say, S believes that her boots are still in the basement but in fact her mother threw them out). Here S believes that P but S does not know that P.

But *sometimes* ...

'Merely' true belief is enough for knowledge!

> Spy Natasha is trying to communicate to spy Boris where the microfilms are buried but does not want to say exactly. Bullwinkle, who is a big goofball, somehow got it into his head that they are buried in the pumpkin patch (and in fact they are). Natasha knows that Bullwinkle has this belief. She tells Boris 'Bullwinkle knows where the microfilms are buried'. Should Boris complain 'no, he just got lucky!'? No he does not. So sometimes a completely lucky true belief can count as knowledge???

What is going on here? This is a big mystery, and one we will be investigating over the semester.

7 Knowledge, belief, and the first-person

Consider the following dialogues:

A: Do you know where Prill Avenue is?

Z: No.

B: Let me try another tack: where is Prill Avenue?

Z: Ah, it's two lights up, on the right.

A: I thought you told me you didn't know where it is!

Z: Well I don't.

A: ???

C: Where is Prill Avenue?

Y: [shrugs shoulders]

C: So you don't know where it is?

Y: Hey *I* know better than anyone that Prill Avenue is two lights up, on the right.

C: Why didn't you tell me when I asked you where it is?

Y: How do you expect me to answer that?

C: weirdo!

D: Where is Prill Avenue?

X: [shrugs shoulders]

D: So you don't know where it is?

X: Nope, sorry.

D: Absolutely no opinion on the matter?

X: Oh yeah, of course—I'm confident that it's two lights up, on the right.

D: Why didn't you say so?

X: Do I look like I know where Prill Avenue is??

D: Nee-nee-nee-nee, nee-nee-nee-nee ...

X, Y, and Z are of course being completely bizarre.

First-person expectations

⇒ We expect a person to treat 'I believe that P' and 'I know that P' in exactly the same way: either affirming both or denying both.

⇒ And we expect a person to affirm P, 'I believe that P', and 'I know that P' under exactly the same circumstances.

⇒ Though notice that *denying* that P is something we do under fewer circumstances than we deny 'I believe that P' or 'I know that P'. If we are uncertain about P, we would not deny P but would deny 'I believe/know that P'. We deny P = affirm $\neg P$ only when we not only deny 'I believe/know that P' but more strongly affirm 'I believe/know that $\neg P$'.

- Someone is *incoherent* under the sorts of circumstances we see with X, Y, and Z: if we just can't make sense of what they are about.
- $P \vdash Q$ = if someone affirms P, they are incoherent unless they affirm Q.

Then let's use a special variant of our notation for knowledge in the 'first-person' case, when someone is talking about themself:

- $\mathsf{K}P$ = I know that P
- $\mathsf{B}P$ = I believe that P

  Note the absence of subscript.

Our thought that a person should affirm either all or none of P, 'I believe that P', and 'I know that P' can be expressed like this:

First-person equivalences

$$P \dashv\vdash \mathsf{K}P \dashv\vdash \mathsf{B}P$$

8 Moore's paradox

Now we can notice a range of weird effects.

8.1

First, although

$$\mathsf{K}P \dashv\vdash P \dashv\vdash \mathsf{B}P$$

also

$$\neg\mathsf{K}P \dashv\not\vdash \neg P \not\dashv\vdash \neg\mathsf{B}P$$

That is to say, although ordinarily when two claims are equivalent their negations are equivalent, that is not true in this case.

8.2

Second, what goes for the first-person does not go for the third-person. Although as we just saw

$$P \dashv\vdash \mathsf{K}P \dashv\vdash \mathsf{B}P$$

as we know

$$P \dashv\nvdash \mathsf{K}_S P \not\dashv\vdash \mathsf{B}_S P$$

8.3

To see a third odd result, consider the following dialogue:

> E: Where is Prill Avenue?
>
> W: I don't know.
>
> E: What about that guy over there?
>
> W: Oh, you mean V? He knows that Prill Avenue is two lights up, on the right.
>
> E: Why didn't you say so?
>
> W: He knows, I don't. What do I look like, his mother?
>
> E: Remind me not to come back to this neighborhood any time soon.

Here we think that

$$\mathsf{K}_S P \vdash \mathsf{K}P$$

This is weird enough in itself: there is no incoherence in saying that Bob is over seven feet but I am not.

But of course there is no problem with the idea that I know that *P* but no one else does. Hence:

$$\mathsf{K}_S P \not\dashv\vdash \mathsf{K}P$$

Three weird effects

1. From the first person, belief in, knowledge of, and truth of *P* are equivalent *but* their negations are not
2. From the first person but *not* the third person, belief, knowledge, and truth are equivalent
3. Third person knowledge is contagious to first-person knowledge but not vice versa

The second of these is what is properly known as 'Moore's paradox', but all of them could be thought of as 'Moore-paradoxical' in a broad sense.

---

9 Summary

This table summarizes the logic of knowledge:

$$\begin{array}{ccccc}
 & & \mathsf{K}_S P & \vdash & \mathsf{B}_S P \\
 & & \top & & \\
P & \dashv\vdash & \mathsf{K}P & \dashv\vdash & \mathsf{B}P
\end{array}$$