Chapter 4

A Tiny Bit of Informal Logic

We have come somewhat proficient in using informal logic in our arguments about aspects of discrete mathematics.

Although we have used quantifiers, $\exists$ and $\forall$ we did so mostly viewing them as symbolic abbreviations of English texts about mathematics. In this chapter we will expand our techniques in logic, extending them to include manipulation of quantifiers including the versatile Induction —or mathematical induction— technique used to prove properties of the natural numbers.

We know how to detect fallacious statements formulated in Boolean logic: Simply show by a truth table that the statement is not a tautology. (talk about tautological implication too)

We will show in the domain of quantifiers not only how to prove statements that include quantifiers but also how to disprove false statements that happen to include quantifiers.

4.1. Enriching our proofs to manipulate quantifiers

Manipulation of quantifiers boils down to "how can I remove a quantifier from the beginning of a formula?" and "how can I add a quantifier at the beginning of a formula?" Once we learn this technique we will be able to reason within mathematics with ease.

But first let us define once and for all what a mathematical proof *looks like*: its *correct*, *expected syntax* or *form*.

We will need some concepts to begin with.

1. The alphabet and structure of formulas. Formulas are strings. The alpha-

bet of *symbols* that we use contain, *at a minimum*,

$$=, \neg, \wedge, \vee, \rightarrow, \equiv, (,), \forall, \exists, \text{object variables}^{\dagger}$$

We finitely generate the infinite set of object variables using single letters, if necessary with primes and/or subscripts: $A, x, y'', w'''_{23}, u_{501}$.

2. One normally works in a mathematical area of interest, or *mathematical theory* —such as Geometry, Set Theory, Number Theory, Algebra, Calculus— where one needs additional symbols to write down formulas, like

$$0, \emptyset, \in, \subset, \int, \circ, +, \times$$

and many others.

3. Mathematicians as a rule get to recognise the *formulas* and *terms* in the math areas of their interest without being necessarily taught the recursive definition of the syntax of these. We will not give the syntax in these notes either (but see [Tou08] if you want to know!). Thus one learns to be content with getting to know formulas and terms by their behaviour and through use, rather than by their exact definition of syntax.

   - *Terms* are "function calls", in the jargon of the computer savvy person. These calls take math objects as inputs and return math objects as outputs. *Examples* are: $x + y$, $x \times 3$, $0 \times x + 1$ (one is told that $\times$ is stronger than $+$, so, notwithstanding the bracket-parsimonious notation "$0 \times x + 1$", we know it means "$(0 \times x) + 1$", so this call returns 1, no matter what we plugged into x).
   - *Formulas* are also function calls, but their output is *restricted* (by their syntax that I will not define carefully!) to be one or the other of the truth values <u>true</u> or <u>false</u> (**t** or **f**) but nothing else! Their input, just as in the case for terms, is any math object. *Examples* are: $2 < 3$ (**t**), $(\forall x)x = x$ (**t**), $(\forall x)x = 0$ (**f**), $(\exists x)x = 0$ (**t**), $x = 0$ neither true nor false; answer depends on the input in x!

     More: $x = x$ (**t**) answer is independent of input; $x = 0 \rightarrow x = 0$ (**t**) answer is independent of input; $x = 0 \rightarrow (\forall x)x = 0$ neither true nor false; answer depends on input in x! The input variable is the *leftmost* x; the other two are *bound* and *unavailable* to accept inputs. See below.
   - If an **occurrence** of formula variable *is* available for input it would normally be called an occurrence as an *input variable*. Rather, such occurrences are called *free occurrences* in the literature.

     At the expense of writing style, "occurrence" occurred no less than four times in the short passage above. The aim is *emphasis*: It is not a *variable* x that is free or bound in a formula, but it is the occurrences

[†] That is, variables that denote *objects* such as numbers, arrays, matrices, sets, trees, etc.

of said variable that we are speaking of, as the immediately preceding example makes clear.

4. In $(\forall x)x = 0$ the variable x is non input, it is "bound" we say. Just like this: $\Sigma_{i=1}^{4} i$, which means $1 + 2 + 3 + 4$ and "i" is not available for input: Something like $\Sigma_{3=1}^{4} 3$ is nonsense! Similar comment for $\exists$.

5. We call $\forall, \exists, \neg, \wedge, \vee, \rightarrow, \equiv$ the "*logical connectives*", the last 5 of them being called *Boolean connectives*. People avoid cluttering notation with a lot of brackets by agreeing that the first 3 have the same "strength" or "priority"; the *highest*. The remaining connectives have priorities decreasing as we walk to the right.

   Thus, if A and B are (denote) formulas, then $\neg A \vee B$ means $(\neg A) \vee B$; $\neg$ wins the claim for A. If we want $(\forall x)$ to apply to the entire $A \rightarrow B$ we must write $(\forall x)(A \rightarrow B)$.

   What about $A \rightarrow B \rightarrow C$ and $A \equiv B \equiv C$? Brackets are *implied from right to left*: $A \rightarrow (B \rightarrow C)$ and $A \equiv (B \equiv C)$. And this? $(\exists y)(\forall x)\neg A$. Brackets are *implied, again, from right to left*: $\Big((\exists y)\big((\forall x)(\neg A)\big)\Big)$.

   BTW, the part where a $\forall x$ or $\exists x$ *acts* —the "$(\ldots)$" in $(\forall x)(\ldots)$ and $(\exists x)(\ldots)$— is called their scope.

6. **Boolean deconstruction**. A formula like $(\forall x)A \rightarrow B$ can be *deconstructed* Boolean-wise into $(\forall x)A$ and B. If I knew more about B —say, it is $x = 3 \rightarrow x = 7$, then I can deconstruct further.

   So, now I have got

   $$(\forall x)A, \quad x = 3, \quad x = 7$$

   The last two have NO Boolean structure so deconstructing stops with them. How about $(\forall x)A$? This cannot be deconstructed either, even if A had Boolean structure! Such structure is locked up in the scope of $(\forall x)$.

   We call the formulas where deconstruction stops "*prime*". A prime formula is one with no Boolean structure, e.g., $x < 8$, or one of the form $(\forall x)A$ (A is the scope) or $(\exists x)A$ (A is the scope).

   Every formula is either prime or can be deconstructed into prime components.

4.1.1 Remark. (Tautologies) A formula A is a *tautology* iff it **is true due to its Boolean structure**, according to truth tables (2.3.4) no matter what the values of its prime formulas into which it is deconstructed are assumed to be. **Assumed to be**: We do **NOT** compute the *intrinsic* truth value of a prime formula when we check whether A is a tautology or not.

For example, $x = x$ is a prime formula and thus its assumed value could be ANY of **t** or **f**. Thus it is NOT a tautology, even though, *intrinsically IS true*, no matter what the value of x. □

4.1.2 Example.

1. $(\forall x)A$ is not a tautology as it has two possible truth values (being a prime formula) in principle.
2. $x = 0 \to x = 0$ is a tautology. Which are its prime (sub) formulas? □
3. $(\forall x)x = 0 \to x = 0$ is not a tautology. I repeat (**once**): To determine tautologyhood we *DO NOT evaluate prime formulas*; we just consider *each* of the two scenarios, **t** or **f**, for each prime formula and use truth tables to compute the overall truth value.

If we DID evaluate $(\forall x)x = 0$ we would see that (say over the natural numbers, or reals, or complex numbers) it is false.† So the implication is true! But we DON'T do that! **Not true** as a Boolean formula!

So, how do we show that $(\forall x)A$ is true (if it is)? Well, in easy cases we try to see if A is true for all values of x. That failing, we will use a proof (see 4.1.9).

Similarly for $(\exists x)A$. To show it is true (if it is) we try to see if A is true for some value of x. Often we just guess one such value that works, say c, and verify the truth of A when $x = c$. That failing, we will use a proof.

4.1.3 Definition. (Tautological implication)

We say that the formulas $A_1, A_2, \ldots, A_n$ *tautologically imply* a formula B —in symbols $A_1, A_2, \ldots, A_n \models_{taut} B$— meaning

"the truth of $A_1 \wedge A_2 \wedge \ldots \wedge A_n$ implies the truth of B"

that is, that

$A_1 \wedge A_2 \wedge \ldots \wedge A_n \to B$ is a tautology

□

4.1.4 Remark. Note that we do NOT care to *check*, or even *state*, what happens if $A_1 \wedge A_2 \wedge \ldots \wedge A_n$ is false.

The implication in blue type is true regardless of the truth value of B

So, a tautological implication $A_1, A_2, \ldots, A_n \models_{taut} B$ says that B is true provided we proved (or accepted) that the lhs of $\models_{taut}$ is true.

$\models_{taut}$ propagates truth from left to right. □

4.1.5 Example. Here are some easy and some involved tautological implications. They can all be verified using truth tables, either building the tables in full, or taking shortcuts.

†If we are doing our mathematics restricted to the set {0}, then, in this "theory" the formula IS true!

1. $A \models_{taut} A$
2. $A \models_{taut} A \vee B$
3. $A \models_{taut} B \to A$
4. $A, \neg A \models_{taut} B$ —any B. Because I do *work* only if $A \wedge \neg A$ is true! See above.
5. $\mathbf{f} \models_{taut} B$ —any B. Because I do *work* only if lhs is true! See above.
6. Is this a valid tautological implication? $B, A \to B \models_{taut} A$, where A and B are distinct.

   No, for if A is false and B is true, then the lhs is true, but the rhs is false!
7. Is this a valid tautological implication? $A, A \to B \models_{taut} B$? Yes! Say $A = \mathbf{t}$ and $(A \to B) = \mathbf{t}$. Then, from the truth table of $\to$, it *must* be $B = \mathbf{t}$.
8. How about this? $A, A \equiv B \models_{taut} B$? Yes! Verify!
9. How about this? $A \vee B \equiv B \models_{taut} A \to B$? Yes! I verify:

   First off, **assume** lhs of $\models_{taut}$ —that is, $A \vee B \equiv B$— is true.

   Two cases:

   - $B = \mathbf{f}$. Then I need the lhs of $\equiv$ to be true to satisfy the bolded "assume". So $A = \mathbf{f}$ as well and clearly the rhs of $\models_{taut}$ is true with these values.
   - $B = \mathbf{t}$. Then I need not worry about A on the lhs. The rhs of $\models_{taut}$ is true by truth table of $\to$.
10. $A \wedge (\mathbf{f} \equiv A) \models_{taut} B$, for any B. Well, just note that the lhs of $\models_{taut}$ is $\mathbf{f}$ so we need to do no work with B to conclude that the implication is valid.
11. $$A \to B, C \to B \models_{taut} A \vee C \to B$$

    This is nicknamed "proof by cases" for the obvious reasons. Verify this tautological implication! □

The job of a mathematical proof is to start from established (previous theorems) truths, or assumed truths (axioms) and unfailingly preserve truths in all its steps as it is developed. Thus, it will have produced, in particular, a truth at its very last step. A theorem.

What are our axioms, our starting assumptions, when we do proofs?

4.1.6 Definition. First off, in *any proof that we will write in math* there are axioms that are independent of the type of math that we do, whether it is set theory, number theory, algebra, calculus, etc.

Our logical axioms are

1. All tautologies; these need no defence as "start-up truths".

2. Formulas of the form $(\forall x)A[x] \to A[t]$, for any formula A, variable x and "object" t.

   This object can be as simple as a variable y (might be same as x), constant c, or as complex as a "*function call*", $f(t_1, t_2, \ldots, f_n)$ where f accepts n-inputs, and the inputs shown here are already available objects.

   Two comments: This is a *bona fide* start-up *truth* as its says "if $A[x]$ is true for all x-values,[†] then it is true also if we plug a specific value/object into x".

   The other comment is that I write $A[x]$ to indicate *a variable of interest*. This may or may not occur in A, which may also have other variables that it depends on. I would write $A(x, y, z)$ —round brackets— if I knew that x, y, z are *all* the variables on which A depends.

3. Formulas of the form $A[x] \to (\forall x)A[x]$, **for any formula A where the variable x does not occur in it.** For example say A is $3 = 3$. This axiom says then, "if $3 = 3$ is true, then so is $(\forall x)3 = 3$". Sure! $3 = 3$ does not depend on x. So saying "for all values of x we have $3 = 3$" is the same as saying just "we have $3 = 3$".

4. Formulas of the form $A[t] \to (\exists x)A[x]$, for any formula A, variable x and "object t. This is a good start-up *truth*: It says that if we know that some object plugged into x makes $A[x]$ true, then it is correct to say "there is some value x that makes $A[x]$ true —in symbols $(\exists x)A[x]$."

5. $x = x$ is the *identity* axiom, no matter what "x" I use to express it. So, $y = y$ and $w = w$ are also instances of the axiom.

6. $x = y \to y = x$ and $x = y \wedge y = z \to x = z$ are the *equality* axioms.

   They can be expressed equally well using variables other than x and y (e.g., u, v and w).

7. The $\exists$ vs. $\forall$ axiom. For any formula A, $(\exists x)A[x] \equiv \neg(\forall x)\neg A[x]$ is an axiom. □

The "rules of proving", or rules of inference. These are two up in front —you will find I am grossly miscounting:

4.1.7 Definition. (Rules)

1. From $A[x]$ I may infer $(\forall x)A[x]$. Logicians write the up-in-front ("primary") rules as fractions without words:

$$\frac{A[x]}{(\forall x)A[x]}$$

   this rule we call *generalisation*, or Gen for short.

[†]People usually say "for all x", meaning for all **values** of x.

2. I may *construct* (and use) using any tautological implication *that I verified*, say, this one
$$A_1, A_2, \ldots, A_n \models_{taut} B$$
the rule
$$\frac{A_1, A_2, \ldots, A_n}{B}$$
Seeing readily that $A, A \to B \models_{taut} B$, we have the rule
$$\frac{A, A \to B}{B}$$
This is a very popular rule, known as *modus ponens*, for short MP.

□

1. **HOW** do you *use* rules? See Definition 4.1.9 below. If in a proof you are writing you have reached the numerator of a rule, then *it is correct* to write next (or later) the denominator of the rule. We say that you applied the rule.

2. The second "rule" above is a rule constructor. Any tautological implication we come up with is fair game: It leads to a *valid rule* since the name of the game (in a proof) is *preservation/propagation of truth.*

   This is NOT an invitation to learn and memorise infinitely many rules (!) but is rather a license to build your own rules as you go, *as long as you bothered to verify* the validity of the tautological implication they are derived from.

3. Gen is a rule that indeed propagates truth: If $A[x]$ is true, that *means* that it is so for all values of x and all values of any other variables on which A depends but I did not show in the $[\ldots]$ notation. But then so is $(\forall x)A[x]$ true, as it says precisely the same thing: "$A[x]$ is true, for all values of x and all values of any other variables on which A depends but I did not show in the $[\ldots]$ notation".

   The only difference between the two notations is that I added some notational *emphasis* in the second —$(\forall x)$.

   For example, if I know that B has just two variables, u and v, I can write it as $B(u, v)$. Then
$$B(u, v) \text{ } \mathbf{t} \text{ iff } (\forall u)B(u, v) \text{ } \mathbf{t} \text{ iff } (\forall v)B(u, v) \text{ } \mathbf{t} \text{ iff } (\forall u)(\forall v)B(u, v) \text{ } \mathbf{t}$$

4. Hmm. So is $\forall x$ redundant? Yes, but *only as a formula prefix.* In something like this
$$x = 0 \to (\forall x)x = 0 \tag{1}$$

it is NOT redundant!

Dropping $\forall$ we change the meaning of (1).

As is, (1) is *not* a true statement. For example, for $x = 0$ it is false. However dropping $\forall x$, (1) changes to $x = 0 \to x = 0$ which is a tautology; always true.

5. The axioms in 4.1.6 are indispensable to do just logic; that is why we call them *logical axioms*.

   You also use them in *all* math reasoning no matter what type of math it is. However, the latter has its own ***additional*** axioms! These are called *special*, but most often "*mathematical axioms*".

   We are not going to list them. Why? Because every math branch, or "theory" as we say, has different axioms!

4.1.8 Example. Here is a *sample* of axioms from *math* (*theories*):

1. Number theory for $\mathbb{N}$:
   - $x < y \vee x = y \vee x > y$ (*trichotomy*)
   - $\neg x < 0$ this axiom indicates that 0 is *minimal* in $\mathbb{N}$. Adding the previous one makes $<$ a total order, so 0 is also *minimum*.
   - Many others that we omit.
2. Euclidean geometry:
   - From two distinct points passes one and only one line.
   - ("Axiom of parallels") From a point A off a line named k —both A and k being on the same plane— passes a unique line on said plane that is parallel to k.
   - Many others that we omit.
3. Axiomatic set theory:
   - For any set A,

     $$(\exists y) y \in A \to (\exists x)\Big(x \in A \wedge \neg(\exists z)(z \in x \wedge z \in A)\Big)$$

     This is the axiom of "foundation" from which one can prove things like $A \in A$ is always *false*.
     It says that *IF* there is *any* element in A *at all* —this is the hypothesis part "$(\exists y) y \in A$"— *THEN* there is some element —this is the part "$(\exists x)\Big(x \in A$"— *below which*, if you follow "$\in$" backwards from it, you will *not* find a z ("$\neg(\exists z)$") that is *both* below x along $\in$ backwards, *and* also a member of A —this part is "$(z \in x \wedge z \in A)$".

4. And a few others that we omit. □

So what is the *shape* of proofs?

4.1.9 Definition. (Proofs and theorems) A proof is a finite sequence of formulas

$$F_1, F_2, \ldots, F_i, \ldots, F_n \tag{1}$$

such that, for *each* $i = 1, 2, \ldots, n$, F_i is obtained as ONE of:

1. It is an axiom from among the ones we listed in 4.1.6.
2. It is an axiom of the theory (area of Math) that we are working in.
3. It is a PREVIOUSLY proved theorem.
4. It is the result of "Gen" applied to a previous formula F_j. That is, $F_i = (\forall x)F_j$, for some x and $j < i$.
5. It is the result of "$\models_{taut}$" applied to previous formulas F_{j_k}, $k = 1, 2, \ldots, m$. That is, $F_{j_1}, F_{j_2}, F_{j_3}, \ldots, F_{j_m} \models_{taut} F_i$, and all j_r for $r = 1, 2, \ldots, m$ are $< i$.

Such proofs are known as "Hilbert-style proofs". We write them vertically, ONE formula per line, every formula consecutively numbered, with annotation to the right of formulas (the "why did I write this?"). Like this

1)	F_1	⟨because⟩
2)	F_2	⟨because⟩
⋮	⋮	⋮
n)	F_n	⟨because⟩

Every F_n in (1) is called a theorem. Thus we define

A theorem is a formula that **appears** in a proof.

Often one writes $\vdash A$ to symbolically say that A is a theorem. If we must indicate that we worked in some specific theory, say ZFC (set theory), then we may indicate this as

$$\vdash_{ZFC} A$$

If moreover we have had some "*non-axiom* assumptions" (read on to see when this happens!) that form a set Σ, then we may indicate so by writing

$$\Sigma \vdash_{ZFC} A$$

□

Why Σ for a set of (*non-axiom*) assumptions? Because we reserve upper case latin letters for formulas. For *sets* of formulas we use a *distinguishable* capital letter, so, we chose distinguishable Greek capital letters, such as $\Gamma, \Sigma, \Delta, \Phi, \Theta, \Psi, \Omega$. Obviously, Greek capital letters like A, B, E, Z will not do!

4.1.10 Example. (New (derived) rules) A derived rule is one we were not given —in 4.1.7— to bootstrap logic, but we can still prove they propagate truth.

1. We have a new (derived) rule: $(\forall x)A[x] \vdash A[t]$.

   This is called *Specialisation*, or *Spec*.

   Aha! We used a non-axiom assumption here! I write a Hilbert proof to show that $A[t]$ is a theorem if $(\forall x)A[x]$ is a (non-axiom) hypothesis (assumption) —shortened to "hyp".

1)	$(\forall x)A[x]$	$\langle$hyp$\rangle$
2)	$(\forall x)A[x] \to A[t]$	$\langle$axiom$\rangle$
3)	$A[t]$	$\langle 1 + 2 + \text{MP}\rangle$

2. Taking t to be x we have $(\forall x)A[x] \vdash A[x]$, simply written as $(\forall x)A \vdash A$.

3. The Dual Spec derived rule: $A[t] \vdash (\exists x)A[x]$. We prove it:

1)	$A[t]$	$\langle$hyp$\rangle$
2)	$A[t] \to (\exists x)A[x]$	$\langle$axiom$\rangle$
3)	$(\exists x)A[x]$	$\langle 1 + 2 + \text{MP}\rangle$

Taking t to be x we have $A[x] \vdash (\exists x)A[x]$, simply written as $A \vdash (\exists x)A$. □

There are two principles of proof that we state without proving them (see [Tou08] if curious).

4.1.11 Remark. (Deduction theorem and proof by contradiction)

1. The *deduction theorem* (also known as "proof by assuming the antecedent") states, if
   $$\Gamma, A \vdash B \tag{1}$$
   then also $\Gamma \vdash A \to B$, provided that in the proof of (1), all free variables of A were treated as constants: That is we neither used them to do a Gen, nor substituted objects into them.

   The notation "Γ, A" is standard for the more cumbersome $\Gamma \cup \{A\}$.

   In practice, this principle is applied to prove $\Gamma \vdash A \to B$, by doing instead the "easier" (1). Why easier? We are helped by an extra hypothesis, A, and the formula to prove, B, is less complex than $A \to B$.

2. Proof by contradiction. To prove $\Gamma \vdash A$ is equivalent to proving the "constant formula" $\mathbf{f}$ from hypothesis $\Gamma, \neg A$.

3. Why the burden of the non-axiom hypotheses Γ? Because in applying the deduction theorem we usually start with a task like "do $\vdash A \to B \to C \to D$".

   So we go like this:

- By DThm, it suffices to prove $A \vdash B \to C \to D$ instead (here "Γ" was $\emptyset$).
- Again, by DThm, it suffices to prove $A, B \vdash C \to D$ instead (here "Γ" was A).
- Again, by DThm, it suffices to prove $A, B, C \vdash D$ instead (here "Γ" was A, B).

□ 

I referred you to [Tou08] for some things. However, the short intro here adopted the so-called "strong generalisation", which has the side-effect of making the deduction theorem to hedge: In proving B from Γ, A one *must* ensure that no variable of A was subject to generalisation or substitution. [Tou08] trades some power of generalisation in order to get an easier to apply deduction theorem, with no hedging.

So this is a choice on what we want to be "easy", and what we want to "not be so easy". There are two options!

4.1.12 Remark. (Ping-Pong) For any formulas A and B, the formula — where I am using way more brackets than I have to, ironically, to *improve* readability—

$$(A \equiv B) \equiv \Big((A \to B) \wedge (B \to A)\Big)$$

is a tautology (draw up a truth table with one row for each of the possible values of A and B and verify that the equivalence is always $\mathbf{t}$).

Thus to prove the lhs of the $\equiv$ suffices to prove the rhs:

	$\vdots$	$\vdots$
1)	$(A \to B) \wedge (B \to A)$	⟨suppose I proved this⟩
2)	$(A \equiv B) \equiv \Big((A \to B) \wedge (B \to A)\Big)$	⟨axiom⟩
3)	$A \equiv B$	⟨1 + 2 + tautological implication⟩

In turn, to prove the rhs it suffices to prove each of $A \to B$ and $B \to A$ separately. This last idea encapsulates the ping-pong approach to proving equivalences.

Here are a few applications. □

4.1.13 Example. 1. Establish $\vdash (\forall x)(A \wedge B) \equiv (\forall x)A \wedge (\forall x)B$.

By ping-pong.

- Prove $\vdash (\forall x)(A \wedge B) \to (\forall x)A \wedge (\forall x)B$. By DThm suffices to do $(\forall x)(A \wedge B) \vdash (\forall x)A \wedge (\forall x)B$ instead.

1)	$(\forall x)(A \wedge B)$	⟨hyp⟩
2)	$A \wedge B$	⟨1 + Spec⟩

3)	A	⟨2 + tautological implication⟩
4)	B	⟨2 + tautological implication⟩
5)	$(\forall x)A$	⟨3 + Gen; OK: x is not free in line 1⟩
6)	$(\forall x)B$	⟨4 + Gen; OK: x is not free in line 1⟩
7)	$(\forall x)A \wedge (\forall x)B$	⟨5 + 6 + tautological implication⟩

- Prove $\vdash (\forall x)A \wedge (\forall x)B \rightarrow (\forall x)(A \wedge B)$. By DThm suffices to do $(\forall x)A \wedge (\forall x)B \vdash (\forall x)(A \wedge B)$ instead.

1)	$(\forall x)A \wedge (\forall x)B$	⟨hyp⟩
2)	$(\forall x)A$	⟨1 + tautological implication⟩
3)	$(\forall x)B$	⟨1 + tautological implication⟩

  Complete the above proof!

2. Prove $\vdash (\forall x)(\forall y)A \equiv (\forall y)(\forall x)A$. By ping-pong.

- Prove $\vdash (\forall x)(\forall y)A \rightarrow (\forall y)(\forall x)A$.
  By DThm suffices to do $(\forall x)(\forall y)A \vdash (\forall y)(\forall x)A$ instead.

1)	$(\forall x)(\forall y)A$	⟨hyp⟩
2)	$(\forall y)A$	⟨1 + Spec⟩
3)	A	⟨2 + Spec⟩
4)	$(\forall x)A$	⟨3 + Gen; OK, no free x in line 1⟩
5)	$(\forall y)(\forall x)A$	⟨4 + Gen; OK, no free y in line 1⟩

- Prove $\vdash (\forall y)(\forall x)A \rightarrow (\forall x)(\forall y)A$.
  Exercise! □

We have seen how to *add* an $(\exists x)$ in front of a formula (4.1.10 3.).

How about *removing* an $(\exists x)$-prefix? This is much more complex than removing a $(\forall x)$-prefix:

The technique can be *proved* to be correct (eg., [Tou03a]) but I will omit the proof here as I did omit the proof of the deduction theorem technique and the proof by contradiction technique. I could say "see [Tou03a] if you want to learn the proof", but this reference is too advanced for a first year course on discrete math. So, why not look at [Tou08]? These two books have chosen *incompatible* "generalisation" rules, which results to *incompatible* deduction theorem versions.

The proof of the technique of eliminating $\exists$-prefixes *relies on the deduction theorem.*

Technique of removing an $\exists$-prefix: Suppose I have that $(\exists x)A[x]$ is true — either as an **assumption** or a **theorem I proved earlier**— and I want to prove B.

Then I **assume** that —for *some* constant c that does not occur in B— $A[c]$ is true.

That is, I **add** $A[c]$ for an unknown c *NOT* in B as a non-axiom hypothesis.

People annotate this step in a proof as "aux. hyp. caused by $(\exists x)A[x]$."

Now proceed to prove B using all that is known to you —that is, the axioms of the theory $\mathcal{T}$ that you work in, perhaps some non-axiom hypotheses Γ, and $(\exists x)A[x]$, *and the non-axiom hypothesis* $A[c]$.

Do so by using all free (input-) variables of $A[c]$ as constants in your proof![†]

The *technique of removing an* $\exists$*-prefix* guarantees that you did better than

$$\Gamma, (\exists x)A[x], \boxed{A[c]} \vdash_{\mathcal{T}} B$$

that actually you achieved

$$\Gamma, (\exists x)A[x] \vdash_{\mathcal{T}} B$$

as if you never assumed nor used $A[c]$!

That is why they call it "auxiliary hypothesis". Once it helps you prove B it drops out; it does not stay around to get credit!

4.1.14 Example. Prove $\vdash (\exists y)(\forall x)A[x,y] \to (\forall x)(\exists y)A[x,y]$.

By the DThm it suffices to prove $(\exists y)(\forall x)A[x,y] \vdash (\forall x)(\exists y)A[x,y]$ instead.

1)	$(\exists y)(\forall x)A[x,y]$	⟨hyp⟩
2)	$(\forall x)A[x,c]$	⟨aux. hyp. caused by 1; for some constant c not in the conclusion⟩
3)	$A[x,c]$	⟨2 + Spec⟩
4)	$(\exists y)A[x,y]$	⟨3 + Dual Spec⟩
5)	$(\forall x)(\exists y)A[x,y]$	⟨4 + Gen; OK, no free x in lines 1 and 2⟩

□

4.1.15 Example. Can I also prove the converse of the above? That is $\vdash (\forall x)(\exists y)A[x,y] \to (\exists y)(\forall x)A[x,y]$.

I will try.

By the DThm it suffices to prove $(\forall x)(\exists y)A[x,y] \vdash (\exists y)(\forall x)A[x,y]$ instead.

1)	$(\forall x)(\exists y)A[x,y]$	⟨hyp⟩
2)	$(\exists y)A[x,y]$	⟨1 + spec⟩
3)	$A[x,c]$	⟨aux. hyp. for 2; c not in the conclusion⟩
4)	$(\forall x)A[x,c]$	⟨3 + Gen; Hmmm! Illegal: I should treat the free x of aux. hyp. as a constant!⟩

Still, can anyone PROVE this even if I cannot?

[†]This is a side-effect of using the deduction theorem in the proof of correctness of the technique.

A question like this, if you are to answer "NO", must be resolved by offering a *counterexample*. That is, a special case of A for which I can clearly see that the claim is not true.

Here is one such:

$$(\forall x)(\exists y)x = y \to (\exists y)(\forall x)x = y \tag{1}$$

Say we work in $\mathbb{N}$. The lhs of $\to$ is true, but the rhs is false as it claims that there is a number such that *all* numbers are equal to it. □

Another useful principle that can be proved, but we will not do so, is that one can *replace equivalents-by-equivalents*. That is, if C is some formula, and if I have

1. $A \equiv B$, via proof, or via assumption, and also
2. A is a subformula of C

then I can *replace* one (or more) occurrence(s) of A in C (as subformula(s)) by B and call the resulting formula C', and be guaranteed the conclusion $C \equiv C'$. That is, from $A \equiv B$, I can prove $C \equiv C'$.

This principle is called the *equivalence theorem*.

Let's do a couple of ad hoc additional examples before we move to the section on Induction.

4.1.16 Example. $A \to B \vdash (\forall x)A \to (\forall x)B$.

By the DThm it suffices to prove $A \to B, (\forall x)A \vdash (\forall x)B$ instead.

1)	$A \to B$	⟨hyp⟩
2)	$(\forall x)A$	⟨hyp⟩
3)	A	⟨2 + Spec⟩
4)	B	⟨1 + 3 + MP⟩
5)	$(\forall x)B$	⟨4 + Gen; OK as the DThm hyp. (line 2) has no free x⟩

□

4.1.17 Example. Refer to 4.1.6(7). Let us apply it to $\neg A$ for arbitrary A. We get

$$\vdash (\exists x)\neg A \equiv \neg(\forall x)\neg\neg A \tag{1}$$

Pause. Why "$\vdash$"?◀

Since $A \equiv \neg\neg A$ is a tautology, hence a theorem

Pause. Why "hence a theorem"?◀

we apply the equivalence theorem above and tautological implication to obtain from (1):

$$\vdash (\exists x)\neg A \equiv \neg(\forall x)A \tag{2}$$

Applying another tautological implication to (2) we get

$$\vdash (\forall x)A \equiv \neg(\exists x)\neg A$$

which is of the same form as 4.1.6(7) with the roles of $\exists$ and $\forall$ reversed. □

4.1.18 Example. $A \equiv B \vdash (\forall x)A \equiv (\forall x)B$.

True due to the equivalence theorem! "C" is "$(\forall x)A$". We replaced (one occurrence of) A by B in C, and we have assumed as starting point that $A \equiv B$. □

4.1.19 Exercise. Prove $A \equiv B \vdash (\forall x)A \equiv (\forall x)B$ without relying on the equivalence theorem. Rather use 4.1.16 in your proof, remembering the ping-pong tautology (4.1.12). □

Bibliography

[Dav65] M. Davis, *The undecidable*, Raven Press, Hewlett, NY, 1965.

[Hin78] P. G. Hinman, *Recursion-theoretic hierarchies*, Springer-Verlag, New York, 1978.

[Kle43] S.C. Kleene, *Recursive predicates and quantifiers*, Transactions of the Amer. Math. Soc. **53** (1943), 41–73, [Also in [Dav65], 255–287].

[Kur63] A.G. Kurosh, *Lectures on General Algebra*, Chelsea Publishing Company, New York, 1963.

[Tou03a] G. Tourlakis, *Lectures in Logic and Set Theory, Volume 1: Mathematical Logic*, Cambridge University Press, Cambridge, 2003.

[Tou03b] ———, *Lectures in Logic and Set Theory, Volume 2: Set Theory*, Cambridge University Press, Cambridge, 2003.

[Tou08] ———, *Mathematical Logic*, John Wiley & Sons, Hoboken, NJ, 2008.