Proving the soundness of natural deduction

What is to be proved?

To establish soundness, we seek to prove that, if

$\Phi_1, \Phi_2, \ldots, \Phi_n \vdash \Psi$

is a valid sequent, then $\Phi_1, \Phi_2, \ldots, \Phi_n \models \Psi$.

The premise here is that we can use natural deduction to derive Ψ from $\Phi_1, \Phi_2, \ldots, \Phi_n$. The conclusion we wish to draw is that any assignment of truth values to propositional atoms that makes all the formulae $\Phi_1, \Phi_2, \ldots, \Phi_n$ true also makes the formula Ψ true.

A sketch proof

We sketch a proof that works by induction on the number of steps in the deduction used to establish the validity of the given sequent. Call this number k. (Note carefully that the inductive hypothesis asserts that we can assume the above implication holds so long as the number of steps in deducing the validity of a sequent is $< k$. This places no restriction on the *number of premises*, which - when applying the inductive hypothesis to a sequent with n premises - may introduce sequents with more than n premises.)

Base case: The base case is k=1. Because of the nature of proof by natural deduction, there is at least one step, and if there is just 1 this must link premise directly to conclusion, so that in effect one of the formulae $\Phi_1, \Phi_2, \ldots, \Phi_n$ must actually be Ψ. In this case, the conclusion we wish to draw is obvious.

Inductive step: Suppose that there is a proof of validity of the sequent

$\Phi_1, \Phi_2, \ldots, \Phi_n \vdash \Psi$

that requires $k > 1$ steps. This deduction must end with the application of one of the rules of natural deduction. Informally (as stated in H&R), we need to show that "the rules of natural deduction behave semantically in the same way as the corresponding truth tables evaluate". We can make this statement more precise!

Any rule of natural deduction takes the form:

antecedents
----
consequent

In a proof of the validity of the given sequent, the last consequent must be Ψ. Each antecedent to this consequent in the last rule applied is either an explicit formula Ψ_i, or takes the form of a "boxed" sequence of deductions in which there is an assumption α, and associated conclusion Γ. For the original sequent to be valid, the steps in the natural deduction must incorporate proofs of the validity of the sequents:

$\Phi_1, \Phi_2, \ldots, \Phi_n \vdash \Psi_i$

for all the explicit antecedents of the form Ψ_i, and of the sequents

$\Phi_1, \Phi_2, \ldots, \Phi_n, \alpha \vdash \Gamma$

associated with boxed antecedents. The proofs of the validity of these sequents, which correspond merely to segments of the proof of the validity of the original sequent, satisfy the inductive hypothesis. We can conclude that any assignment of truth values to propositional atoms that makes all the formula $\Phi_1, \Phi_2, \ldots, \Phi_n$ true also makes the explicit antecedents Ψ_i in the final deduction step true, and also renders the sequents implicit in the boxed antecedents valid, so that any assignment making α true also renders Γ true. By inspecting each of the individual rules for natural deduction, it is then apparent that any assignment of truth values to propositional atoms that makes all the formulae $\Phi_1, \Phi_2, \ldots, \Phi_n$ true also makes the explicit antecedents Ψ true.

By way of illustration:

- if the last rule is $\wedge$ introduction, and the conclusion is $\Psi = \Psi_1 \wedge \Psi_2$, then the inductive hypothesis establishes that any assignment of truth values to propositional atoms that makes all the formulae $\Phi_1, \Phi_2, \ldots, \Phi_n$ true also makes the explicit antecedents Ψ_1 and Ψ_2 true, so guaranteeing the truth of Ψ.

- if the last rule is $\vee$ introduction, and the conclusion is $\Psi = \Psi_1 \vee \Psi_2$, then the inductive hypothesis establishes that any assignment of truth values to propositional atoms that makes all the formulae $\Phi_1, \Phi_2, \ldots, \Phi_n$ true also makes at least one of the explicit antecedents Ψ_1 or Ψ_2 true, whilst also ensuring that if in addition either Ψ_1 or Ψ_2 is true then so also is Ψ.

Checking the other rules for natural deduction is left as an exercise.

The above proof illustrates *course-of-values* induction, since it is necessary to invoke the inductive hypothesis not merely for sequents whose validity can be established in k-1 steps, but in any number of steps $< k$.
