If \( \Sigma \) is a modal system, then the soundness theorem establishes that if \( \Sigma \vdash \varphi \), then \( \varphi \) is valid in any class \( C \) of models in which all instances of all formulas in \( \Sigma \) are valid. In particular that means that if \( K \vdash \varphi \) then \( \varphi \) is true in all models; if \( KT \vdash \varphi \) then \( \varphi \) is true in all reflexive models; if \( KD \vdash \varphi \) then \( \varphi \) is true in all serial models, etc.

Completeness is the converse of soundness: that \( K \) is complete means that if a formula \( \varphi \) is valid, \( \vdash \varphi \), for instance. Proving completeness is a lot harder to do than proving soundness. It is useful, first, to consider the contrapositive: \( K \) is complete iff whenever \( \not \vdash \varphi \), there is a countermodel, i.e., a model \( \mathfrak{M} \) such that \( \mathfrak{M} \not\models \varphi \). Equivalently (negating \( \varphi \)), we could prove that whenever \( \not \vdash \neg \varphi \), there is a model of \( \varphi \). In the construction of such a model, we can use information contained in \( \varphi \). When we find models for specific formulas we often do the same: E.g., if we want to find a countermodel to \( p \rightarrow \Box q \), we know that it has to contain a world where \( p \) is true and \( \Box q \) is false. And a world where \( \Box q \) is false means there has to be a world accessible from it where \( q \) is false. And that's all we need to know: which worlds make the propositional variables true, and which worlds are accessible from which worlds.

In the case of proving completeness, however, we don’t have a specific formula \( \varphi \) for which we are constructing a model. We want to establish that a model exists for every \( \varphi \) such that \( \not \vdash \Sigma \neg \varphi \). This is a minimal requirement, since if \( \not \vdash \Sigma \neg \varphi \), by soundness, there is no model for \( \varphi \) (in which \( \Sigma \) is true). Now note that \( \not \vdash \Sigma \neg \varphi \) iff \( \varphi \) is \( \Sigma \)-consistent. (Recall that \( \Sigma \not \vdash \varphi \) and \( \varphi \not \vdash \Sigma \bot \) are equivalent.) So our task is to construct a model for every \( \Sigma \)-consistent formula.

The trick we’ll use is to find a \( \Sigma \)-consistent set of formulas that contains \( \varphi \), but also other formulas which tell us what the world that makes \( \varphi \) true has to look like. Such sets are complete \( \Sigma \)-consistent sets. It’s not enough to construct a model with a single world to make \( \varphi \) true, it will have to contain multiple worlds and an accessibility relation. The complete \( \Sigma \)-consistent set containing \( \varphi \) will also contain other formulas of the form \( \Box \psi \) and \( \Diamond \chi \). In all accessible worlds, \( \psi \) has to be true; in at least one, \( \chi \) has to be true. In order to accomplish this, we’ll simply take all possible complete \( \Sigma \)-consistent sets as the basis for the set of worlds. A tricky part will be to figure out when a complete \( \Sigma \)-consistent set should count as being accessible from another in our model.

We’ll show that in the model so defined, \( \varphi \) is true at a world—which is also a complete \( \Sigma \)-consistent set—iff \( \varphi \) is an element of that set. If \( \varphi \) is \( \Sigma \)-consistent, it will be an element of at least one complete \( \Sigma \)-consistent set (a fact we’ll prove), and so there will be a world where \( \varphi \) is true. So we will have a single model where every \( \Sigma \)-consistent formula \( \varphi \) is true at some world. This single model is the canonical model for \( \Sigma \).
Photo Credits

Bibliography