Propositional Modal Logic

Modal logic is the logic of necessity, possibility and other related notions. We introduce two new sentential operators ◊ and □, to be understood as 'Possibly' and 'Necessarily' respectively. (Sentential operators are devices that take sentences to form new sentences: thus, in adding ◊ and □ to standard propositional logic we can form such sentences as □P, ◊(P ∧ Q), P ∧ □Q, □◊P.) We use these to formalize a number of locutions of English. If P is the sentence 'Pigs can fly', then ◊P formalizes such English sentences as 'Pigs might have flown' 'Pigs might have been able to fly' 'It could have been the case that pigs fly', 'It's possible for pigs to have flown'. Similarly if Q is the sentence '1+1=2' then □Q formalizes such sentences as '1+1 has to be 2' '1+1 is necessarily 2' '1+1 couldn't have failed to be 2'.

In understanding possibility and necessity we employ the device of *possible worlds*. A possible world is a way that the world, i.e. everything, might have been. ('Possible *world'* is thus a bit of a misnomer; 'possible *universe'* would have been better.) There is a great deal of controversy over what possible worlds are: whether they really exist, or whether they are just things that we somehow construct in the imagination; however we do not need to get into this here. (Those interested might have a look at David Lewis's book *On the Plurality of Worlds*; he thinks that they really exist.) What is important for our purposes is that they provide a way of understanding talk of possibility and necessity. Roughly, sentence that is necessarily true is one that is true in every possible world, and a sentence that is possibly true is one that is true in at least one possible world.

Before making this idea more precise, let us pause to mention a use of possibility talk which cannot be accommodated in the possible worlds framework; and some other uses which, despite initial appearances to the contrary, can. First, the use which cannot be accommodated. Suppose I say to you, out of the blue, 'Is 329 x 747 equal to 254,763?' Unless you are remarkably good at mental arithmetic, you are likely to say something like 'I don't know; it's possible that it is, but I'd have to work it out'. However, as you'll find if you do work it out, 329 x 747 is in fact equal to 245,763. But if that is true in this world, then it is true in any possible world. So when you said it was possible that it was equal to 254,763, your utterance cannot be understood in possible worlds terms. In fact this kind of possibility, sometimes called 'epistemic possibility' since it concerns what one does not know is not the case, is not at all easy to analyse, and there is no agreed account of it.

We now turn to a set of possibility claims that can be understood in the possible worlds framework, despite initial appearances to the contrary. Suppose we say that it's not possible for someone to travel faster than light, or that, given their anatomy, it's not possible for pigs to fly, or that we can't possibly go to the cinema at 1 o'clock on Wednesday since we have an important metaphysics class. In each of these cases we are not asserting that there is no possible world in which someone goes faster than light, or that pigs fly, or in which we do go to the cinema at 1 o'clock on Wednesday. But we can still understand these claims in possible world terms; we

simply need to restrict the set of possible worlds that we are considering. Thus, in the first case, there is no possible world in which the actual laws of physics hold in which someone travels faster than light; in the second, there is no possible world in which the actual laws of physics hold, and pigs have their actual anatomy, and pigs fly; in the third, there is no possible world in which my obligation to go to classes is as it actually is, and I live up to those obligations, and I go to the cinema at 1 o'clock on Wednesday. (This last example might need some fine tuning; but I trust that the idea is clear.)

We now turn to ask how this talk of possible worlds might be formalized. What we have said so far might suggest that we can think of the modal operators as quantifiers that range over possible worlds: $\Box$P is true just in case $\forall$w (P is true in w); $\Diamond$P is true just in case $\exists$w (P is true in w). Unfortunately things are not quite so simple. In understanding the modal operators in this way we have prevented ourselves from saying various things that we might well want to say. The point becomes clear once we come to iterated (repeated) modalities. Suppose we want say that whilst a certain state of affairs isn't possible, it could have been possible. In short: it is possible that it is possible. That seems to make some sense: it seems plausible that $\Diamond\Diamond$P is weaker that $\Diamond$P. How can we spell this out in terms of possible world? The standard way is this: rather than just saying that $\Diamond$P is true just in case there is a world in which P is true, say that it is true just in case there a world *which we can get to from this world*, in which P is true. Then we can make sense of the idea that it is possible that it is possible that P without it being possible that P as follows: whilst from this possible world we can't get to another world in which P, we can get to a world in which possibly P; and from that world we can get to another world in which P. What we need for this then, is the idea of there being some worlds that you can get to directly from this world, and others that you cannot; or, in the jargon, an *accessibility relation*. Certain worlds are *accessible* from others.

We can then state all this a little more generally and formally. $\Box$P is true at a world w_1 iff every world which is accessible from w_1 is a world in which P is true; similarly $\Diamond$P is true at w_1 just in case there exists a world w_2 which is accessible from w_1 in which P is true. $\Diamond\Diamond$P will be true at a world w_1 just in case there exists a world w_2 which is accessible from w_1, in which $\Diamond$P is true, which in turn will be true just in case there exists a world w_3 which is accessible from w_2, and in which P is true. Thus, in order for $\Diamond$P to be false and $\Diamond\Diamond$P to be true, it must be possible for w_2 to be accessible from w_1, and for w_3 to be accessible from w_2, whilst w_3 is not accessible from w_1. In short, the accessibility relation must not be transitive. Once we insist that the accessibility relation is transitive, $\Diamond\Diamond$P will be equivalent to $\Diamond$P; and ($\Diamond\Diamond$P $\leftrightarrow$ $\Diamond$P) will be a tautology.

Once we think this way, we see that there are a host of different conditions that we might put on the accessibility relation. We could insist that it be symmetric (for all worlds, w_1, w_2, if w_1 is accessible from w_2 then w_2 is accessible from w_1), or reflexive (every world is assessible from itself), or transitive (for all worlds, w_1, w_2, w_3, if w_1 is accessible from w_2 and w_2 is accessible from w_3, then w_1 is accessible from w_3) or any combination of these. Each different combination will generate a different logical system, with its own set of tautologies (or *theorems* as they are more standardly called). In fact some of these combinations are more interesting and more widely discussed than others. We shall look at just five.

There are no restrictions on the accessibility relation: K

Characteristic theorems: $\Box(A \rightarrow B) \rightarrow (\Box\, A \rightarrow \Box\, B)$ *The K axiom*

$\Box\,(A \wedge B) \leftrightarrow (\Box\, A \wedge \Box\, B)$

The accessibility relation is reflexive: T

Characteristic theorems: $(\Box\, A \rightarrow A)$ *The T axiom*

$(A \rightarrow \Diamond\, A)$

$(\Box\, A \rightarrow \Diamond\, A)$

The accessibility relation is reflexive and symmetric: B

Characteristic theorems: $(A \rightarrow \Box\, \Diamond\, A)$ *The B axiom*

The accessibility relation is reflexive and transitive: S4

Characteristic theorems: $(\Box\, A \rightarrow \Box\, \Box\, A)$ *The S4 axiom*

$(\Diamond\, \Diamond\, A \rightarrow \Diamond\, A)$

The accessibility relation is reflexive, transitive and symmetric: S5

Characteristic theorems: $(\Diamond\, A \rightarrow \Box\, \Diamond\, A)$ *The S5 axiom*

$(\Diamond\, \Box\, A \rightarrow \Box\, A)$

Requiring the accessibility relation to be reflexive, transitive and symmetric is to require that it be an equivalence relation. This isn't the same as saying that every world is accessible from every other. But it is to say that the class of worlds is split up into classes within which every world is accessible from every other; and there is no access between these classes. S5, the system that results, is in many ways the most intuitive of the modal systems, and is the closest to the naive ideas with which we started.

Further Reading

Samuel Guttenplan, *The Language of Logic* (Oxford, 1986). Has a short chapter on modal logic, which contains a useful discussion of the different things that we might mean by possibility. However, in giving a proof system it doesn't go as far as we go in these notes.

Graeme Forbes, *The Metaphysics of Modality* (Oxford, 1985). A useful and interesting book on modal issues in general. The first chapter has a discussion of propositional modal logic which goes a bit further than we go here; the second chapter contains a discussion of predicate modal logic.

G.E.Hughes and M.J.Cresswell, *An Introduction to Modal Logic* (London 1968). The classic formal modal logic text, covering both propositional and predicate logics.

Brian Chellas, *Modal Logic: An Introduction* (Cambridge 1980). Another introductory formal text that places fairly high demands on the reader. Covers propositional modal logic only, but has a very complete discussion of the various systems that come between K and S5.