Notes on "Accessibility" and Modality

Branden Fitelson
11/05/03

In these notes, I will flesh out the relation of "accessibility" and the role it plays in both modal logic (and modal semantics), and applications of modal logic (including modal metaphysics, and other applications). First, I will discuss the *technical* work that "accessibility" does in unifying and providing an extensional formal semantics for the plethora of modal logics that have come into existence since Aristotle. In this first stage, "accessibility" will be an *uninterpreted* formal device. Then, I will discuss some applications of modal logic (to modal metaphysics and beyond). It is here that various *interpretations* of the "accessibility relation" will be introduced and put to work in concrete applications.

"Accessibility" in Modal Logic and Formal Semantics. As I mentioned in class, there have been a great many modal logics floating around since Aristotle. A modal logic is just a set of axioms and rules that govern the modal operators. There are two modal operators: "Necessarily" and "Possibly". I will use p, q, *etc.* to stand for statements, x, y, *etc.* to stand for objects (concrete particulars), and I will use P, Q, *etc.* to stand for predicates or properties. Almost all modal logics agree on the following five basic principles (or axioms):

- Necessarily p if and only if not possibly not p.
- Possibly p if and only if not necessarily not p.
- If p is a logical truth (*e.g.*, $x = x$), then necessarily p.
- Necessarily (p and q) if and only if (necessarily p and necessarily q).
- Necessarily, (p if and only if p).

Most other principles ("axioms") concerning the modal operators have been controversial, and not widely agreed upon or accepted. The most well-known and widely used of these are the following (with their tradtional names):

(T) If necessarily p, then p.

(4) If necessarily p, then necessarily necessarily p.

(E) If possibly p, then necessarily possibly p.

(B) If p, then necessarily possibly p.

Until the advent of possible worlds semantics, there was no unified understanding of the logical (and formal semantical) relationships between these various additional modal axioms (or the alternative modal logics they partake in). Moreover, there was no extensional semantics for modal claims such as these. Possible worlds semantics changed all that. Formally, possible worlds semantics (PWS) introduces "possible worlds" or "points" (for now, these w's are *uninterpreted* formal devices, but various interpretations of them will be discussed below). PWS also introduces a binary "accessibility relation" R between pairs of possible worlds (again, for now, this is an uninterpreted formal device, but various interpretations of R will be discussed below). With these devices in hand, an extensional semantics can be given which unifies the various axioms and logics listed above. The basic technique in PWS is quantification over possible worlds, subject to the relation of "accessibility". Letting $w*$ denote the actual world, we have the following two fundamental translational schema for PWS:

(†) Necessarily $p \mapsto p$ is true at *every* possible world w such that $R(w*, w)$.
Possibly $p \mapsto p$ is true at *some* possible world w such that $R(w*, w)$.

In other words, p is *necessarily* true if p is true at all possible worlds w that are accessible from the actual world $w*$. And, p is possibly true if p is true at some possible worlds w that are accessible from the actual world. For now, we will not concern ourselves with what these "possible worlds" w *are*, or with what the "accessibility relation" R *is*. These questions of the interpretation of the formal machinery of PWS will be taken up, below, in *applications* of modal logic.

For now, we will focus on the logical and semantical work that PWS can do. The main achievement is to provide a unified extensional framework for talking about modal claims. In particular, it can be shown that the following obtain:

- Axiom (T) will hold (in all PWS models) just in case the relation R is *reflexive*. That is, (T) will hold just in case $R(w, w)$, for all worlds w. Or, intuitively, if all possible worlds are "accessible from themselves." Most accessibility relations in applications of modal logic are reflexive.
- Axiom (4) will hold iff R is *transitive*. That is, (4) will hold iff $R(w_1, w_2)$ and $R(w_2, w_3)$ implies that $R(w_1, w_3)$. This is important in applications. We'll see that several important accessibility relations that appear in applications are not transitive. This implies that (4) fails in such applications.
- Axiom (E) will hold iff R is *euclidean*. That is, (E) will hold iff $R(w_1, w_2)$ and $R(w_1, w_3)$ implies that $R(w_2, w_3)$. Some interesting relations that are non-transitive (*e.g.*, similarity-based relations) are also non-euclidean. Below, we'll see some relations that are neither transitive nor euclidean.
- Axiom (B) will hold iff R is *symmetric*, That is, (B) will hold iff $R(w_1, w_2)$ implies $R(w_2, w_1)$. I won't discuss this one. Can you think of any candidate accessibility relations that are asymmetric? Hint: Lewis' accessibility relation "w_1 obeys the physical laws of w_2" (see below for discussion) is not symmetric. Can you explain why such a relation would be asymmetric, by giving a counterexample under this interpretation of "accessibility"?

So, PWS effects a reduction and unification of all disputes about purported modal axioms to disputes about the properties of the accessibility relation R.

> Old disputes give way to new. Instead of asking the baffling question whether whatever is actual is necessarily possible, we could try asking: is the relation R symmetric? [David Lewis, *The Plurality of Worlds*, p. 19]

In this sense, PWS is a neat, formal apparatus for unifying and "entensionalizing" modal discourse and the plethora of modal logics. But, the ultimate *philosophical payoff* of PWS will be in its *applications* to problems in metaphysics, epistemology, value theory, *etc.* In such applications, the *interpretation* of the "possible worlds" and the "accessibility relation" that appear in the *formal schema* (†) will be crucial. I now turn to some philosophical applications of PWS. These will be in metaphysics and epistemology, but applications of PWS and modal logic to value theory (under the rubric "deontic logic") also exist.

"Accessibility" in Applications of PWS. Lewis describes many applications of PWS (and modal logic) to metaphysics. I will discuss a couple of these, with an eye toward understanding how the translation schema (†) [and its w's and R] gets interpreted and applied. First, consider applications of PWS to modal claims in physical science. Physical necessity is usually called *nomological* necessity. The standard instantiation of (†) in this context is the following:

($\dagger_N$) p is nomologically necessary $\mapsto$ p is true at all possible worlds that are *nomologically accessible* from the actual world — *i.e.*, p is true at all possible worlds *that obey the physical laws of the actual world* $w*$.

Now, we might ask: "Does nomological necessity/possibility satisfy axioms such as (T), (4), (E), and (B)?" It would be difficult to answer this question just by thinking about the axioms. For instance, it's unclear what to say about whether nomological necessity satisfies axiom (E). Is something that is nomololgically possible nomologically necessarily possible? PWS — via ($\dagger_N$) — allows us to rephrase this question as: "Is the relation of nomological accessibility symmetric?". That is, is it so that whenever w_1 obeys the physical laws of w_2 then also w_2 obeys the physical laws of w_1? This question will be answered differently by different theories of physical lawhood. For instance, Lewis' theory of physical laws answers this question in the negative. But, virtually all theories of physical law (including Lewis') say that nomological accessibility is *reflexive* — that *all worlds obey their own physical laws*. This implies — via PWS and ($\dagger_N$) — that nomological necessity satisfies (T). That is, if p is nomologically necessary, then p is (actually) true. This is a nice example of how PWS, *when combined with an appropriate metaphysical theory (or theories)*, can illuminate the nature of various kinds of modality and the meaning of various sorts of modal claims. Of course, in such metaphysical *applications*, PWS must be *supplemented* with an *interpretation* of "possible world" and "accesssibility". In this example, "possible worlds" were understood as mereological space-time (physical) wholes, and "w_1 is accessible from w_2" was understood as "w_1 obeys the physical laws of w_2".

Similar analyses can be given of *other kinds* of modality, including logical modality, biological modality, psychological modality, etc. And, PWS, together with theories of logical, biological, psychological (or other kinds of) *laws*, will allow us to get a grip on the properties that various sorts of modality have. For instance, it is typically assumed that *all* worlds (full stop) satisfy the same logical laws. If that is true, then (assuming PWS) logical necessity will satisfy *all* of the axioms (T), (4), (E), and (B). Here, again, "accessibility" plays a key role. The standard assumption is that all worlds are *logically accessible* from all others.

And, that explains why (given PWS) logical necessity satisfies all these axioms.

PWS has applications not only in metaphysics and logic, but also in *epistemology*. We can talk about the relation of *espistemic accessibility*. A world w' is epistemically accessible from w for an agent S (in w) iff S knows nothing that would rule out the hypothesis that $w' = w$. Then, p is epistemically necessary (for S) iff p is true at all possible worlds that are epistemically accessible from w (for S). And, for instance, if we want to know whether epistemic necessity satisfies axiom (4), we can ask whether espistemic accessibility is *transitive*. If S knows nothing that rules out $w' = w$, and S knows nothing that rules out $w'' = w'$, then does it follow that S knows nothing that rules out $w'' = w$? It seems not. After all, S may not be able to distinguish w from w', and he may not be able to distinguish w' from w'', but it does not follow that he cannot distinguish w from w''. Small, imperceptible differences can add up to big, perceptible ones. You may not be able to distinguish a cup of coffee with $\frac{1}{2}$ a teaspoon of sugar from a black cup, and you may not be able to distinguish a cup with $\frac{1}{2}$ a teaspoon of sugar from a cup with 1 teaspoon. But, you may, nonetheless, be able to distinguish a black cup from a cup with 1 teaspoon of sugar. These sorts of intransitivites in perception and knowledge are common.

There are other interesting examples of intransitive accessibility relations, which lead to types of modality that violate axiom (4). Any accessibility relation that rests on a type of *similarity* between worlds will be intransitive. This is because similarity is (in general) intransitive. As a canonical and clear example, think of the relation of approximate equality $\approx$ between numbers. Specifically, consider the relation defined as follows: $n \approx n'$ iff $|n - n'| < \frac{1}{2}$. It is easy to see that $\approx$ is intransitive. Here's a counter-example: $\frac{1}{4} \approx \frac{1}{2}$ and $\frac{1}{2} \approx \frac{3}{4}$, but $\frac{1}{4} \not\approx \frac{3}{4}$. This, in a nutshell, is why similarity relations are not transitive, and why accessibility relations based on similarity lead to modalities that violate (4).

Lewis' counterpart theory of *de re* modality is based on an accessibility relation (the counterpart relation) defined in terms of *similarity*. This results in an intransitive accessibility relation, and, as a result, a theory of *de re* modality that violates the *de re* analogue of axiom (4). For Lewis, x is necessarily P iff all of x's counterparts are P. But, for Lewis, being a counterpart of x involves being *similar to* x (in some sense of similarity). As a result, x' can be a counterpart of x, and x'' a counterpart of x', without x'' being a counterpart of x. So, x can be necessarily P without x being necessarily necessarily P. That is to say, all of x's counterparts x' can be P without all the x''s counterparts x'' being P. This is made possible by the fact that the counterpart relation is intransitive. If that relation were transitive, then x's counterparts x' being P without the x''s counterparts x'' being P would violate the indiscernibility of identicals, since we would then have some x'' being P and not being P (prove this!). On Kripke's account of *de re* modality, x is necessarily P if x is P-in-w, for all possible worlds w in which x exists. Here, the salient relation of accessibility is $R(w*, w)$ iff x exists in w (where, it is assumed that x exists in $w*$). Intuitively, this relation is transitive, since it is based on *identity* and not mere *similarity* (clarify this!).