

8. Natural Deduction for First Order Logic

8.1. Rules of Inference

In the last chapter, we discussed the language of first-order logic, and the rules that govern their use. We summarize them here:

The universal quantifier:

$$\frac{A(x)}{\forall y A(y)} \forall I \qquad \frac{\forall x A(x)}{A(t)} \forall E$$

In the introduction rule, x should not be free in any uncanceled hypothesis. In the elimination rule, t can be any term that does not clash with any of the bound variables in A .

The existential quantifier:

$$\frac{A(t)}{\exists x A(x)} \exists I \qquad \frac{\frac{\frac{\overline{A(y)}^1}{\vdots} B}{\exists x A(x)} 1}{B} \exists E$$

In the introduction rule, t can be any term that does not clash with any of the bound variables in A . In the elimination rule, y should not be free in B or any uncanceled hypothesis.

Equality:

$$\frac{}{t = t} \text{ refl} \qquad \frac{s = t}{t = s} \text{ symm} \qquad \frac{r = s \quad s = t}{r = t} \text{ trans}$$

$$\frac{s = t}{r(s) = r(t)} \text{ subst} \qquad \frac{s = t \quad P(s)}{P(t)} \text{ subst}$$

Strictly speaking, only refl and the second substitution rule are necessary. The others can be derived from them.

8.2. The Universal Quantifier

The following example of a proof in natural deduction shows that if, for every x , $A(x)$ holds, and for every x , $B(x)$ holds, then for every x , they both hold:

$$\frac{\frac{\frac{\overline{\forall x A(x)}^1}{A(y)}}{\forall y (A(y) \wedge B(y))} \frac{\overline{\forall x B(x)}^2}{B(y)}}{\forall x B(x) \rightarrow \forall y (A(y) \wedge B(y))} 2$$

$$\frac{}{\forall x A(x) \rightarrow (\forall x B(x) \rightarrow \forall y (A(y) \wedge B(y)))} 1$$

Notice that neither of the assumptions 1 or 2 mention y , so that y is really “arbitrary” at the point where the universal quantifiers are introduced.

Here is another example:

$$\frac{\frac{\frac{\overline{\forall x A(x)}^1}{A(y)}}{A(y) \vee B(y)}}{\forall x (A(x) \vee B(x))} \quad \frac{\forall x (A(x) \vee B(x))}{\forall x A(x) \rightarrow \forall x (A(x) \vee B(x))}^1$$

As an exercise, try proving the following:

$$\forall x (A(x) \rightarrow B(x)) \rightarrow (\forall x A(x) \rightarrow \forall x B(x)).$$

Here is a more challenging exercise. Suppose I tell you that, in a town, there is a (male) barber that shaves all and only the men who do not shave themselves. You can show that this is a contradiction, arguing informally, as follows:

By the assumption, the barber shaves himself if and only if he does not shave himself. Call this statement (*).

Suppose the barber shaves himself. By (*), this implies that he does not shave himself, a contradiction. So, the barber does not shave himself.

But using (*) again, this implies that the barber shaves himself, which contradicts the fact we just showed, namely, that the barber does not shave himself.

Try to turn this into a formal argument in natural deduction.

Let us return to the example of the natural numbers, to see how deductive notions play out there. Suppose we have defined *even* and *odd* in such a way that we can prove:

- $\forall n (\neg \text{even}(n) \rightarrow \text{odd}(n))$
- $\forall n (\text{odd}(n) \rightarrow \neg \text{even}(n))$

Then we can go on to derive $\forall n (\text{even}(n) \vee \text{odd}(n))$ as follows:

$$\frac{\frac{\overline{\text{even}(n) \vee \neg \text{even}(n)}}{\text{even}(n)} \quad \frac{\overline{\text{even}(n)}^1}{\text{even}(n) \vee \text{odd}(n)} \quad \frac{\frac{\overline{\forall n (\neg \text{even}(n) \rightarrow \text{odd}(n))}}{\neg \text{even}(n) \rightarrow \text{odd}(n)} \quad \frac{\overline{\neg \text{even}(n)}}{\text{odd}(n)}}{\text{even}(n) \vee \text{odd}(n)}}{\text{even}(n) \vee \text{odd}(n)}^1 \quad \frac{\text{even}(n) \vee \text{odd}(n)}{\forall n (\text{even}(n) \vee \text{odd}(n))}$$

We can also prove and $\forall n \neg(\text{even}(n) \wedge \text{odd}(n))$:

$$\frac{\frac{\overline{\text{odd}(n) \rightarrow \neg \text{even}(n)}}{\neg \text{even}(n)} \quad \frac{\overline{\text{even}(n) \wedge \text{odd}(n)}^H}{\text{odd}(n)} \quad \frac{\overline{\text{even}(n) \wedge \text{odd}(n)}^H}{\text{even}(n)}}{\perp} \quad \frac{\perp}{\neg(\text{even}(n) \wedge \text{odd}(n))}^H \quad \frac{\neg(\text{even}(n) \wedge \text{odd}(n))}{\forall n \neg(\text{even}(n) \wedge \text{odd}(n))}$$

As we move from modeling basic rules of inference to modeling actual mathematical proofs, we will tend to shift focus from natural deduction to formal proofs in Lean. Natural deduction has its uses: as a model of logical reasoning, it provides us with a convenient means to study metatheoretic properties such as soundness and completeness. For working *within* the system, however, proof languages like Lean's tend to scale better, and produce more readable proofs.

8.3. The Existential Quantifier

Remember that the intuition behind the elimination rule for the existential quantifier is that if we know $\exists x A(x)$, we can temporarily reason about an arbitrary element y satisfying $A(y)$ in order to prove a conclusion that doesn't depend on y . Here is an example of how it can be used. The next proof says that if we know there is something satisfying both A and B , then we know, in particular, that there is something satisfying A .

$$\frac{\frac{\frac{\overline{\exists x(A(x) \wedge B(x))}^1}{\exists x A(x)} \quad \frac{\frac{\overline{A(y) \wedge B(y)}^2}{A(y)}}{\exists x A(x)}^2}{\exists x(A(x) \wedge B(x)) \rightarrow \exists x A(x)}^1$$

The following proof shows that if there is something satisfying either A or B , then either there is something satisfying A , or there is something satisfying B .

$$\frac{\frac{\overline{\exists x(A(x) \vee B(x))}^1}{\exists x A(x) \vee \exists x B(x)}^2 \quad \frac{\frac{\overline{A(y) \vee B(y)}^2}{\exists x A(x) \vee \exists x B(x)}^3 \quad \frac{\frac{\overline{B(y)}}{\exists x B(x)}^3}{\exists x A(x) \vee \exists x B(x)}^3}{\exists x(A(x) \vee B(x)) \rightarrow \exists x A(x) \vee \exists x B(x)}^1$$

The following example is more involved:

$$\frac{\frac{\frac{\overline{\forall x(A(x) \rightarrow \neg B(x))}^1}{A(x) \rightarrow \neg B(x)} \quad \frac{\frac{\overline{A(x) \wedge B(x)}^3}{A(x)} \quad \frac{\overline{A(x) \wedge B(x)}^3}{B(x)}}{\neg B(x)}^3}{\frac{\overline{\exists x(A(x) \wedge B(x))}^2}{\perp}}^3 \quad \frac{\perp}{\neg \exists x(A(x) \wedge B(x))}^2}{\forall x(A(x) \rightarrow \neg B(x)) \rightarrow \neg \exists x(A(x) \wedge B(x))}^1$$

In this proof, the existential elimination rule (the line labeled 3) is used to cancel two hypotheses at the same time. Note that when this rule is applied, the hypothesis $\forall x(A(x) \rightarrow \neg B(x))$ has not yet been canceled. So we have to make sure that this formula doesn't contain the variable x freely. But this is o.k., since this hypothesis contains x only as a bound variable.

Another example is that if x does not occur in P , then $\exists x P$ is equivalent to P :

$$\frac{\frac{\overline{\exists x P}^1}{P}^2 \quad \frac{\overline{P}^2}{\exists x P}^1}{\exists x P \leftrightarrow P}^1$$

This is short but tricky, so let us go through it carefully. On the left, we assume $\exists x P$ to conclude P . We assume P , and now we can immediately cancel this assumption by existential elimination, since x does not occur in P , so it doesn't occur freely in any assumption or in the conclusion. On the right we use existential introduction to conclude $\exists x P$ from P .

8.4. Equality

Recall the natural deduction rules for equality:

$$\frac{}{t = t} \quad \frac{s = t}{t = s} \quad \frac{r = s \quad s = t}{r = t} \\ \frac{s = t}{r(s) = r(t)} \quad \frac{s = t \quad P(s)}{P(t)}$$

Keep in mind that we have implicitly fixed some first-order language, and r , s , and t are any terms in that language. Recall also that we have adopted the practice of using functional notation with terms. For example, if we think of $r(x)$ as the term $(x + y) \times (z + 0)$ in the language of arithmetic, then $r(0)$ is the term $(0 + y) \times (z + 0)$ and $r(u + v)$ is $((u + v) + y) \times (z + 0)$. So one example of the first inference on the second line is this:

$$\frac{u + v = 0}{((u + v) + y) \times (z + 0) = (0 + y) \times (z + 0)}$$

The second axiom on that line is similar, except now $P(x)$ stands for any *formula*, as in the following inference:

$$\frac{u + v = 0 \quad x + (u + v) < y}{x + 0 < y}$$

Notice that we have written the reflexivity axiom, $t = t$, as a rule with no premises. If you use it in a proof, it does not count as a hypothesis; it is built into the logic.

In fact, we can think of the first inference on the second line as a special case of the first. Consider, for example, the formula $((u + v) + y) \times (z + 0) = (x + y) \times (z + 0)$. If we plug $u + v$ in for x , we get an instance of reflexivity. If we plug in 0, we get the conclusion of the first example above. The following is therefore a derivation of the first inference, using only reflexivity and the second substitution rule above:

$$\frac{u + v = 0 \quad ((u + v) + y) \times (z + 0) = ((u + v) + y) \times (z + 0)}{((u + v) + y) \times (z + 0) = (0 + y) \times (z + 0)}$$

Roughly speaking, we are replacing the second instance of $u + v$ in an instance of reflexivity with 0 to get the conclusion we want.

Equality rules let us carry out calculations in symbolic logic. This typically amounts to using the equality rules we have already discussed, together with a list of general identities. For example, the following identities hold for any real numbers x , y , and z :

- commutativity of addition: $x + y = y + x$
- associativity of addition: $(x + y) + z = x + (y + z)$
- additive identity: $x + 0 = 0 + x = x$
- additive inverse: $-x + x = x + -x = 0$
- multiplicative identity: $x \cdot 1 = 1 \cdot x = x$
- commutativity of multiplication: $x \cdot y = y \cdot x$
- associativity of multiplication: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
- distributivity: $x \cdot (y + z) = x \cdot y + x \cdot z$, $(x + y) \cdot z = x \cdot z + y \cdot z$

You should imagine that there are implicit universal quantifiers in front of each statement, asserting that the statement holds for *any* values of x , y , and z . Note that x , y , and z can, in particular, be integers or rational numbers as well. Calculations involving real numbers, rational numbers, or integers generally involve identities like this.

The strategy is to use the elimination rule for the universal quantifier to instantiate general identities, use symmetry, if necessary, to orient an equation in the right direction, and then using the substitution rule for equality to change something in a previous result. For example, here is a natural deduction proof of a simple identity, $\forall x, y, z ((x + y) + z = (x + z) + y)$, using only commutativity and associativity of addition. We have taken the liberty of using a brief name to denote the relevant identities, and combining multiple instances of the universal quantifier introduction and elimination rules into a single step.

$$\begin{array}{c}
 \frac{\frac{\text{assoc}}{(x + z) + y = x + (z + y)}}{x + (z + y) = (x + z) + y} \quad \frac{\text{comm}}{y + z = z + y} \quad \frac{\text{assoc}}{(x + y) + z = x + (y + z)} \\
 \hline
 (x + y) + z = (x + z) + y \\
 \hline
 \forall x, y, z ((x + y) + z = (x + z) + y)
 \end{array}$$

There is generally nothing interesting to be learned from carrying out such calculations in natural deduction, but you should try one or two examples to get the hang of it, and then take pleasure in knowing that it is possible.

8.5. Counterexamples and Relativized Quantifiers

Consider the statement:

Every prime number is odd.

In first-order logic, we could formulate this as $\forall p (prime(p) \rightarrow odd(p))$. This statement is false, because there is a prime number which is even, namely the number 2. This is called a *counterexample* to the statement.

More generally, given a formula $\forall x A(x)$, a counterexample is a value t such that $\neg A(t)$ holds. Such a counterexample shows that the original formula is false, because we have the following equivalence: $\neg \forall x A(x) \leftrightarrow \exists x \neg A(x)$. So if we find a value t such that $\neg A(t)$ holds, then by the existential introduction rule we can conclude that $\exists x \neg A(x)$, and then by the above equivalence we have $\neg \forall x A(x)$. Here is a proof of the equivalence:

$$\begin{array}{c}
 \frac{\frac{\neg \forall x A(x)}{1} \quad \frac{\frac{\frac{\neg(\exists x \neg A(x))}{4} \quad \frac{\neg A(x)}{5}}{\exists x \neg A(x)}}{\perp} \quad \frac{\frac{\perp}{5} \quad A(x)}{\forall x A(x)} \\
 \hline
 \frac{\perp}{\exists x \neg A(x)} \quad 4 \quad \neg \forall x A(x) \leftrightarrow \exists x \neg A(x)
 \end{array}
 \qquad
 \begin{array}{c}
 \frac{\frac{\exists x \neg A(x)}{1} \quad \frac{\neg A(y)}{3}}{\perp} \\
 \hline
 \frac{\perp}{\neg \forall x A(x)} \quad 2 \quad 1
 \end{array}$$

One remark about the proof: at the step marked by 4 we *cannot* use the existential introduction rule, because at that point our only assumption is $\neg \forall x A(x)$, and from that assumption we cannot prove $\neg A(t)$ for a particular term t . So we use a proof by contradiction there.

As an exercise, prove the “dual” equivalence yourself: $\neg \exists x A(x) \leftrightarrow \forall x \neg A(x)$. This can be done without using proof by contradiction.

In [Chapter 7](#) we saw examples of how to use relativization to restrict the scope of a universal quantifier. Suppose we want to say “every prime number is greater than 1”. In first order logic this can be written as $\forall n(\text{prime}(n) \rightarrow n > 1)$. The reason is that the original statement is equivalent to the statement “for every natural number, if it is prime, then it is greater than 1”. Similarly, suppose we want to say “there exists a prime number greater than 100.” This is equivalent to saying “there exists a natural number which is prime and greater than 100,” which can be expressed as $\exists n (\text{prime}(n) \wedge n > 100)$.

As an exercise you can prove the above results about negations of quantifiers also for relativized quantifiers. Specifically, prove the following statements:

- $\neg \exists x (A(x) \wedge B(x)) \leftrightarrow \forall x (A(x) \rightarrow \neg B(x))$
- $\neg \forall x (A(x) \rightarrow B(x)) \leftrightarrow \exists x (A(x) \wedge \neg B(x))$

For reference, here is a list of valid sentences involving quantifiers:

- $\forall x A \leftrightarrow A$ if x is not free in A
- $\exists x A \leftrightarrow A$ if x is not free in A
- $\forall x (A(x) \wedge B(x)) \leftrightarrow \forall x A(x) \wedge \forall x B(x)$
- $\exists x (A(x) \wedge B) \leftrightarrow \exists x A(x) \wedge B$ if x is not free in B
- $\exists x (A(x) \vee B(x)) \leftrightarrow \exists x A(x) \vee \exists x B(x)$
- $\forall x (A(x) \vee B) \leftrightarrow \forall x A(x) \vee B$ if x is not free in B
- $\forall x (A(x) \rightarrow B) \leftrightarrow (\exists x A(x) \rightarrow B)$ if x is not free in B
- $\exists x (A(x) \rightarrow B) \leftrightarrow (\forall x A(x) \rightarrow B)$ if x is not free in B
- $\forall x (A \rightarrow B(x)) \leftrightarrow (A \rightarrow \forall x B(x))$ if x is not free in A
- $\exists x (A(x) \rightarrow B) \leftrightarrow (A(x) \rightarrow \exists x B)$ if x is not free in B
- $\exists x A(x) \leftrightarrow \neg \forall x \neg A(x)$
- $\forall x A(x) \leftrightarrow \neg \exists x \neg A(x)$
- $\neg \exists x A(x) \leftrightarrow \forall x \neg A(x)$
- $\neg \forall x A(x) \leftrightarrow \exists x \neg A(x)$

All of these can be derived in natural deduction. The last two allow us to push negations inwards, so we can continue to put first-order formulas in negation normal form. Other rules allow us to bring quantifiers to the front of any formula, though, in general, there will be multiple ways of doing this. For example, the formula

$$\forall x A(x) \rightarrow \exists y \forall z B(y, z)$$

is equivalent to both

$$\exists x, y \forall z (A(x) \rightarrow B(y, z))$$

and

$$\exists y \forall z \exists x (A(x) \rightarrow B(y, z)).$$

A formula with all the quantifiers in front is said to be in *prenex* form.

8.6. Exercises

1. Give a natural deduction proof of

$$\forall x (A(x) \rightarrow B(x)) \rightarrow (\forall x A(x) \rightarrow \forall x B(x)).$$

2. Give a natural deduction proof of $\forall x B(x)$ from hypotheses $\forall x (A(x) \vee B(x))$ and $\forall y \neg A(y)$.
3. From hypotheses $\forall x (\text{even}(x) \vee \text{odd}(x))$ and $\forall x (\text{odd}(x) \rightarrow \text{even}(s(x)))$ give a natural deduction proof $\forall x (\text{even}(x) \vee \text{even}(s(x)))$. (It might help to think of $s(x)$ as the function defined by $s(x) = x + 1$.)
4. Give a natural deduction proof of $\exists x A(x) \vee \exists x B(x) \rightarrow \exists x (A(x) \vee B(x))$.

5. Give a natural deduction proof of $\exists x (A(x) \wedge C(x))$ from the assumptions $\exists x (A(x) \wedge B(x))$ and $\forall x (A(x) \wedge B(x) \rightarrow C(x))$.
6. Prove some of the other equivalences in the last section.
7. Consider some of the various ways of expressing “nobody trusts a politician” in first-order logic:
 - $\forall x (\text{politician}(x) \rightarrow \forall y (\neg \text{trusts}(y, x)))$
 - $\forall x, y (\text{politician}(x) \rightarrow \neg \text{trusts}(y, x))$
 - $\neg \exists x, y (\text{politician}(x) \wedge \text{trusts}(y, x))$
 - $\forall x, y (\text{trusts}(y, x) \rightarrow \neg \text{politician}(x))$

They are all logically equivalent. Show this for the second and the fourth, by giving natural deduction proofs of each from the other. (As a shortcut, in the \forall introduction and elimination rules, you can introduce / eliminate both variables in one step.)

8. Formalize the following statements, and give a natural deduction proof in which the first three statements appear as (uncancelled) hypotheses, and the last line is the conclusion:
 - Every young and healthy person likes baseball.
 - Every active person is healthy.
 - Someone is young and active.
 - Therefore, someone likes baseball.

Use $Y(x)$ for “is young,” $H(x)$ for “is healthy,” $A(x)$ for “is active,” and $B(x)$ for “likes baseball.”

9. Give a natural deduction proof of $\forall x, y, z (x = z \rightarrow (y = z \rightarrow x = y))$ using the equality rules in [Section 8.4](#).
10. Give a natural deduction proof of $\forall x, y (x = y \rightarrow y = x)$ using only these two hypotheses (and none of the new equality rules):

- $\forall x (x = x)$
- $\forall u, v, w (u = w \rightarrow (v = w \rightarrow u = v))$

(Hint: Choose instantiations of u , v , and w carefully. You can instantiate all the universal quantifiers in one step, as on the last homework assignment.)

11. Give a natural deduction proof of $\neg \exists x (A(x) \wedge B(x)) \leftrightarrow \forall x (A(x) \rightarrow \neg B(x))$
12. Give a natural deduction proof of $\neg \forall x (A(x) \rightarrow B(x)) \leftrightarrow \exists x (A(x) \wedge \neg B(x))$
13. Remember that both the following express $\exists! x A(x)$, that is, the statement that there is a unique x satisfying $A(x)$:
 - $\exists x (A(x) \wedge \forall y (A(y) \rightarrow y = x))$
 - $\exists x A(x) \wedge \forall y \forall y' (A(y) \wedge A(y') \rightarrow y = y')$

Do the following:

- Give a natural deduction proof of the second, assuming the first as a hypothesis.
- Give a natural deduction proof of the first, assuming the second as a hypothesis.

(Warning: these are long.)