THE DEDUCTION PARADOX

Draft of November 03, 2024

Matheus Silva

A deduction is an inference that aims for validity and can be either valid or invalid. An invalid deduction can never be valid, because if an inference is valid in one possible world, it must be valid in all. One possible world where an inference is valid implies that there are no worlds where the inference is invalid. Therefore, only valid inferences can truly be deductive.

The only plausible invalid deductive inferences are those mistakenly perceived as possibly valid. Examples include formal fallacies like affirming the consequent or denying the antecedent, which might be confused with valid forms like *modus ponens* or *modus tollens*. However, this implies that invalid deductive inferences are simply mistaken as deductive, which is contradictory. An erroneous intuition of validity is insufficient to label an inference deductive, since deductive inferences are, by definition, possibly valid.

This should give us pause: if the only genuine deductions are the valid ones, then our talk about deduction is an indirect and thoughtless way of referring to validity rather than to an inference type. There is simply no deduction to speak off—only validity. But then again, it's clear that validity is an attribute that some inferences possess, while others do not. This is a paradox.

One solution is to argue that the paradox arises from accepting S5's treatment of the accessibility relation, where all worlds are equally accessible. In S5, validity in one world implies validity in all. By adopting a weaker system like S4, where accessibility is not symmetric, an inference can be valid in some worlds without being valid in all. If not all worlds are fully accessible to each other, validity can be restricted to certain worlds, avoiding the paradox. This aligns with formal implication, which is transitive and reflexive, but not symmetric.

It could be objected that an inference valid in some worlds only because it lacks access to all worlds is not truly valid. It is this lack of access that creates the appearance of validity. If we accept that solution, we could declare an inference valid in one world simply by removing its access to the remaining worlds. This is implausible, as it generates validity too easily. Thus, such an inference would not only be invalid from the start but could never be valid, as required for a genuine deduction. Moreover, proponents of S5 would argue that weaker systems, such as S4, are too weak to capture our modal intuitions about the logic of necessity and possibility.

Perhaps a way out of this paradox is to maintain that the supposed paradigmatic examples of deductions, such as *modus ponens* or hypothetical syllogism, are not inferential forms, but coherence requirements for inferences. The actual inferential forms in such requirements are the conditionals that are misinterpreted as premises[1]. For example, a *modus ponens* should be reinterpreted as follows:

> **Inference:** If *A* is true in $world_n$, *B* is true in $world_n$.
> **Premise:** *A* is true in $world_n$.
> **Conclusion:** *B* is true in $world_n$.

The reference to a given $world_n$ is made to ensure that both premises and conclusion can be about any world. If I accept the inference and the premise turns out to be true, I have to accept the conclusion as a matter of coherence in order to keep my inference commitment. Of course, inferences can be invalid. In some cases, the premise is true and the conclusion is false. What can't be "invalid" is the acceptance of both an inference and its premise accompanied by the denial of its conclusion. But since we are considering a combination of an inference and a premise, any talk about validity would be a category mistake in this case. The intended examples of deductions are coherence requirements. So an invalid deduction would have to be an incoherent set of statements that can be coherent, but this is impossible. Now, let's reconsider the supposed examples of invalid deductive inferences. First, consider affirming the consequent:

> **Inference:** If *A* is true in $world_n$, *B* is true in $world_n$.
> **Premise:** *B* is true in $world_n$.
> **Conclusion:** *A* is true in $world_n$.

This is an incoherent set of statements because the premises and the conclusion are inverted. So no inference is actually made. The only charitable interpretation of this example would require the addition of another inference:

1 The thesis that conditional sentences are inferences instead of premises is presented in *"If-then" as a version of "Implies".* For a more in-depth defense of the notion that deductive patterns are coherence requirements for inferences, see my *Coherence of Inferences*.

Inference: If *A* is true in $world_n$, *B* is true in $world_n$.

Inference: If *B* is true in $world_n$, A is true in $world_n$.

If we unravel both inferences, we get two inferential commitments that are coherent:

Inference: If *A* is true in $world_n$, *B* is true in $world_n$.

Premise: *A* is true in $world_n$.

Conclusion: *B* is true in $world_n$.

Inference: If *B* is true in $world_n$, *A* is true in $world_n$.

Premise: *B* is true in $world_n$.

Conclusion: *A* is true in $world_n$.

The other paradigmatic example of invalid deduction is denying the antecedent, which can be presented as follows:

Inference: If *A* is true in $world_n$, *B* is true in $world_n$.

Premise: *A* is false in $world_n$.

Conclusion: *B* is false in $world_n$.

Once again there is a mismatch between the description of the premise and the conclusion offered in the inference and the subsequent individual descriptions of each one of them. So this is an incoherent set of statements without any inference taking place. The charitable interpretation of denying the antecedent would result in two inferences:

Inference: If *A* is true in $world_n$, *B* is true in $world_n$.

Inference: If *A* is false in $world_n$, *B* is false in $world_n$.

Similarly to affirming the consequent, the premise and the conclusion are perceived to be logically equivalent. The only difference is that both inferences can't take place in the same world, because if the premise and conclusion of the first inference are true, the premise and the conclusion of the latter inference can't be true. I'm afraid a similar fate would await any candidate for invalid deduction.

Since the paradigmatic examples of deductions are actually coherence requirements for inferences, shouldn't we abandon the notion of deduction altogether? The reason why we should keep the concept of deduction is that for an inference to be coherent the conclusion must be necessitated by the premise in some modal range. Otherwise the reasoner could make the inference, accept the premise, but still remain reticent about the conclusion. That the conclusion must be true given the premise is intrinsic to the notion of inference.