Deductive Logic

Kenny Easwaran

December 6, 2018

We have already discussed the relation $\models$, which is a semantic relation between a set of sentences Γ and a sentence ϕ, holding iff there is no interpretation where every sentence in Γ is true and ϕ is false. However, there is no obvious way to figure out when this relation holds between a given Γ and ϕ (except in special cases), since the set of possible interpretations for a language is not just infinite, but non-enumerable. (Consider the language of arithmetic with a successor function and a relation symbol "$<$" — even if we fix the standard domain of the natural numbers, and the standard interpretation of the successor function, there are uncountably many sets of natural numbers that could bear the "$<$" relation to any given number.) Thus, it would be nice to find a coextensive relationship that can be checked by looking at something finite. Additionally, this semantic relationship directly connects some premises and a conclusion, while much ordinary reasoning passes through various intermediate steps on the way.

A logical deduction system is a set of syntactic procedures for defining a relationship $\vdash$ iff there is a "derivation" of ϕ using only assumptions in Γ. In this case we will say that ϕ is a "theorem" of Γ. This relation will be a purely syntactic one — that is, it can be defined entirely in terms of the symbols that appear in various sentences, without reference to any particular interpretation of the symbols.

A note on terminology: Many textbooks use the word "proof" for these formal derivations, but I will reserve the word "proof" for an ordinary language explanation of why something must be true, and will always use the word "derivation" for a formal manipulation of symbols of this relevant sort, in order to try to keep the two concepts distinct. Formal derivations are intended to represent informal proofs, but they have some very important differences, starting with the fact that formal derivations have a very precise set of rules regarding what steps are allowed. I will use the lower case word "theorem" for the result of a formal derivation, and the lower case word "meta-theorem" for the result of an informal proof about what theorems are derivable within a given formal system. But in keeping with standard informal terminology, in the capitalized names of some meta-theorems, I will just use the word "Theorem".

There are many different ways to define a deduction system. In this handout, I will discuss two common systems, called Natural Deduction and Axiomatic Deduction. The Natural Deduction system gives us a relation $\vdash_{ND}$ and the axiomatic system gives us a relation $\vdash_A$, and at the end we will show that $\Gamma \vdash_{ND} \phi$ iff $\Gamma \vdash_A \phi$. Thus, we will eventually be justified in ignoring the distinction between the two relations and saying $\Gamma \vdash \phi$.

The goal is to make sure that the deduction system is in fact coextensive with the semantic consequence relation. That is, we have already said that $\Gamma \models \phi$ iff there is no interpretation of the language that makes all sentences of Γ true while making ϕ false. Our eventual goal is to show that we have defined $\vdash$ in such a way that $\Gamma \models \phi$ iff $\Gamma \vdash \phi$. The claim that if $\Gamma \vdash \phi$ then $\Gamma \models \phi$ is called the Soundness Theorem — it shows that the deduction system is "sound" in the sense that it only derives things that are actually consequences of the premises. The converse claim, that if $\Gamma \models \phi$ then $\Gamma \vdash \phi$, is called the Completeness Theorem — it shows that the deduction system is "complete" in the sense that it actually allows you to derive all consequences of the premises. This is not to be confused with another sense of "completeness" from Gödel's Incompleteness Theorem — there may still be a sentence ϕ such that neither ϕ nor $\neg\phi$ is a consequence of Γ.

The significance of this result will be that the semantic relationship $\models$ can be expressed in a purely syntactic way. Instead of having to check every interpretation of a language to see if a sentence is a consequence of some premises, one can just construct a derivation.

There are a few points of notation to keep in mind. Capital Greek letters like "Γ" will stand for a *set* of sentences, and not just one sentence, while lowercase Greek letters like "ϕ" and "ψ" will stand for individual sentences. Furthermore, although the cases we have talked about so far usually only use finite sets Γ, we will also consider *infinite* sets of sentences. Furthermore, just as with the semantic relationship $\models$, there will be certain abbreviations used with $\vdash$. Instead of writing $\Gamma \cup \{\phi\} \vdash \psi$, we will often write $\Gamma, \phi \vdash \psi$. And instead of writing $\emptyset \vdash \psi$ we will often write $\vdash \psi$.

1 Natural Deduction

In a "natural deduction" system, a derivation consists of a finite chart like the following:

1		
2	$\quad A \wedge B$	
3	$\quad A$	$\wedge$ Elim, 2
4	$\quad B$	$\wedge$ Elim, 2
5	$\quad B \wedge A$	$\wedge$ Intro, 4,3
6	$(A \wedge B) \to (B \wedge A)$	$\to$ Intro, 2-5

Each line in this chart represents a step in the reasoning. Some lines (the ones above the short horizontal dashes) represent assumptions in a derivation (or subderivation), while the others represent inferences. The vertical lines represent the scope of a given assumption — each sentence in the derivation depends on the assumptions made at the top of each vertical line that is still running at that point, but doesn't depend on any assumptions made in subderivations that have since been closed. A new subderivation may be started at any time, with any assumption whatsoever introduced at its beginning. A subderivation may also be ended at any time. But for the steps that represent actual inferences, a rule must be cited to justify the inference, and one must also cite the particular earlier steps that justify the relevant case of that rule.

In the example above, the overall derivation does not depend on any assumptions, but there is a subderivation that depends on one assumption. There are three distinct rules of inference that are applied (one of them twice, on lines 3 and 4). I will explain these rules later on — for now, this is just an example to show the general look of a derivation in a natural deduction system.

We can then use this notion of a derivation and subderivations to define the relation $\vdash_{ND}$.

> $\Gamma \vdash_{ND} \phi$ iff there is a derivation that includes ϕ as a sentence on some line, and Γ contains all the assumptions of the main derivation and any subderivations that are still running at that line.

Thus, the above derivation gives us $\{(A \wedge B), C\} \vdash_{ND} B$ by considering line 4 (since $(A \wedge B)$ is a member of $\{(A \wedge B), C\}$), and $\vdash_{ND} (A \wedge B) \rightarrow (B \wedge A)$ by considering line 6 (since there are no assumptions needed at that point). The vertical lines are visual guides that help you find out what assumptions each sentence in the derivation depends on.

Each rule in a natural deduction system gives some conditions under which one can create a new line in such a deduction. These conditions that can justify the application of a particular rule always depend on certain things having occurred earlier in the derivation. However, the basic rule is that what goes on inside a subderivation stays inside the subderivation — things that occur inside a subderivation can't be used to justify an inference that is made outside that subderivation. Officially, this is just an arbitrary rule, but informally, this is because things inside a subderivation depend on additional assumptions that are no longer in effect when one is outside the subderivation. (This informal explanation tells us that the arbitrary rules of the derivation system were set up in a particular way because we want this syntactic derivation system to behave like the semantic notion of validity, where these sorts of dependencies matter.)

Some rules require that certain *sentences* have occurred earlier (outside the scope of any subderivation that has since closed), but other rules require that a *subderivation* of a particular form must have occurred before. Although this subderivation must itself have ended by the time it is used, the background assumptions under which the subderivation was carried out must all still be in effect — the other vertical lines that are running alongside the subderivation must all still be running at the step where the rule appealing to the existence of this subderivation can be used.

Importantly, the only conditions on the application of a rule of inference depend on what came earlier in the derivation, and there are no conditions on the introduction of a new assumption and subderivation, so truncating a natural deduction derivation at any point also results in a valid natural deduction derivation.

1.1 Simplifying derivations

At this point, even before we have stated any of the rules, we can already say something about the $\vdash_{ND}$ relation. (You may find it helpful to skip ahead to the actual rules and some example derivations though.) We have defined it in terms of any line within a valid derivation, that has any number of subderivations still running. However, we should pay attention to the restrictions mentioned above on when a previous statement in the derivation can be used for one of the rules. Using these restrictions, we can show:

> $\Gamma \vdash_{ND} \phi$ iff there is a derivation where ϕ is the very last step in the derivation, there are no subderivations still running at the end, and Γ contains all the assumptions of the main derivation.

To see this, first recall that we can truncate the derivation at the step on which ϕ is stated. So now start with any derivation that allows us to establish $\Gamma \vdash_{ND} \phi$, and assume that there are n subderivations still running when ϕ appears, on the very last step. (That is, there are $n+1$ vertical lines in front of ϕ on the final step.) We will show that regardless of what n is, there is another derivation that allows us to establish $\Gamma \vdash_{ND} \phi$, in which there are no subderivations still in effect — that is, ϕ occurs with just one vertical line in front of it, and there are no assumptions at the top of this new derivation other than the ones that occurred at the top of the original derivation, or at the top of subderivations that were still in effect at the end.

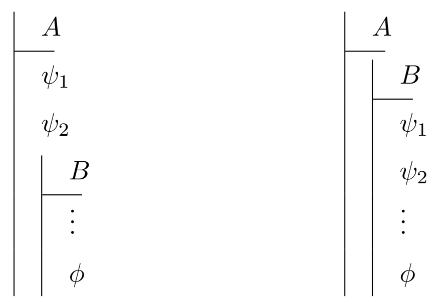
This proof will be by induction on n. We will show that this holds for any derivation where $n = 0$, and then show that if it holds for all derivations where there are n subderivations still running when ϕ appears, then it also holds for all derivations where there are $n + 1$ subderivations still running when ϕ appears.

The case where $n = 0$ is trivial — the original derivation has no subderivations still running when ϕ appears, so it is already the desired final derivation.

So we must just show the induction step. So assume that for any derivation P where there are n subderivations still running when ϕ appears, there is another derivation P' that only starts with assumptions that occurred at the beginning of P or at the beginning of subderivations that are still in effect at the end of P, in which there are no subderivations still running when ϕ appears.

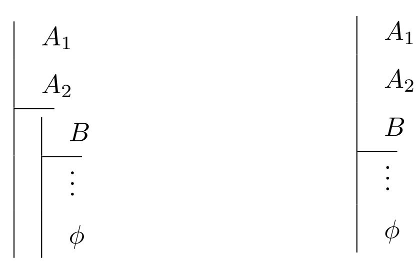
Now consider a particular derivation Q in which ϕ appears on the last line, with $n+1$ subderivations still running at the end. We will show that there is another derivation Q' (with the same assumptions) that has only n subderivations still running at the end, when ϕ still occurs. Then the induction assumption will show that there is yet another derivation Q'' that has no subderivations still running when ϕ appears, and has only the same assumptions as those involved in Q (or Q') and any subderivations that were still running at the end.

So now we must show that if there is a derivation that allows us to establish $\Gamma \vdash_{ND} \phi$ in which $n+1$ subderivations are still running at the point that ϕ appears, and this is the last step, then there is also such a derivation in which only n subderivations are still running. So take the outermost subderivation that is still running at the end. If there are steps before this subderivation opens, then all of those steps can be moved into the subderivation:



Anything that depended on those steps must either already have been within the subderivation, or must be moved into the subderivation along with those, and therefore can still access them. And anything that these steps depended on must still be earlier than them, and hasn't been moved into the scope of a different subderivation that can't be accessed.

Once there are no steps in the derivation before this subderivation begins, we can just add the assumptions of this subderivation to the initial assumptions, and compress the vertical line for this subderivation into the vertical line for the derivation as a whole:



All applications of rules in the derivation will have been within the scope of this subderivation, and will thus still be able to access all statements that they could access before. Thus, we can decrease the number of subderivations running by one, which allows us to complete the proof.

Thus, we can now say that $\Gamma \vdash_{ND} \phi$ iff there is a valid derivation, all of whose assumptions are in Γ, that ends with ϕ, and has no subderivations still running at the end. We have just proven that this is equivalent to the earlier definition of $\vdash_{ND}$, without relying on any features of the particular natural deduction rules, other than the fact that a step is justified entirely in terms of what came earlier.

1.2 The rules

The rules for a natural deduction system traditionally come in pairs — each logical symbol of the language has an introduction rule and an elimination rule. The only exceptions to this are the Double-Negation rule and the Reiteration rule, which stand by themselves. In the descriptions of these rules I have used schematic letters ϕ and ψ to stand for any sentences in the language. The rules all say that if sentences or subderivations of some syntactic form have occurred at an accessible location earlier in the derivation, then a new sentence of some related syntactic form may occur on the present line.

In all of the statements of these rules, recall the *rule of accessibility*:

> A sentence or subderivation that is cited in the use of one of these rules must not be inside any subderivation that has ended. Equivalently, all subderivations that were running when a cited statement occurred must still be running at the time the statement is cited.

1.2.1 Reiteration rule

If ϕ occurs on line n, then we can put ϕ on a later line by citing "Reit n".

1.2.2 Conjunction rules

If ϕ occurs on line m and ψ occurs on line n, then we can put $(\phi \wedge \psi)$ on a later line by citing "$\wedge$ Intro, m, n".

If $(\phi \wedge \psi)$ occurs on line m, then we can either put ϕ or ψ on a later line by citing "$\wedge$ Elim, m".

1.2.3 Disjunction rules

If either ϕ or ψ occurs on line m, then we can put $(\phi \vee \psi)$ on a later line by citing "$\vee$ Intro, m".

If $(\phi \vee \psi)$ occurs on line i, and there is a subderivation whose assumption is ϕ that ends with χ on lines j to k, and there is a subderivation whose assumption is ψ that ends with χ on lines m to n, then we can put χ on a later line by citing "$\vee$ Elim, i, j–k, m–n".

1.2.4 Conditional rules

If there is a subderivation whose assumption is ϕ that ends with ψ on lines m to n, then we can put $(\phi \to \psi)$ on a later line by citing "$\to$ Intro, m–n".

If ϕ occurs on line m, and $(\phi \to \psi)$ occurs on line n, then we can put ψ on a later line by citing "$\to$ Elim, m, n".

1.2.5 Negation rules

If there is a subderivation whose assumption is ϕ that ends with $\perp$ on lines m to n, then we can put $\neg\phi$ on a later line by citing "$\neg$ Intro, m–n".

If ϕ occurs on line m and $\neg\phi$ occurs on line n, then we can put $\perp$ on a later line by citing "$\neg$ Elim, m, n".

If $\neg\neg\phi$ occurs on line m, then we can put ϕ on a later line by citing "Double-Negation, m".

Note that these rules include a "sentence" $\perp$. We can either treat this as an abbreviation for any contradiction (like $((0 < 0) \land \neg(0 < 0)))$, or we can treat it as a special sentence of the language that is false on any interpretation.

Some other authors refer to the $\neg$ elimination rule as $\perp$ introduction, and refer to the double-negation rule as $\neg$ elimination. However, that naming convention ruins a certain pattern of "harmony" between elimination rules and their associated introduction rules. (The relevant notion of harmony is in fact a technical term discussed by Dag Prawitz and Michael Dummett, in arguing for Intuitionist logic on the basis of double-negation being a rule that doesn't harmonize with any introduction rule.)

1.2.6 Quantifier rule definitions

In the statement of the existential and universal rules, v will be any variable, $\phi(v)$ will be a formula with no free variable other than v, t will be any term containing no variables, and $\phi(t)$ will be the sentence that results from replacing all free instances of v in $\phi(v)$ by t. Note that v doesn't actually have to appear in $\phi(v)$ (if it doesn't, then $\phi(v)$ and $\phi(t)$ are identical), just that no other variable can occur free.

c will be any constant that doesn't appear earlier in the derivation (on any accessible line), or in $\phi(v)$, and $\phi(c)$ will be the formula that results from replacing all instances of v in $\phi(v)$ by c. These restrictions on c are somewhat tricky — there are no such restrictions on t, which is allowed to occur in $\phi(v)$ or earlier in the derivation.

1.2.7 Universal rules

If there is a subderivation with no assumption, that ends with $\phi(c)$, on lines m to n, then we can put $\forall v\phi(v)$ on a later line by citing "$\forall$ Intro, m–n".

If $\forall v\phi(v)$ occurs on line n, then we can put $\phi(t)$ on a later line by citing "$\forall$ Elim, n".

1.2.8 Existential rules

If $\phi(t)$ appears on line n, then we can put $\exists v\phi(v)$ on a later line by citing "$\exists$ Intro, n".

If $\exists v \phi(v)$ occurs on line i, and there is a subderivation whose assumption is $\phi(c)$ that ends with ψ on lines j to k, and c doesn't occur in ψ, then we can put ψ on a later line by citing "$\exists$ Elim, i, j–k".

1.2.9 Identity rules

On any line, we can put $t = t$ by citing "$=$ Intro".

If $t_1 = t_2$ occurs on line m, and $\phi(t_1)$ occurs on line n, then we can put $\phi(t_2)$ on a later line by citing "$=$ Elim, m, n".

1.3 Sample derivations

These first three derivations are each somewhat tricky. In the first and third derivations, v doesn't actually occur in $\phi(v)$, and so the constant c never appears in $\phi(c)$ either. In the second derivation, in the definition of $=$ Elim, we let t_1 be c, t_2 be d, and $\phi(v)$ be $v = c$. This is a case where one of the terms does appear in $\phi(v)$.

1	$(0'' < 0''')$	
2	$\quad$	
3	$\quad (0'' < 0''')$	Reit 1
4	$\forall z(0'' < 0''')$	$\forall$ Intro 2-3

1	$c = d$	
2	$c = c$	$=$ Intro
3	$d = c$	$=$ Elim 1,2

1	$(0'' < 0''')$	
2	$\exists y(0'' < 0''')$	$\exists$ Intro 1

The rest of these derivations are more straightforward, if longer, but they illustrate how more of the rules work.

1	$\exists x(((x \cdot x) = 0'') \vee ((x + x) = 0'))$	
2	$\quad (((c \cdot c) = 0'') \vee ((c + c) = 0'))$	
3	$\quad\quad ((c \cdot c) = 0'')$	
4	$\quad\quad \exists y((y \cdot y) = 0'')$	$\exists$ Intro 3
5	$\quad\quad (\exists y((y \cdot y) = 0'') \vee \exists z((z + z) = 0'))$	$\vee$ Intro 4
6	$\quad\quad ((c + c) = 0')$	
7	$\quad\quad \exists z((z + z) = 0')$	$\exists$ Intro 6
8	$\quad\quad (\exists y((y \cdot y) = 0'') \vee \exists z((z + z) = 0'))$	$\vee$ Intro 7
9	$\quad (\exists y((y \cdot y) = 0'') \vee \exists z((z + z) = 0'))$	$\vee$ Elim 2, 3-5, 6-8
10	$(\exists y((y \cdot y) = 0'') \vee \exists z((z + z) = 0'))$	$\exists$ Elim 1, 2-9

1		
2	$\neg((0 < 0) \lor \neg(0 < 0))$	
3	$(0 < 0)$	
4	$((0 < 0) \lor \neg(0 < 0))$	$\lor$ Intro 3
5	$\bot$	$\neg$ Elim 4, 2
6	$\neg(0 < 0)$	$\neg$ Intro 3-5
7	$((0 < 0) \lor \neg(0 < 0))$	$\lor$ Intro 6
8	$\bot$	$\neg$ Elim 7, 2
9	$\neg\neg((0 < 0) \lor \neg(0 < 0))$	$\neg$ Intro 2-8
10	$((0 < 0) \lor \neg(0 < 0))$	Double-Negation 9

1	$\forall x \forall y \forall z(((x < y) \land (y < z)) \to (x < z))$	
2	$\forall x \neg(x < x)$	
3		
4		
5	$(c < d)$	
6	$(d < c)$	
7	$\forall y \forall z(((c < y) \land (y < z)) \to (c < z))$	$\forall$ Elim 1
8	$\forall z(((c < d) \land (d < z)) \to (c < z))$	$\forall$ Elim 7
9	$(((c < d) \land (d < c)) \to (c < c))$	$\forall$ Elim 8
10	$((c < d) \land (d < c))$	$\land$ Intro 5, 6
11	$(c < c)$	$\to$ Elim 10, 9
12	$\neg(c < c)$	$\forall$ Elim 2
13	$\bot$	$\neg$ Elim 11, 12
14	$\neg(d < c)$	$\neg$ Intro 6-13
15	$(c < d) \to \neg(d < c)$	$\to$ Intro 5-14
16	$\forall y((c < y) \to \neg(y < c))$	$\forall$ Intro 4-15
17	$\forall x \forall y((x < y) \to \neg(y < x))$	$\forall$ Intro 3-16

1.4 Exercises

1. Show that

$$\forall x(x < x'), \forall x \forall y \forall z(((x < y) \wedge (y < z)) \rightarrow (x < z)) \vdash_{ND} (0 < 0'')$$

2. Show that

$$\forall x(x < x'), \forall x \forall y \forall z(((x < y) \wedge (y < z)) \rightarrow (x < z)) \vdash_{ND} \forall x(x < x'')$$

3. Show that

$$\exists x((0' + x'') = 0') \vdash_{ND} \exists y \exists z((y' + z') = 0')$$

1.5 Summary of the rules

It will be suggestive soon to represent much of the content of these rules in different notation. If a rule depends on some particular sentences having occurred earlier, then that means that given those sentences as assumptions, one can derive the conclusion in one step. If a rule depends on a particular subderivation occurring earlier, then the rule states that if one derivation is possible, then another is as well. Thus, the rules have the following consequences:

- Reiteration: $\phi \vdash_{ND} \phi$.
- $\wedge$ Intro: $\phi, \psi \vdash_{ND} (\phi \wedge \psi)$.
- $\wedge$ Elim: $(\phi \wedge \psi) \vdash_{ND} \phi$, and $(\phi \wedge \psi) \vdash_{ND} \psi$.
- $\vee$ Intro: $\phi \vdash_{ND} (\phi \vee \psi)$, and $\psi \vdash_{ND} (\phi \vee \psi)$.
- $\vee$ Elim: If $\phi \vdash_{ND} \chi$ and $\psi \vdash_{ND} \chi$, then $(\phi \vee \psi) \vdash_{ND} \chi$.
- $\rightarrow$ Intro: If $\phi \vdash_{ND} \psi$ then $\vdash_{ND} (\phi \rightarrow \psi)$.
- $\rightarrow$ Elim: $\phi, (\phi \rightarrow \psi) \vdash_{ND} \psi$.
- $\neg$ Intro: If $\phi \vdash_{ND} \perp$ then $\vdash_{ND} \neg\phi$.
- $\neg$ Elim: $\phi, \neg\phi \vdash_{ND} \perp$.
- Double Negation: $\neg\neg\phi \vdash_{ND} \phi$.
- $\forall$ Intro: If $\vdash_{ND} \phi(c)$ then $\vdash_{ND} \forall v\phi(v)$.
- $\forall$ Elim: $\forall v\phi(v) \vdash_{ND} \phi(t)$.
- $\exists$ Intro: $\phi(t) \vdash_{ND} \exists v\phi(v)$.
- $\exists$ Elim: If $\phi(c) \vdash_{ND} \psi$ then $\exists v\phi(v) \vdash_{ND} \psi$.
- $=$ Intro: $\vdash_{ND} t = t$.

- $=$ Elim: $t_1 = t_2, \phi(t_1) \vdash_{ND} \phi(t_2)$.

Of course, in the quantifier rules, we assume that t is a term with no variables, and c is a constant that appears nowhere else in any of the mentioned sentences.

Each of these inferences is also good if any set of additional sentences (not containing c) is added to the left side of every $\vdash_{ND}$.

2 Axiomatic Derivations

In an "axiomatic" deduction system, a derivation consists of a finite sequence of statements, each one of which is either a logical axiom, an assumption of the derivation, or follows from earlier statements by use of one of the rules. That is, we say:

> $\Gamma \vdash_A \phi$ iff there is a (finite) sequence of statements in the language, each of which is either a logical axiom, an element of Γ, or follows from earlier statements in the sequence by use of the rules, whose last sentence is ϕ.

In an axiomatic system, there is no such thing as a subderivation, and thus no need to keep track of the scope of assumptions. We also don't need to worry about a rule of accessibility that pays attention to opening and closing subderivations — a sentence is accessible as long as it occurred earlier. Additionally, in the particular axiomatic system we will use, there is only one rule of inference, which makes many meta-theorems about the deduction system much easier to prove.

The main disadvantage of axiomatic systems is that it is often harder to actually carry out a derivation within the system. However, as we will show later, anything that can be derived using one of the systems can in fact be derived using the other. That is, $\Gamma \vdash_H \phi$ iff $\Gamma \vdash_{ND} \phi$. Thus, we will generally use the natural deduction system to actually construct derivations, while using the axiomatic system to state and prove the important meta-theorems about derivability.

2.1 Modus ponens

The one rule of inference is modus ponens. If ϕ occurs on line m, and $(\phi \to \psi)$ occurs on line n, then we can write ψ on a later line, citing "MP m, n".

2.2 The axioms

If ϕ, ψ, and χ are any formulas, then the following are all axioms, provided that we add enough universal quantifiers at the beginning to bind all the free variables. (If ϕ, ψ, and χ are sentences with no free variables, then we don't need any additional universal quantifiers at the front, but they will be allowed.):

$\wedge$I: $(\phi \to (\psi \to (\phi \wedge \psi)))$

$\wedge$E1: $((\phi \wedge \psi) \to \phi)$

$\wedge$E2: $((\phi \wedge \psi) \to \psi)$

$\vee$I1: $(\phi \to (\phi \vee \psi))$

$\vee$I2: $(\psi \to (\phi \vee \psi))$

$\vee$E: $((\phi \vee \psi) \to ((\phi \to \chi) \to ((\psi \to \chi) \to \chi)))$

$\to$1: $(\phi \to (\psi \to \phi))$

$\to$2: $((\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi)))$

$\forall \to$: $(\forall v(\phi(v) \to \psi(v)) \to (\forall v\phi(v) \to \forall v\psi(v)))$, where $\phi(v)$ and $\psi(v)$ are any formulas with v as the only free variable.

$\neg$I: $((\phi \to \bot) \to \neg\phi)$

$\neg$E: $(\phi \to (\neg\phi \to \bot))$

DN: $(\neg\neg\phi \to \phi)$

$\forall$I: $(\phi \to \forall v\phi)$

$\forall$E: $(\forall v\phi(v) \to \phi(t))$, where v is any variable, t is any term containing no variables, $\phi(v)$ is any formula with no free variables other than v, and $\phi(t)$ is the result of replacing all instances of v in $\phi(v)$ by t.

$\exists$I: $(\phi(t) \to \exists v\phi(v))$, with the same conditions.

$\exists$E: $(\exists v\phi(v) \to (\forall v(\phi(v) \to \psi) \to \psi))$

$=$I: $t = t$, where t is any term with no variables.

$=$E: $((t_1 = t_2) \to (\phi(t_1) \to \phi(t_2)))$, where t_1 and t_2 are any terms with no variables, and $\phi(t_1)$ and $\phi(t_2)$ both derive from a single formula $\phi(v)$ with v as the only free variable, by replacing all instances of v by the appropriate term.

Because of the considerations about free variables, the axioms satisfy the Axiom Generalization Theorem:

> If $\phi(v)$ is a formula with no free variables other than v, c is a constant not occurring in $\phi(v)$, and $\phi(c)$ is an axiom, then so is $\forall v\phi(v)$.

I have given these axioms names that are reminiscent of the names of the rules from the Natural Deduction system — that is because there is a natural correspondence (though it is not entirely direct) between these rules that we will use to prove that the two systems have the same theorems. The parallels with the summarized versions of the rules are fairly clear — "$\vdash_{ND}$" corresponds to "$\to$". The only exceptions are the axioms for the conditional (including $\forall \to$) — because the conditional plays such a special role in this system (it is the only connective that matters for the one inference rule), its axioms don't directly correspond to either of the rules for the conditional in the Natural Deduction system.

I have called each of these form an axiom *schema* (plural: "schemas" or "schemata") because each has infinitely many instances. Any sentence of the appropriate form is an axiom, and it is straightforward (if sometimes tedious) to recognize whether or not a sentence fits a given form.

2.3 Sample derivations

A derivation that $P \vdash_A \neg\neg P$

1	$((\neg P \to \bot) \to \neg\neg P)$	$\neg$I
2	$(((\neg P \to \bot) \to \neg\neg P) \to (P \to ((\neg P \to \bot) \to \neg\neg P)))$	$\to$1
3	$(P \to ((\neg P \to \bot) \to \neg\neg P))$	MP 1,2
4	$((P \to ((\neg P \to \bot) \to \neg\neg P)) \to ((P \to (\neg P \to \bot)) \to (P \to \neg\neg P)))$	$\to$2
5	$((P \to (\neg P \to \bot)) \to (P \to \neg\neg P))$	MP 3,4
6	$(P \to (\neg P \to \bot))$	$\neg$E
7	$(P \to \neg\neg P)$	MP 6,5
8	P	Premise
9	$\neg\neg P$	MP 8,7

There may be a way to shorten this following derivation, but I haven't figured it out. Steps 4 to 28 are basically a derivation that $\neg A \vdash_A (A \to C)$ — perhaps there is a way to shorten that long sequence.

A derivation that $A \vee B, \neg A, (B \to C) \vdash_A C$:

1	$(A \vee B)$	Premise
2	$((A \vee B) \to ((A \to C) \to ((B \to C) \to C)))$	$\vee$E
3	$((A \to C) \to ((B \to C) \to C))$	MP 1,2
4	$((\neg C \to \bot) \to \neg\neg C)$	$\neg$I
5	$(((\neg C \to \bot) \to \neg\neg C) \to (\bot \to ((\neg C \to \bot) \to \neg\neg C)))$	$\to$1
6	$(\bot \to ((\neg C \to \bot) \to \neg\neg C))$	MP 4, 5
7	$((\bot \to ((\neg C \to \bot) \to \neg\neg C)) \to ((\bot \to (\neg C \to \bot)) \to (\bot \to \neg\neg C)))$	$\to$2
8	$((\bot \to (\neg C \to \bot)) \to (\bot \to \neg\neg C))$	MP 6, 7
9	$(\bot \to (\neg C \to \bot))$	$\to$1
10	$(\bot \to \neg\neg C)$	MP 8,9
11	$(\neg\neg C \to C)$	DN
12	$((\neg\neg C \to C) \to (\bot \to (\neg\neg C \to C)))$	$\to$1
13	$(\bot \to (\neg\neg C \to C))$	MP 11, 12
14	$((\bot \to (\neg\neg C \to C)) \to ((\bot \to \neg\neg C) \to (\bot \to C)))$	$\to$2
15	$((\bot \to \neg\neg C) \to (\bot \to C))$	MP 13, 14
16	$(\bot \to C)$	MP 10, 15
17	$((\bot \to C) \to (A \to (\bot \to C)))$	$\to$1
18	$(A \to (\bot \to C))$	MP 16, 17
19	$((A \to (\bot \to C)) \to ((A \to \bot) \to (A \to C)))$	$\to$2
20	$((A \to \bot) \to (A \to C))$	MP 18, 19
21	$((A \to (\neg A \to \bot)) \to ((A \to \neg A) \to (A \to \bot)))$	$\to$2
22	$(A \to (\neg A \to \bot))$	$\neg$E
23	$((A \to \neg A) \to (A \to \bot))$	MP 22, 21
24	$(\neg A \to (A \to \neg A))$	$\to$1
25	$\neg A$	Premise
26	$(A \to \neg A)$	MP 25, 24
27	$(A \to \bot)$	MP 26, 23
28	$(A \to C)$	MP 27, 20
29	$((B \to C) \to C)$	MP 28, 3
30	$(B \to C)$	Premise
31	C	MP 30, 29

3 Equivalence

As I mentioned before, these two derivation systems are equivalent.

3.1 Natural Deduction encompasses Axiomatic Deduction

We want to show that if $\Gamma \vdash_A \phi$ then $\Gamma \vdash_{ND} \phi$. To show this, it will suffice to show that we have $\vdash_{ND} \phi$ for every logical axiom ϕ of the axiomatic system. The reason this is sufficient is that we can take the axiomatic derivation that shows $\Gamma \vdash_A \phi$ and turn it into a Natural Deduction derivation.

First, assume that all sentences of Γ occur at the beginning of the axiomatic derivation of ϕ, if they occur at all. (We can do this because those sentences don't depend on anything else having occurred earlier.) Draw a vertical line next to the entire derivation, and put the horizontal dash after the sentences of Γ and before everything else. Every further step in the sequence is either a logical axiom of the axiomatic system (in which case we can precede it with the Natural Deduction derivation of the axiom) or follows from two earlier steps by modus ponens (in which case we can use just $\rightarrow$ Elim). The result is a Natural Deduction derivation that shows that $\Gamma \vdash_{ND} \phi$.

Thus, we just need to construct a Natural Deduction derivation of each axiom in the axiomatic system. This is a useful series of exercises in using the Natural Deduction derivation system, the beginning of which is provided in the summary of the Natural Deduction rules. I will not spell out any of these derivations in further detail here.

To show that anything given by the Generalization Clause is also an axiom, we note that if $\phi(c)$ is an axiom, then we can derive it within a subderivation, and then by the $\forall$ Intro rule, we can conclude $\forall v \phi(v)$ outside the subderivation, as required.

Thus, anything derivable in the axiomatic system is also derivable in the Natural Deduction system.

3.2 Some meta-theorems about the axiomatic system

Before proving the converse, it will be useful to prove some results about derivations in the axiomatic system. These results will be proved by induction on the derivation. Just as the set of terms is the smallest set containing all variables and constants and closed under the function symbols, and the set of formulas is the smallest set containing all of the atomic formulas and closed under the connectives and quantifiers, the set of axiomatic theorems of a given set Γ is the smallest set that contains all the logical axioms and members of Γ, and is closed under modus ponens.

That is, if $\Gamma \vdash_A \phi$ is an arbitrary statement of theorem-hood, then we have to consider three cases. Either ϕ could be a logical axiom, or ϕ could be a member of Γ, or ϕ was derived by an application of modus ponens from earlier statements ψ and $(\psi \rightarrow \phi)$, which are themselves theorems of Γ. So if we show that a property holds of all logical axioms and all members of Γ, and show that if the property holds of ψ and $(\psi \to \phi)$ then it also holds of ϕ, then we can use these three facts to work our way down any derivation of ϕ from Γ to show that the property also holds of ϕ.

The possibility of these inductive proofs is a major advantage of the axiomatic system over the Natural Deduction system — we have many base cases corresponding to all the types of axioms, but we only have one inductive step, and we don't have to consider any special status of subderivations.

3.2.1 The Soundness Theorem

If $\Gamma \vdash_A \phi$, then $\Gamma \models \phi$.

This result is quite important — it partly explains why we use the particular derivation system that we do. It means that if ϕ is a theorem of Γ, then it is in fact a logical consequence of Γ. When we eventually get to our proof that the axiomatic system encompasses the Natural Deduction system, we will be able to conclude that the Natural Deduction system is sound as well. This is just the first instance of a meta-theorem that is easier to prove about the axiomatic system and then transfer to the Natural Deduction system.

The proof is by induction on the derivation of ϕ from Γ.

There are three possibilities — either ϕ is a logical axiom, or it is a member of Γ, or it is derived from two earlier steps in the derivation by modus ponens.

If ϕ is a logical axiom, then we can check that it is true on all interpretations. (This will guarantee that it is true on all interpretations that make all of Γ true.) For considerations of space, I won't go through all the cases here, but will just check a few cases. For instance, in an axiom of type $\wedge$I, we can consider whether the sentences ϕ and ψ come out true on the interpretation — if ϕ is false, then the whole sentence is clearly true, while if ψ is false, then the consequent of the overall conditional is true so the whole sentence is true, and if both are true, then again the whole sentence is true.

The axioms containing quantifiers require some special consideration. Any axiom of the form $\forall$I is straightforwardly true, because the second part has a vacuous quantification. In checking axioms of the form $\forall$E and $\exists$I, we have to make careful use of the conditions on v and t (and this is in fact the reason for these carefully stated conditions). For the versions of the axioms with free variables in the formulas and universal quantifiers out front, note that if $\phi(v)$ is a formula with no free variables other than v, and $\phi(c)$ is an axiom, then every interpretation must make $\phi(c)$ true (by our previous work), and thus $\phi(v)$ is satisfied by every variable assignment on every interpretation (by extensionality), and thus $\forall v \phi(v)$ is true on every interpretation. Thus, we can check that any axiom is true on all interpretations, and thus in particular is true on all interpretations where every sentence of Γ is true.

If ϕ is a member of Γ, then the result is quite straightforward. Any interpretation that makes all members of Γ true must by definition make ϕ true, so $\Gamma \models \phi$.

Finally, if ϕ comes from two previous lines by modus ponens, then we can call those lines ψ and $(\psi \to \phi)$. By truncating the derivation at these earlier steps, we can see that $\Gamma \vdash_A \psi$ and $\Gamma \vdash_A (\psi \to \phi)$. Because this is a proof by induction, we can assume that the Soundness Theorem applies to these two derivations. Thus, $\Gamma \models \psi$ and $\Gamma \models (\psi \to \phi)$. But then any interpretation that makes all sentences of Γ true must make both ψ and $(\psi \to \phi)$ true, and by checking the truth table for $\to$, we can see that it must also make ϕ true, which completes the proof.

3.2.2 The Deduction Theorem

If $\Gamma, \phi \vdash_H \psi$, then $\Gamma \vdash_H (\phi \to \psi)$.

The proof is by induction on the derivation of ψ from $\Gamma \cup \{\phi\}$.

Note that in this case, ψ could be a logical axiom, a member of $\Gamma \cup \{\phi\}$, or it could be the result of modus ponens applied to two earlier statements. It will be useful to separate the case where ψ is a member of Γ from the case where ψ is ϕ.

If ψ is either a logical axiom or a member of Γ, then we can construct a three-step derivation of $(\phi \to \psi)$ from Γ as follows:

1	ψ	Either axiom or from Γ
2	$(\psi \to (\phi \to \psi))$	$\to$1
3	$(\phi \to \psi)$	MP 1,2

If ψ and ϕ are the same sentence, then we can construct a five-step derivation of $(\phi \to \psi)$ (which is the same as $(\phi \to \phi)$) as follows:

1	$(\phi \to ((\phi \to \phi) \to \phi))$	$\to$1
2	$((\phi \to ((\phi \to \phi) \to \phi)) \to ((\phi \to (\phi \to \phi)) \to (\phi \to \phi)))$	$\to$2
3	$((\phi \to (\phi \to \phi)) \to (\phi \to \phi))$	MP 1,2
4	$(\phi \to (\phi \to \phi))$	$\to$1
5	$(\phi \to \phi)$	MP 3,4

In this case, the substitution instances on lines 1 and 2 are a little tricky to see.

The last possibility is that ψ is itself achieved by modus ponens from two earlier steps, say χ and $(\chi \to \psi)$. By truncating the derivation at the earlier steps, we can see that $\Gamma, \phi \vdash_A \chi$ and $\Gamma, \phi \vdash_A (\chi \to \psi)$. Because this is the induction step, we can assume that the Deduction Theorem applies to these two earlier derivations, so that $\Gamma \vdash_A (\phi \to \chi)$ and $\Gamma \vdash_A (\phi \to (\chi \to \psi))$. But by stringing these two derivations together, and then following them by an instance of $\to$2 (in particular, $((\phi \to (\chi \to \psi)) \to ((\phi \to \chi) \to (\phi \to \psi)))$ is the instance we want), and then applying modus ponens twice, we get a derivation from Γ of $(\phi \to \psi)$, as required.

3.2.3 The Generalization Theorem

In the statement of this meta-theorem, $\phi(v)$ will be a formula that has no free variable other than v, and $\phi(c)$ will be the result of replacing all free occurrences of v (if there are any) with c.

> If c is a constant that does not occur in Γ or $\phi(v)$, and $\Gamma \vdash_A \phi(c)$, then $\Gamma \vdash_A \forall v\phi(v)$.

When ϕ is a sentence with no free variables, this has as a special case the claim that if $\Gamma \vdash_A \phi$ then $\Gamma \vdash_A \forall v\phi$, so that one can put any number of vacuous universal quantifiers at the beginning of a formula. This is very much like the Axiom Generalization Theorem, but it applies to *all* theorems derived from a set of premises that don't involve c, and not just the axioms.

Again, the proof is by induction on the derivation from Γ of $\phi(c)$.

There are three possibilities — either $\phi(c)$ is a logical axiom, or it is a member of Γ, or it was achieved by modus ponens applied to two earlier statements in the derivation.

If $\phi(c)$ is a logical axiom, then in fact $\forall v\phi(v)$ is itself a logical axiom, because of the Axiom Generalization Theorem. (This is where we use the fact that c does not occur in $\phi(v)$.) Thus, it has a one-step derivation.

If $\phi(c)$ is a member of Γ, then the fact that c doesn't occur in Γ means that it doesn't appear in $\phi(c)$, which means that v doesn't occur in $\phi(v)$. That is, $\phi(c)$ and $\phi(v)$ are in fact the same formula, which is a sentence that we might call ϕ. Then $(\phi \to \forall v\phi)$ is an axiom of type $\forall$I. By adding this statement to a derivation from Γ of ϕ, and then applying modus ponens, we get a derivation from Γ of $\forall v\phi$.

Finally, if $\phi(c)$ was derived using modus ponens from two earlier steps, then we can call those steps $\psi(c)$ and $(\psi(c) \to \phi(c))$. I explicitly include the mention of c in $\psi(c)$ because it *may* appear, though of course it isn't guaranteed to. At any rate, we have $\Gamma \vdash_A \psi(c)$ and $\Gamma \vdash_A (\psi(c) \to \phi(c))$, where c does not occur in Γ. We can make sure c doesn't occur in $\psi(v)$ by replacing all instances of c by instances of v (though we may have to be careful about the scope of internal quantifiers involving v). Thus, by the induction assumption, we have $\Gamma \vdash_A \forall v\psi(v)$ and $\Gamma \vdash_A \forall v(\psi(v) \to \phi(v))$. But there is an instance of $\forall \to$ that states $(\forall v(\psi(v) \to \phi(v)) \to (\forall v\psi(v) \to \forall v\phi(v)))$. By taking the two previous derivations, following them with this axiom, and applying modus ponens twice, we get $\forall v\phi(v)$, as required.

3.2.4 The Negation Theorem

We say that a set Γ is consistent iff it is not the case that $\Gamma \vdash_A \bot$. Equivalently, we say that it is inconsistent iff $\Gamma \vdash_A \bot$.

Exercise: Prove that Γ, ϕ is inconsistent iff $\Gamma \vdash_A \neg\phi$, and that $\Gamma, \neg\phi$ is inconsistent iff $\Gamma \vdash_A \phi$.

3.3 Axiomatic Deduction encompasses Natural Deduction

We want to show that if $\Gamma \vdash_{ND} \phi$ then $\Gamma \vdash_A \phi$. To do this, we will note that every line in a natural deduction derivation is associated with a particular claim of the form $\Gamma \vdash_{ND} \phi$, where Γ is the set of all assumptions from the main derivation and all subderivations that are still going at that line, and ϕ is the sentence on that line. We will work by induction on the steps in the Natural Deduction derivation, and show that if the claim associated with all earlier lines holds, then the claim associated with the current line does as well.

So consider the line containing ϕ. ϕ must either be one of the assumptions of the main derivation, the assumption for a subderivation, or be the result of the application of one of the rules.

If ϕ is an assumption of the main derivation or a subderivation, then the corresponding claim $\Gamma \vdash_{ND} \phi$ will have ϕ itself as a member of Γ. But in this case, it is clear that we can construct an axiomatic derivation that shows that $\Gamma \vdash_H \phi$, just by considering the one-line derivation consisting of just ϕ. So we just need to consider the cases where ϕ is a result of applying one of the rules of inference.

Note that the rule of accessibility guarantees that if ψ is cited in justifying the line where ϕ occurs, then all subderivations that were still going when ψ occurred must still be going when ϕ occurs. Thus, if the claim associated with ϕ is $\Gamma \vdash_{ND} \phi$, then the claim associated with ψ must be $\Gamma' \vdash_{ND} \psi$, where Γ' is a subset of Γ. Similarly, if a subderivation beginning with α and ending with β was cited, then the claim associated with β must be $\Gamma', \alpha \vdash \beta$, where Γ' is a subset of Γ. Since all the cited sentences and subderivations must occur earlier than ϕ, we can assume that all of their associated claims have axiomatic derivations as well. These axiomatic derivations will thus also give axiomatic derivations of $\Gamma \vdash_H \psi$ and $\Gamma, \alpha \vdash_H \beta$. These observations already make the Reiteration rule trivial, and now we just have to go through each of the connective and quantifier rules and show that if the cited steps are associated with claims that have axiomatic derivations, then so is ϕ.

$\wedge$ Intro: By assumption, we have $\Gamma \vdash_H \phi$ and $\Gamma \vdash_H \psi$. We need to show that $\Gamma \vdash_A (\phi \wedge \psi)$. We can take a derivation of ϕ, followed by a derivation of ψ, followed by an axiom of the form $\wedge$I, followed by two applications of modus ponens, to get a derivation of $(\phi \wedge \psi)$.

$\wedge$ Elim: By assumption, we have $\Gamma \vdash_H (\phi \wedge \psi)$. We need to show that $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$. But we can just take a derivation of $(\phi \wedge \psi)$ from Γ, follow it by an axiom of the form $\wedge$E1 or $\wedge$E2, and use a single application of modus ponens.

$\vee$ Intro: This is very similar to the previous cases, with with axioms of the form $\vee$I1 or $\vee$I2.

$\vee$ Elim: By assumption, we have $\Gamma \vdash_H (\phi \vee \psi)$. Additionally, by considering the two subderivations ending in χ, we have $\Gamma, \phi \vdash_A \chi$ and $\Gamma, \psi \vdash_A \chi$. By the Deduction Theorem, we have $\Gamma \vdash_A (\phi \to \chi)$ and $\Gamma \vdash_A (\psi \to \chi)$. By putting these three derivations together with an axiom of type $\vee$E, and applying modus ponens three times, we get a derivation from Γ of χ.

$\to$ Intro: By considering the subderivation, we have $\Gamma, \phi \vdash_A \psi$. By the

Deduction Theorem, we have $\Gamma \vdash_A (\phi \to \psi)$, as required.

$\to$ Elim: By assumption, we have $\Gamma \vdash_A \phi$ and $\Gamma \vdash_A (\phi \to \psi)$. By a simple application of modus ponens, we have $\Gamma \vdash_A \psi$.

$\neg$ Intro: This is very similar to $\vee$ Elim and $\to$ Intro, where we consider a subderivation, and apply an axiom of type $\neg$I.

$\neg$ Elim: This is very similar to the rules without subderivations, but with an axiom of type $\neg$E.

Double-Negation: This is also very similar, but with an axiom of type DN.

$\forall$ Intro: By assumption we have $\Gamma \vdash_A \phi(c)$, and c doesn't occur in either Γ or $\phi(v)$. Thus, by the Generalization Theorem, we have $\Gamma \vdash_A \forall v \phi(v)$.

$\forall$ Elim: This is like earlier rules, but with an axiom of type $\forall$E.

$\exists$ Intro: This is like earlier rules, but with an axiom of type $\exists$I.

$\exists$ Elim: By assumption we have $\Gamma, \phi(c) \vdash_A \psi$, where c doesn't occur in Γ or $\phi(v)$ or ψ. Thus, by the Deduction Theorem, we have $\Gamma \vdash_A (\phi(c) \to \psi)$. By the Generalization Theorem, we have $\Gamma \vdash_A \forall v(\phi(v) \to \psi)$. Also by assumption, we have $\Gamma \vdash_A \exists v \phi(v)$. Putting these two derivations together, followed by an axiom of the form $\exists$E, and two uses of modus ponens, gives a derivation that $\Gamma \vdash_A \psi$, as required.

$=$ Intro: $t = t$ is in fact an axiom of type $=$I, so $\Gamma \vdash_A t = t$.

$=$ Elim: This is like earlier rules, but with an axiom of type $=$E.

Thus, if all earlier steps in a Natural Deduction derivation correspond to axiomatic derivations, then so does the current line, and thus by induction, if $\Gamma \vdash_{ND} \phi$, then $\Gamma \vdash_A \psi$, and we can finally see that the two systems are equivalent.

4 Completeness Theorem

Recall that the Completeness Theorem states that if $\Gamma \models \phi$, then $\Gamma \vdash \phi$. By the Negation Theorem, we know that $\Gamma \vdash \phi$ iff $\Gamma, \neg\phi$ is inconsistent. Similarly, we know that $\Gamma \models \phi$ iff $\Gamma, \neg\phi$ is unsatisfiable. Thus, to prove the Completeness Theorem, it will suffice to show that if $\Gamma, \neg\phi$ is unsatisfiable, then it is inconsistent. Taking the contrapositive, it suffices to show that if $\Gamma, \neg\phi$ is consistent, then it is satisfiable. Thus, we want to show that every consistent set of a sentences has an interpretation that makes every sentence in the set true.

If we start with an arbitrary consistent set of sentences, we will construct the interpretation as follows. First, we will "add witnesses" to the set, and show that the resulting set is still consistent. Then we will extend this set to a "maximal" consistent set that has witnesses. Finally, we will show that from a maximal consistent set that has witnesses, one can define an interpretation that makes every sentence in that set true. Since our initial set is a subset of the final one, this interpretation makes every sentence in the initial set true.

So let Γ be the initial consistent set.

4.1 Adding Witnesses

If Γ contains a sentence of the form $(\exists v\phi(v) \rightarrow \phi(c))$, then we will say that c is a "witness" for $\phi(v)$. Sentences of this form are called "Henkin sentences", because they were introduced by Leon Henkin for the purpose of this proof of the Completeness Theorem. (Gödel's original proof was even more technical than this one.) The reason the constant is called a "witness" is because any interpretation satisfying this Henkin sentence that gives any object the property expressed by $\phi(v)$ will have to make sure that c is interpreted as one of those objects. Thus, for any interpretation that makes all of Γ true, we can figure out whether or not the existential claim is true just by looking at the particular object denoted by c on that interpretation — it is the only "witness" we have to question to see if the sentence is true.

What we want to show in this section is that if Γ is any consistent set of sentences, then there is a consistent set Δ containing Γ that has witnesses for *every* formula in the language. Of course, in order to make sure that these sentences are all consistent, we will have to have a different witness in each Henkin sentence. Additionally, we will need to make sure that these witnesses are not constants that Γ already talks about, or else the Henkin sentence may contradict some other facts about it.

Thus, we do the following. First, we will add an infinite enumerable set of new constants to the language, to make sure that they aren't all already used in Γ. Because even this new language is enumerable, we can enumerate all formulas in this language that have one free variable, as $\phi_0(v), \phi_1(v), \phi_2(v), \ldots$. Because each of these formulas is finitely long, they only mention finitely many of the new constants. Thus, for each formula, we can choose a new constant that we will use in the corresponding Henkin sentence, and we can always make sure that the constant hasn't appeared in Γ (since it is one of the new constants) or one of the earlier Henkin sentences (because only finitely many of them have appeared in earlier sentences). That is, we will start with the set $\Gamma_0 = \Gamma$ and then, one by one, add the sentences $\exists v\phi_1(v) \rightarrow \phi_1(c_1)$, $\exists v\phi_2(v) \rightarrow \phi_2(c_2)$, and so on. That is, $\Gamma_{n+1} = \Gamma_n \cup \{(\exists v\phi_n(v) \rightarrow \phi_n(c_n))\}$.

First, we show that if Γ_n is consistent, then Γ_{n+1} is consistent as well. That is, if we start with a consistent set of sentences and add a single Henkin sentence involving a new constant, then the resulting set is consistent as well. Equivalently, we will show that if the resulting set is inconsistent, then the starting set was as well. If $\Gamma_n, (\exists v\phi_n(v) \rightarrow \phi_n(c_n)) \vdash \perp$, then by the Negation Theorem, $\Gamma_n \vdash \neg(\exists v\phi_n(v) \rightarrow \phi(c_n))$. By some manipulations, we can then show that $\Gamma_n \vdash \exists v\phi_n(v)$ and $\Gamma_n \vdash \neg\phi_n(c_n)$. (This just involves constructing a derivation that shows $\neg(A \rightarrow B) \vdash (A \wedge \neg B)$.) By the Negation Theorem, $\Gamma_n, \phi_n(c) \vdash \perp$, and by the Deduction Theorem, $\Gamma_n \vdash (\phi_n(c_n) \rightarrow \perp)$. But then, by the Generalization Theorem, since c_n doesn't appear in Γ_n or $\phi_n(v)$, we see that $\Gamma_n \vdash \forall v(\phi_n(v) \rightarrow \perp)$. But since $\Gamma_n \vdash \exists v\phi_n(v)$, we can use an instance of $\exists$E to show that $\Gamma_n \vdash (\forall v(\phi_n(v) \rightarrow \perp) \rightarrow \perp)$. Thus, by one more modus ponens, we see that $\Gamma_n \vdash \perp$. Thus, if Γ_n is consistent, and c_n does not occur in Γ_n or $\phi_n(v)$, then $\Gamma_n, (\exists v\phi_n(v) \rightarrow \phi(c_n))$ is consistent as well.

Now, let Δ be the union of the Γ_n, so that it contains Γ together with all infinitely many Henkin sentences. Next, we will use a fact, called "compactness", stating that if every finite subset of a set is consistent, then the full set is as well. In this case, every finite subset of Δ is a subset of one of the Γ_n, and is thus consistent. To see that consistency of each finite subset implies consistency of the whole set, consider the contrapositive. If the whole set were inconsistent, then there would be a derivation of $\perp$ from it. But this derivation would be finitely long, and thus would include only finitely many sentences. The finitely many members of Δ that were used would then be a finite subset of Δ that would be inconsistent. So if every finite subset is consistent, then the whole set is consistent. (This fact is called "compactness" because there are many other related mathematical facts in the field of topology that also involve infinite sets with a property being reducible to a finite subset with the same property, that intuitively have more to do with our geometric concept of being "compact".) Thus, if Γ is consistent, then each Γ_n in the process is consistent, and so the result Δ (which has witnesses for every formula) is as well.

Thus, if Γ is a consistent set of sentences, then there is a consistent set Δ that contains all of Γ, and has witnesses for every formula in the language.

4.2 Maximal Consistent Set

The next step in the construction of an interpretation is to take a consistent set Δ that has witnesses for every formula in the language, and extend it to a *maximal* consistent set Ω. What it means for a consistent set to be maximal is that for any sentence of the language not in the set, adding that sentence to the set would result in an inconsistent set. The fact that every consistent set of sentences can be extended to a maximal consistent set is known as Lindenbaum's Lemma, after Adolf Lindenbaum, the colleague of Tarski that proved it. (Lindenbaum was married to Janina Hosiasson, who did important work in Bayesian confirmation theory. Both of them were murdered by the Nazis during World War II.)

To prove Lindenbaum's Lemma, we first notice that the sentences of the language are enumerable, so we can enumerate them as $\phi_0, \phi_1, \phi_2, \ldots$. Then we define the sets $\Delta_0, \Delta_1, \Delta_2, \ldots$ as follows. $\Delta_0 = \Delta$, and $\Delta_{n+1} = \Delta_n \cup \{\phi_n\}$ if that set is consistent, and $\Delta_{n+1} = \Delta_n$ otherwise. Note that by this construction, each Δ_n will be consistent, if Δ itself was consistent. Let Ω be the union of all the Δ_n (that is, it is the limit of the process). By the compactness result mentioned above, since any derivation of a contradiction from Ω would use only finitely many of the sentences in Ω, we can see that if Ω were inconsistent, then one of the Δ_n would be as well, and since each of those is consistent, then Ω must be as well.

Since Ω is consistent, we must just show that it is maximal. So consider any sentence that is not in Ω. This sentence must show up at some point in our enumeration, as ϕ_n. Since ϕ_n is not a member of Ω, this means it must not have been added at the nth stage of our construction, which means that Δ_n, ϕ_n is not consistent. But since Δ_n is a subset of Ω, any derivation of a contradiction from Δ_n and ϕ_n must also be a valid derivation from Ω and ϕ_n, and thus adding ϕ_n to Ω would result in an inconsistent set. Thus, Ω is a maximal consistent set.

Therefore, starting with a consistent set Γ, we can add witnesses for every formula to get a consistent Δ, and then extend this to a maximal consistent set Ω, which still has witnesses for every formula.

4.3 Results about a Maximal Consistent Set with Witnesses

Now we will prove some results about Ω, from the fact that it is a maximal consistent set with witnesses.

4.3.1 For any sentence ϕ in the language, ϕ is a member of Ω iff $\Omega \vdash \phi$.

The left to right direction is clear — if ϕ is a member of Ω, then there is a one-step derivation of ϕ. So we must prove the other direction. So assume that $\Omega \vdash \phi$. But then anything one can derive from Ω, ϕ can also be derived from Ω alone, by first deriving ϕ. Thus, since Ω is consistent, it doesn't derive $\bot$, and therefore Ω, ϕ doesn't either. Thus, since Ω is maximal, ϕ must already be a member of Ω, as required.

Now, using this fact, we will prove results about which sentences are members of Ω, one for each logical symbol of the language.

4.3.2 $(\phi \wedge \psi)$ is a member of Ω iff ϕ and ψ are both members of Ω.

This is clear because $\Omega \vdash (\phi \wedge \psi)$ iff $\Omega \vdash \phi$ and $\Omega \vdash \psi$.

4.3.3 $(\phi \vee \psi)$ is a member of Ω iff at least one of ϕ or ψ is a member of Ω.

The right to left direction is clear, because if Ω derives at least one of those two sentences, then it will also derive their disjunction. So we must show the left to right direction. So assume that $(\phi \vee \psi)$ is a member of Ω. If ϕ is not a member of Ω, then Ω, ϕ is inconsistent (because Ω is maximal). Similarly, if ψ is not a member of Ω, then Ω, ψ is inconsistent. Thus, if neither is, then we would have $\Omega \vdash (\phi \vee \psi)$, and $\Omega \vdash (\phi \rightarrow \bot)$ and $\Omega \vdash (\psi \rightarrow \bot)$ (by the Negation Theorem), and thus by $\vee$E, we would have $\Omega \vdash \bot$, which is a contradiction. Therefore, if $(\phi \vee \psi)$ is in Ω, then at least one of its disjuncts must be as well.

4.3.4 $(\phi \rightarrow \psi)$ is a member of Ω iff either ϕ is not a member of Ω or ψ is.

Left to right: If $(\phi \rightarrow \psi)$ is a member of Ω, and ϕ is a member of Ω, then an application of modus ponens would show that $\Omega \vdash \psi$, and thus by the earlier result, ψ would be a member of Ω. Thus, if $(\phi \to \psi)$ is a member of Ω, then either ϕ is not a member, or ψ is.

Right to left: If ϕ is not a member of Ω, then Ω, ϕ is inconsistent. Since an inconsistent set of sentences derives everything (this is straightforward to check on the Natural Deduction system, but a derivation of this sort can be done in the axiomatic system as well), this means that $\Omega, \phi \vdash \psi$. But then $\Omega \vdash (\phi \to \psi)$ by the Deduction Theorem, so $(\phi \to \psi)$ is in Ω. If ψ is a member of Ω, then since $(\psi \to (\phi \to \psi))$ is an axiom of type $\to$1, we would have $\Omega \vdash (\phi \to \psi)$, so again it is in Ω.

4.3.5 $\neg\phi$ is a member of Ω iff ϕ isn't a member of Ω.

This is a straightforward consequence of the Negation Theorem and the fact that Ω is maximal.

4.3.6 $\forall v \phi(v)$ is a member of Ω iff $\phi(c_i)$ is a member of Ω for every constant c_i added at the stage of adding witnesses.

The left to right direction is straightforward, since $\forall v \phi(v) \vdash \phi(c_i)$. Right to left: This direction will be easier to do by contrapositive, so assume that $\forall v \phi(v)$ is not a member of Ω. Then by the result about negation, we see that $\neg\forall v \phi(v)$ is a member of Ω. Since $\neg\forall v \phi(v) \vdash \exists v \neg\phi(v)$ (constructing this derivation is left as an exercise), we see that this sentence is in Ω as well. Since $\neg\phi(v)$ is a formula with one free variable, and Ω contains witnesses for every formula, we see that some sentence of the form $(\exists v \neg\phi(v) \to \neg\phi(c_i))$ is a member of Ω, and so by modus ponens, we see that $\Omega \vdash \neg\phi(c_i)$, and thus it is a member of Ω as well. But then $\phi(c_i)$ is not in Ω, since it is consistent. Thus, if $\forall v \phi(v)$ is not a member of Ω, then there is some constant c_i added with the witnesses, such that $\phi(c_i)$ is not in Ω, which completes the right to left direction.

4.3.7 $\exists v \phi(v)$ is a member of Ω iff there is some constant c_i added at the stage of adding witnesses such that $\phi(c_i)$ is a member of Ω.

The left to right direction again follows from the fact that Ω contains witnesses, and the right to left direction follows from the $\exists$I axiom.

This result has a straightforward consequence that for every term t in the language, there is a constant c_i added as one of the witnesses such that $t = c_i$ is in Ω. This follows because $t = t$ is an axiom of $=$I, and thus $\exists v(t = v)$ is in Ω, by $\exists$I.

4.4 Constructing the Interpretation

The interpretation itself will be somewhat strange. To specify an interpretation we must specify the domain of the interpretation, assign an object in this domain to each constant in the language, assign a function on this domain to each function symbol, and assign a set of n-tuples in the domain to each n-place relation symbol. We will also eventually want to show that a sentence is true on this interpretation iff the sentence is in Ω, and since Γ is a subset of Ω, we can see that the interpretation will make every element of Γ true. Thus, constructing this interpretation and showing that it makes all and only the sentences of Ω true will show that if Γ is consistent, then there is an interpretation that makes all of its sentences true.

Domain: The domain of the interpretation will be a subset of the natural numbers. In particular, consider the constants $c_0, c_1, c_2, \ldots$ added to the language along with the witnesses. We will say that a number n is in the domain iff no sentence of the form $(c_n = c_i)$ is in Ω for any $i < n$.

As another way to think about this, we can define a relation on the natural number $m \approx n$ iff $(c_m = c_n)$ is in Ω. We can show that this relation is an equivalence relation, and then the domain will consist of the smallest number in each equivalence class.

Constants: To assign the interpretation of each constant, we will say that for constant c, we find the lowest n such that $(c = c_n)$ is a sentence in Ω, and let this n be the interpretation of c. We can see that this does in fact assign a member of the domain to every constant: by the last result of the previous section, there is always a sentence of the form $(c = c_i)$ in Ω for some witness constant c_i, and then either i is in the domain, or else there is some $n < i$ that is in the domain such that $(c_n = c_i)$ is in Ω, and then $(c = c_n)$ must also be in Ω, because Ω derives the transitivity of identity.

Functions: We will also use Ω to specify the interpretation of each function symbol. For instance, if f is a binary function symbol in the language, and m and n are two numbers in the domain, then we need to say what the function assigned to f does to this pair of numbers as inputs. But then we note that $f(c_m, c_n)$ is a term in the language, and so there is some sentence of the form $(f(c_m, c_n) = c_i)$ in Ω, and we say that the interpretation of f sends (m, n) to i, if i is the lowest such number (and thus in the domain of the interpretation). We will do the same for functions with arbitrarily many places.

Lemma: For any term t with no variables, and for i in the domain of the interpretation, the interpretation of t is i iff $(t = c_i)$ is in Ω.

Proof: If t is just a constant, then this is true by definition. Otherwise, t is the result of applying an n-place function symbol f_n to terms $t_1, \ldots, t_n$. Let the interpretations of these terms be $i_1, \ldots, i_n$. By the interpretation of f_n, this means that the interpretation of t is the lowest number i such that the sentence $(f_n(c_{i_1}, \ldots, c_{i_n}) = c_i)$ is in Ω. We can assume the lemma is true for $t_1, \ldots, t_n$, so that sentences $(t_1 = c_{i_1}), \ldots, (t_n = c_{i_n})$ are all in Ω. By applying axiom =E enough times, we can show that the sentence $(t = f_n(c_{i_1}, \ldots, c_{i_n}))$ is a consequence of these sentences, and is thus in Ω. By applying this one more time, we can show that $(t = c_i)$ is in Ω, as required. This proof can be carried out in reverse as well.

Relations:

We will also use Ω to specify the interpretation of each relation symbol. For instance, if R is a binary relation symbol in the language, and m and n are two numbers in the domain, then we need to say whether the relation assigned to R holds of this pair of numbers. But then we note that $R(c_m, c_n)$ is a sentence in the language, and so either it or its negation is in Ω. We say that the relation holds of m and n iff this sentence is in Ω.

Lemma: For any atomic sentence ϕ, ϕ is true on the interpretation iff ϕ is in Ω.

Proof: Since ϕ is an atomic sentence, it is $R_n(t_1, \ldots, t_n)$, where R_n is some n-place relation symbol, and $t_1, \ldots, t_n$ are terms with no variables. Let the interpretations of these terms be $i_1, \ldots, i_n$. By the interpretation of R_n, the sentence is true iff $R_n(c_{i_1}, \ldots, c_{i_n})$ is in Ω. By the previous lemma, the sentences $(t_1 = c_{i_1}), \ldots, (t_n = c_{i_n})$ are all in Ω. By applying axiom =E enough times, we can show that the sentence $R(t_1, \ldots, t_n)$ is in Ω iff $R_n(c_{i_1}, \ldots, c_{i_n})$ is. Putting these biconditionals together, we get the claimed result.

Main Theorem: For *any* sentence ϕ, ϕ is true on the interpretation iff ϕ is in Ω.

Proof: By induction. The second lemma proves the base case, where ϕ is atomic. The cases for the connectives are all done by means of the minor results from the previous subsection. If ϕ is $\exists v \psi(v)$, then ϕ is true iff there is some object i in the domain that satisfies $\psi(v)$. Since c_i denotes i on this interpretation, the Extensionality Theorem (from the book) guarantees that ϕ is true iff $\psi(c_i)$ is, for some i in the domain. By our induction assumption, $\psi(c_i)$ is true iff it is in Ω. By the result about $\exists$ from the previous subsection, $\exists v \psi(v)$ is in Ω iff one of these $\psi(c_i)$ is. Putting the biconditionals together, we see that ϕ is in Ω iff it is true. By similar means, we can prove the induction step for the universal quantifier.

Thus, we have constructed an interpretation and shown that a sentence is true on this interpretation iff it is in Ω. Thus, Ω is satisfiable, as are all of its subsets, including Γ. Since the construction of Ω from Γ can be done for any consistent Γ, we can construct an interpretation satisfying any consistent set of sentences, which is the Completeness Theorem.