PHIL 100G: Logic and Critical Analysis
Chris Dorst

Soundness and Completeness (of a Derivation System)

Our derivation system is the collection of inference rules (like MP, DM, R, etc.) that we have been using, coupled with the *structure* of our derivations themselves (like when and how you can cancel a "Show," and which lines are available from which other lines).

We use our derivation system to show that arguments are valid. But arguments themselves are things that occur in ordinary language, not first and foremost in the symbolic language that we use in our derivation system. And whether an argument in ordinary language is valid or not depends only on the connection between the premises and the conclusion. If it is impossible for all of the premises to be true and for the conclusion to be false, then the argument is valid. Otherwise, if it's *possible* for all of the premises to be true and the conclusion false, the argument is invalid. The key point here is that what *makes* an argument valid or invalid is the relationship between the premises and the conclusion, not whether we can do a derivation of it.

Back to our derivation system. Our derivation system lets us do certain sorts of derivations but not other ones. We could easily do a derivation from the premises ~P and (P v Q) to the conclusion Q. On the other hand, try as we might, we'll never be able to do a derivation from the premises P and Q to the conclusion R. And that's because there's just no collection of inference rules and sub-derivations in our derivation system that will ever get us to that conclusion from those premises.

So the question naturally arises: is our derivation system a good one? In particular, we can sharpen this question in two ways:

1. Given a derivation in our derivation system, is the corresponding argument in ordinary language guaranteed to be valid?
2. Given a valid argument in ordinary language, is there guaranteed to be a derivation of it in our derivation system?

You could imagine a derivation system for which the answer to these questions is "No." For example, you could imagine a derivation system that allowed certain derivations that did not correspond to valid arguments (thus answering "No" to question 1). Suppose, say, that we added a new rule called "Conditional Introduction," which lets us use P to infer (P -> Q). That's clearly a bad rule of inference, because just knowing that the antecedent of a conditional is true doesn't guarantee that the entire conditional is true. But in principle we *could* add this rule of inference to our derivation system. If we did, our derivation system would allow us to do derivations that don't correspond to any valid argument in ordinary language.

Such a derivation system would not be *sound*. So *soundness* (of a derivation system) is what results if the answer to question (1) above is "Yes." In other words, **a derivation system is *sound* if every derivation that it lets us do corresponds to a valid argument**. A sound derivation system is never going to lead us wrong—it's never going to allow us to do a derivation of an invalid argument.[1]

Return to our questions (1) and (2) above. We just imagined a derivation system for which the answer to (1) was "No," and we saw that such a derivation system would be unsound. You could also imagine a derivation system for which the answer to question (2) is "No." It might be, for example, that our derivation system is pretty good and lets us do derivations of lots of valid arguments, but nevertheless there are some valid arguments that we cannot do derivations of. If we took our current derivation system and removed a bunch of its rules of inference, then we would be in such a situation.

Such a derivation system would not be *complete*. So *completeness* (of a derivation system) is what results if the answer to question (2) above is "Yes." In other words, **a derivation system is *complete* if every valid argument can be given a derivation in that system.[2]** A complete derivation system is "good enough" in the sense that, given a valid argument, there's guaranteed to be a derivation of that argument in our system. (Though it may sometimes be very difficult to *find* such a derivation. It may be very long and convoluted, for example.)

It's easy to get soundness without completeness, and it's also easy to get completeness without soundness. For the former, you could imagine a derivation system that only has the rule Repetition. That system is clearly never going to allow you to do a derivation of an *invalid* argument (so it's sound), but it's just as clearly not going to allow you to do derivations of *every valid* argument (so it's not complete).

Conversely, you could imagine a derivation system that has the following rule of inference: "From any premise(s), you may infer any conclusion." That derivation system would be complete, because every valid argument would have a derivation in that system. But that's just because *every* argument, valid or invalid, has a derivation in that system. So the system would not be sound; it would allow us to do derivations of arguments that are not valid.

So soundness by itself comes cheap, and completeness by itself comes cheap. The challenge is to design a derivation system that is *both* sound and complete. Though we won't prove it in this course, our derivation system *is* both sound and complete. (If you take the next course in this sequence [PHIL 301] you'll learn how to prove both soundness and completeness for a variety of different derivation systems.)

---

[1] Make sure to distinguish the soundness *of a derivation system* from the soundness of an *argument*. Remember, an *argument* is sound if it is valid and it has all true premises. This doesn't have much at all to do with the soundness of a derivation system. It's a bit unfortunate that logicians have chosen the same term for both of these distinct notions.

[2] As with soundness, "completeness" has multiple meanings in logic. In particular, make sure to distinguish the completeness *of a derivation system* from *truth-functional completeness*. Truth-functional completeness is another term for what we've been calling *expressive adequacy*, which is a property of the set of connectives in the symbolic language that we're using, not a property of the derivation system itself. Remember, a set of connectives is expressively adequate (i.e. truth-functionally complete) if it can be used to express every possible truth function. This notion of truth-functional completeness doesn't have much to do with the notion of completeness as applied to derivation systems, so try to keep them distinct in your mind.