Logic

Aditya Neeraje

June 2024

1 What is Logic?

A logic is a language equipped with rules for deducing the truth of one sentence from another. They are designed to avoid the ambiguities and paradoxes associated with natural languages. For instance, consider the following sentences:

Define n as the smallest natural number that cannot be defined in fewer than 20 words.

If this sentence is true, then Germany borders China.

or the simpler sentence: "I am lying". Logic avoids these paradoxes and enables sentences to be processed by programming languages such as Prolog. First Order Logic contains certain standard symbols: $\neg$ for "not", $\vee$ for "or", $\wedge$ for "and", $\rightarrow$ for "implies", $\leftrightarrow$ for "iff", $\forall$ for "for all", $\exists$ for "there exists", $\top$ to represent a tautology and $\perp$ to represent a contradiction.
In FOL, every variable (eg. x, y, z) and every constant (eg. a, b, c) is a term, as is $f(t_1, \cdots, t_n)$ for a function (we will explain functions soon).
Atomic formulas are sentences of the form $t_1 = t_2$ or $R(t_1, \cdots, t_n)$ for some relation (again, to be explained soon). Formulas are derived from atomic formulas by repeated application of negation, conjunction and disjunction, existential and universal quantifiers. To eliminate brackets, it is generally assumed that $\neg$, $\exists$ and $\forall$ have priority over $\vee$ and $\wedge$, and that between symbols of equal priority, they are evaluated from right to left.

2 Semantics of First Order Logic:

Definition 2.1. A $\mathcal{V}$-structure M consists of a non-empty set called the Universe (U) and describes certain constants, functions and relations. Each constant is assigned a particular number from the Universe, each n-ary function takes every element of U^n to U and every n-ary relation is a subset of U^n. Symbols for these constants, functions and relations are taken from the Vocabulary $\mathcal{V}$ of M.

For instance, consider the structures $M = (\mathbb{N}|S)$ where S is the binary successor relation on natural numbers (aSb if b is the successor of A). Let $M' = (\mathbb{N}|s)$ where s is the unary successor function (f(a) = a+1). These are structures for the vocabulary $\mathcal{V} = \{S\}$ and $\mathcal{V} = \{s\}$ respectively.

Definition 2.2. Loosely, a sentence is a formula in which every variable is bound by quantifiers. For instance, $\forall x \exists y(x+1 = y)$ is a sentence, but $\forall x(x+1 = y)$ is not, since y is not bound. Every variable must be bound each time it occurs, hence $\forall x(\exists y(x + 1 = y) \vee (x = y))$ is not a sentence since y is not bound in its second occurence[1].

Definition 2.3. A structure is said to model a First Order Logic sentence ψ if: Base Case: If ψ is an atomic formula, it is of the form $t_1 = t_2$ or $R(t_1, \cdots, t_n)$. In the first case, since ψ is a sentence, t_1 and t_2 are both constants. M models ψ if it interprets the constants t_1 and t_2 as the same element of U. In the second case, M models $R(t_1, \cdots, t_n$ if $(a_i, \cdots, a_n)$ is in the subset of U^n assigned to R, where $(a_i, \cdots, a_n)$ is what M interprets $(t_i, \cdots, t_n)$ as.
The truth tables for negation, conjunction and disjunction are commonly known, so we skip those. If ψ is of the form $\exists x \phi(x)$ where M modeling $\phi(c)$ is well-defined under induction on the complexity of formulas (i.e induction on the number of occurences of the special symbols), M $\models \phi$ iff M models $\phi(c)$ for some constant c in the expanded vocabulary $\mathcal{V}$(M) (intuitively, the expanded vocabulary simply assigns constants to every element in the Universe of M. For instance, if the Universe of M is a singleton, $\mathcal{V}$(M) could be the union of the original vocabulary with a single constant 'a' that M interprets as the element of U)[2][3].

Definition 2.4. Let $S = A_1, \cdots A_n$ be a set of atomic formulas. An *assignment* of S is a function $\mathcal{A} : s \to 0, 1$.

Let $\mathcal{A}$ be an assignment of S and let F be a formula with atomic formulas from S. $\mathcal{A}$ *models* F if the truth value of F when each atomic formula is assigned the truth value that A assigns to that atomic formula is 1. We denote this as $\mathcal{A} \models F$.

Definition 2.5. A sentence is valid or is a tautology if it holds under every possible structure. For instance, the formula (a=b)$\vee\neg$(a=b) is a tautology.

[1] Formally, the set of free (not bound) variables in a formula is inductively created as follows: 1) all variables in a quantifier-free formula are free. 2) The negation of a formula has the same set of free variables. 3) The conjunction of two formulas has the union of their individual sets of free variables as the resultant set. 4) Adding a quantifier for a particular variables removes it from the set of free variables constructed until that point

[2] An expansion of the vocabulary can actually include more relations or functions, too, but we will restrict ourselves to adding constants for now

[3] The expanded vocabulary could have multiple symbols for the same constant. For instance, you could add the constants a, b and c, and make M interpret them all as the number 1

Note that what we take as true in convention need not hold. A model could assign the truth value 0 to the statement "2 + 2 = 4"by interpreting '+' as subtraction on reals, taking away George Orville's freedom.

Definition 2.6. A sentence is satisfiable if it is modeled by some structure.

Definition 2.7. A sentence is unsatisfiable or is a contradiction if it is not modeled by any structure.

Example: Is the following sentence satisfiable? Is it valid?

$$\forall x \exists y (y \cdot y = x)$$

Answer: It is satisfiable, as seen by the model which interprets $\cdot$ as complex number multiplication and has the complex numbers as its Universe. It is not valid, since not every real number has a real square root, hence a structure with the reals as its universe and interpreting $\cdot$ as multiplication would not model this.
Example: Let $\mathcal{V}_E$ be the vocabulary {E} consisting of one binary relation. Consider the conjunction of the following sentences:

$$\forall x E(x, x)$$

$$\forall x \forall y (E(x, y) \implies E(y, x))$$

$$\forall x \forall y \forall z (E(x, y) \wedge E(y, z) \rightarrow E(x, z))$$

All three sentences together define E as an equivalence relation. By demonstrating models for the pairwise conjunction of two of these sentences which do not interpret E as an equivalence relation, we can conclude that all 3 statements are necessary.
Answer: For the conjunction of the first two sentences, simply take the structure with $\mathbb{Z}$ as the Universe and E as the set of all ordered pairs of adjacent integers. For the conjunction of the second two sentences, let the relation E be empty on some non-empty Universe, thus not satisfying reflexivity. For the conjunction of sentences 1 and 3, consider E as the $\leq$ relation on the real numbers.

Remark. FOL operators can be completely described using $\neg$, $\wedge$ and $\exists$. The other symbols are conventionally used for convenience.[4]

3 Importance of FOL to Computer Science

There are various degrees of logic - the most basic, Propositional Logic, only allows for variables and the $\neg, \vee$ and $\wedge$ operators on Atomic Formulas that .

[4]This is similar to the observation that the NAND and NOR operators can be used to construct all Logic gates. We can also prove that some combinations do not have this expressive power ($\neg$ and $\oplus$ cannot define a formula which takes (False,True) to a different truth value than it takes from (True,False) to, $\vee$ and $\wedge$ together cannot define a formula which takes (False,$\cdots$, $False$) to True). $\rightarrow$ and $\perp$ are capable of constructing all truth tables (This is a good exercise to try).

For instance, Propositional Logic satisfies the condition that if an infinite set of statements is unsatisfiable, some finite subset must be unsatisfiable as well (compactness) and the condition that a set of formulas derives another formula iff the union of all those formulas is unsatisfiable (completeness). This necessarily comes at a tradeoff with respect to expressive power - as you can imagine, a language with no quantifiers, no ability to describe functions and no structures such as sets has more theoretical implications than practical ones.
FOL greatly expands the expressive power by introducing quantifiers, functions and relations. All finite graphs and finite relational databases can be described uniquely using FOL. However, it does not have the concept of sets (which comes in Second Order Logic), and is still incapable of expressing any difference between the natural numbers and the real numbers, explaining concepts such as Cauchy sequences or expressing that any model of the sentence is a well-ordered set.
Languages such as Prolog utilize FOL to quickly verify the satisfiability of certain statements over a finite relational database and for natural language processing.
Monadic second order logic is helpful in understanding finite state automata and regular expression matching.

4 Useful Properties of Logic

Definition 4.1. A 'theory' is a set of consistent sentences. That is, no contradiction can be derived from the statements in that set.

Theorem 1. *Every theory has a model and every model has a theory.*

Proof: We first prove that every model has a theory. Denote by $Th(M)$ the set of all sentences that hold in M. We need to show that this set is consistent and is thus a theory.
We show that the theory of any model is complete over the vocabulary $\mathcal{V}$ of the theory. This means that for any sentence ψ that can be expressed using only symbols from V, either ψ or $\neg\psi$ is in $Th(M)$ and it is not the case that ψ and $\neg\psi$ are both in $Th(M)$.
We prove this by induction on the number of occurences of the primitive symbols $\neg$, $\wedge$ and $\exists$ in ψ. For the base case, if ψ is atomic, it is of the form $t_1 = t_2$ or $R(t_1, \cdots, t_n)$. Looking at the definitions of $M \models t_1 = t_2$ and $M \models R(t_1, \cdots, t_n)$ above, this is easy to prove.
By the semantics of $\neg$, $M \models \neg\psi$ iff M does not model ψ. By the semantics of $\wedge$, $M \models \psi \wedge \phi$ iff $M \models \psi$ and $M \models \phi$. Finally, suppose ψ has the form $\exists x\phi$. Then $M \models \psi$ iff there exists some c such that the expansion M_c[5] models $\phi(c)$. By this induction, we see that whether or not M models a V-sentence is

[5] First, expand the vocabulary of $\mathcal{V}$, the vocabulary of M, to have a constant for each element in U and let the new vocabulary be $\mathcal{V}_c$. Then let M_c be the expansion of M to a $\mathcal{V}_c$-structure which interprets these new constants as distinct elements of its Universe.

well-defined and $Th(M)$ is complete.
Having concluded that the theories of all models are consistent, we can observe that any satisfiable set of sentences $\mathcal{T}$ is consistent, since it is a subset of the theory of its model, and FOL has the property of monotonicity - if a formula can be derived from a set of formulas taken as hypotheses, then it can be derived from any superset of that set of formulas.

For the other direction, we use Henkin construction:
Let $\mathcal{V}$ be the vocabulary of $\mathcal{T}$. Expand this vocabulary to include as many constants as the cardinality of $\mathcal{V}$ itself and let the new vocabulary be $\mathcal{V}'$. Enumerate all $\mathcal{V}'$ sentences as $\{\psi_\beta \mid \beta < \alpha\}$ where α is the cardinality of the vocabulary[6][7]. Now, we proceed to construct a complete $\mathcal{V}'$-theory T with the following two properties:

1. Every sentence of $\mathcal{T}$ is in T.
2. For every $\mathcal{V}'$-sentence in T of the form $\exists x \psi(x)$, the sentence $\psi(c)$ exists in T for some constant c in $\mathcal{V}'$.

We now construct T iteratively. Let T_0 be $\mathcal{T}$. Now suppose for some nonzero $\beta < \alpha$ T_γ has been defined for all $\gamma < \beta$. We assume that for each $\gamma < \beta$, T_γ uses at most $|\gamma| + |\mathbb{N}|$ constants. We now define T_β. Either $\beta = \gamma + 1$ for some γ (i.e it is a successor ordinal) or it is a limit ordinal. In the first case, we consider three cases:

1. If $T_\gamma \cup \{\neg\psi_\gamma\}$ is consistent, set $T_{\gamma+1} = T_\gamma \cup \{\neg\psi_\gamma\}$.
2. Else if ψ_γ is not of the form $\exists x \phi(x)$, then let $T_{\gamma+1} = T_\gamma \cup \{\psi_\gamma\}$.
3. Else $T_{\gamma+1} = T_\gamma \cup \{\psi_\gamma\} \cup \phi(c)$ for some c that has not been used yet.

If β is a limit ordinal, let T_β be the union of $\mathcal{V}'$ sentences occuring in T_γ for all $\gamma < \beta$. If T_β were to be inconsistent, we can derive a contradiction from it in a finite sequence of steps. Taking all the $\mathcal{V}'$-sentences that play a role in this proof, this finite set of sentences would be a subset of T_γ for some $\gamma < \beta$ (because each T is a subset of the next, and no sentence can be present in T_β which was not present in any T_γ). Also, the number of constants used up by T_β is less than $|\beta| \cdot (|\gamma| + |\mathbb{N}|)$ ¡ $|\beta| \cdot |\beta| = |\beta|$ for infinite ordinals β. Finally, let T be the union of this infinite sequence of T_βs.
It is easy to observe that properties 1 and 2 claimed above hold under our construction. Define the underlying set U of a model M defined on the vocabulary $\mathcal{V}'$. Let M interpret t_1 and t_2 as equal iff $T \models t_1 = t_2$, and let M interpret

[6] The set of all sentences is also of cardinatily α. To prove this, notice that sentences are finite. The number of ways of creating FOL sentences using k symbols for an appropriate ordinal k is strictly lesser than $\sum_{n=1}^{\infty} k^n$, up to countable terms. This is $|\mathbb{N}||k|$, which in turn is $|k|$

[7] This assumes that the Well-Ordering Principle holds.

relations and functions the same way as T does. M is the required model of $\mathcal{T}$[8].

Theorem 2. ***(Completeness)*** *For any sentence ψ and a theory T, every model of the theory T models $\{\psi\}$ iff ψ can be derived from T.*

Proof: At this point, it remains to be explained what a formal derivation is. A formal derivation is a finite sequence of steps by which you prove a claim from a premise (a set of sentences) using only established justifications. Some commonly applied rules can be summarized as follows (G is a formula and $\mathcal{F}$ is a set of formulas):

Premise	**Conclusion**	**Name**
G is in $\mathcal{F}$	$\mathcal{F} \vdash$ G	Assumption
$\mathcal{F} \vdash$ G and $\mathcal{F} \subset \mathcal{F}'$	$\mathcal{F}' \vdash$ G	Monotonicity
$\mathcal{F} \vdash$ G	$\mathcal{F} \vdash \neg\neg$ G	Double Negation
$\mathcal{F} \vdash$ F, $\mathcal{F} \vdash$ G	$\mathcal{F} \vdash$ (F$\wedge$G)	$\wedge$-Introduction
$\mathcal{F} \vdash$ (F$\wedge$G)	$\mathcal{F} \vdash$ F	$\wedge$-Elimination
$\mathcal{F} \vdash$ (F$\wedge$G)	$\mathcal{F} \vdash$ (F$\wedge$G)	$\wedge$-Symmetry
$\mathcal{F} \vdash$ F	$\mathcal{F} \vdash$ (F$\vee$G)	$\vee$-Introduction
$\mathcal{F} \vdash$ (F$\vee$G), $\mathcal{F} \cup \{F\} \vdash$ H, $\mathcal{F} \cup \{G\} \vdash$ H	$\mathcal{F} \vdash$ H	$\vee$-Elimination
$\mathcal{F} \vdash$ (F$\vee$G)	$\mathcal{F} \vdash$ (G$\vee$F)	$\vee$-Symmetry
$\mathcal{F} \cup \{F\} \vdash$ F	$\mathcal{F} \vdash$ (F$\rightarrow$ G)	$\rightarrow$-Introduction
$\mathcal{F} \vdash$ (F$\rightarrow$ G), $\mathcal{F} \vdash$ F	$\mathcal{F} \vdash$ G	$\rightarrow$-Elimination
$\mathcal{F} \vdash$ ($\neg$F$\wedge G$), $\mathcal{F} \vdash$ F	$\mathcal{F} \vdash$ G	$\wedge$-*Modus Ponens*
$\mathcal{F} \vdash$ ($\neg$F$\wedge G$), $\mathcal{F} \vdash \neg$G	$\mathcal{F} \vdash \neg$F	$\wedge$-*Modus Tollens*
$\mathcal{F} \vdash \psi(t)$	$\mathcal{F} \vdash \exists y \psi(y)$	$\exists$-Introduction
$\mathcal{F} \vdash \psi(c)$	$\mathcal{F} \vdash \forall y \psi(y)$	$\forall$-Introduction
$\mathcal{F} \vdash \psi \rightarrow \phi$	$\mathcal{F} \vdash \exists x \psi \rightarrow \exists x \phi$	$\exists$-Distribution
$\mathcal{F} \vdash \psi \rightarrow \phi$	$\mathcal{F} \vdash \forall x \psi \rightarrow \forall x \phi$	$\forall$-Distribution
	$\mathcal{F} \vdash t_1 = t_1$	Reflexivity of equality
$\mathcal{F} \vdash \psi(t_1)$, $\mathcal{F} \vdash t_1 = t_2$	$\mathcal{F} \vdash \psi(t_2)$	Equality substitution

Perhaps the only non-trivial rules to prove are $\exists$-Introduction, $\forall$-Introduction and the distributivity of $\exists$ and $\forall$. For $\exists$-Introduction, let us assume $\mathcal{F} \vdash \psi(t)$ for some t. Let M be a model of $\mathcal{F}$. $M \models \psi(t)$, from which it follows that $M \models \exists y \psi(y)$ by the semantics of $\exists$.
For $\forall$-Introduction, suppose that $\mathcal{T} \models \psi(c)$ where c is a constant that does not occur in $\mathcal{T}$ Let M be a model of $\mathcal{T}$. For any element of the underlying set U of M, let $M_{c=a}$ be the structure having U as the underlying set and interpreting c as a and interpreting every other symbol in the vocabulary of M the same way M does. Since c does not occur in $\mathcal{T}$, $M_{c=a}$ cannot contradict any formula in $\mathcal{T}$

[8]Provable using induction on the complexity of formulas in $\mathcal{T}$, using the fact that T is a complete theory

since M does not. Thus, it follows that $M_{c=a} \models \psi(c)$, hence $M \models \psi(a)$. Since a is an arbitrary element, this property holds for all a in U.
Now, consider $\exists$-Distribution. Let M be a model of T, we need to show that since $M \models \psi \to \phi$, $M \models \exists x\psi$ implies $M \models \exists x\phi$. We take the following cases:

1. If x is not a free variable of ψ. then the quantification of ψ using $\exists x$ only adds the condition that the Model has at least one element in its Universe (which is true by the definition of U anyway). Hence, ψ and $\exists x\psi$ are equivalent. $M \models \exists x\psi$ if $M \models \psi$. By implication, $M \models \phi$. If x is not a free variable in ϕ as well, then as we saw above, $M \models \exists x\phi$. If x is a free variable of ϕ then $M \models \phi$ for arbitrary values of x, hence obviously $M \models \exists x\phi$ if the Universe is non-empty.

2. If x is a free variable of ψ but not of ϕ. In this case, x is a free variable of $\psi \to \phi$. $M \models \psi \to \phi$ implies $M \models \forall x(\psi(x) \to \phi)$. Since $M \models \exists x\psi(x)$, $M \models \psi(a)$ for some a. Hence, $M \models \phi$, and this obviously means $M \models \exists x\phi$.

3. If x is a free variable of both ψ and ϕ, $M \models \psi \to \phi$ means $M \models \forall x(\psi(x) \to \phi(x)$. Since $M \models \exists x\psi(x)$, $M \models \psi(a)$ for some a. $M \models \phi(a)$ as well, by the above implication, and hence $M \models \exists x\phi(x)$.

The proof for $\forall$-Distribution is similar.
Example: Prove the Contrapositive rule. If p implies q, then ¬q implies ¬p.
Solution: Premise: $\mathcal{F} \cup \{F\} \vdash G$
Conclusion: $\mathcal{F} \cup \{\neg G\} \vdash \neg F$

Statement	**Justification**
$\mathcal{F} \cup \{F\} \vdash G$	Premise
$\mathcal{F} \vdash F \to \neg\neg G$	Double Negation
$\mathcal{F} \vdash F \to \neg\neg G$	$\to$-Introduction
$\mathcal{F} \vdash (\neg F \vee \neg\neg G)$	$\to$-Definition
$\mathcal{F} \vdash (\neg\neg G \vee \neg F)$	$\vee$-Symmetry
$\mathcal{F} \vdash (\neg G \to \neg F)$	$\to$-Definition
$\mathcal{F} \cup \{\neg G\} \vdash (\neg G \to \neg F)$	Monotonicity
$\mathcal{F} \cup \{\neg G\} \vdash \neg G$	Assumption
$\mathcal{F} \cup \{\neg G\} \vdash \neg F$	$\vee$-*Modus Ponens*

Example: Prove the Proof by cases rule.
Solution: Premise: $\mathcal{F} \cup \{F\} \vdash G$, $\mathcal{F} \cup \{\neg F\} \vdash G$
Conclusion: $\mathcal{F} \vdash G$

Statement	**Justification**
$\mathcal{F} \cup \{F\} \vdash G$	Premise
$\mathcal{F} \cup \{\neg F\} \vdash G$	Premise
$\mathcal{F} \vdash (F \vee \neg F)$	Tautology rule
$\mathcal{F} \vdash G$	$\vee$-Elimination

Example: If a contradiction can be derived from a set of sentences, then the set is inconsistent. Else, it is consistent. Let $\mathcal{F}$ be an inconsistent set. For each G $\in \mathcal{F}$, let $\mathcal{F}_G$ be the set obtained by removing G. Prove that $\mathcal{F}_G \vdash \neg G$.
Solution: First, we will prove a useful fact: If a contradiction can be derived from a set of sentences, then any formula can be derived from that set[9].

Statement	Justification
$\mathcal{F} \vdash (F \wedge \neg F)$	Premise
$\mathcal{F} \vdash F$	$\wedge$-Elimination
$\mathcal{F} \vdash (F \vee G)$	$\vee$-Introduction
$\mathcal{F} \vdash (\neg F \rightarrow G)$	$\rightarrow$-Definition
$\mathcal{F} \vdash (\neg F)$	$\wedge$-Elimination
$\mathcal{F} \vdash G$	$\vee$-*Modus Ponens*

Having proven this, we proceed with proving the main claim:

Statement	Justification
$\mathcal{F}_G \cup \{G\} \vdash (F \wedge \neg F)$	Definition of inconsistent
$\mathcal{F}_G \cup \{G\} \vdash \neg G$	By the above subclaim
$\mathcal{F}_G \cup \{\neg G\} \vdash \neg G$	Assumption
$\mathcal{F}_G \vdash \neg G$	Proof by Cases

Exercise: Prove or disprove using formal proofs:

1. p→(q∧r) ≡ (p→q)∧(q→r)
2. (p∧q)→r≡(p→r)∧(q→r)
3. ¬(p∧q)≡(¬p∨¬q)[10]

Solution:

1. Subproblem 1:

Statement	Justification
$\mathcal{F} \vdash p \rightarrow (q \wedge r)$	Premise
$\mathcal{F} \cup \{p\} \vdash (q \wedge r)$	$\rightarrow$-Definition
$\mathcal{F} \cup \{p\} \vdash q$	$\wedge$-Elimination
$\mathcal{F} \cup \{p\} \vdash r$	$\wedge$-Elimination
$\mathcal{F} \vdash p \rightarrow q$	$\rightarrow$-Definition
$\mathcal{F} \vdash p \rightarrow r$	$\rightarrow$-Definition
$\mathcal{F} \vdash (p \rightarrow q) \wedge (p \rightarrow r)$	$\wedge$-Introduction

[9] Bertrand Russell famously used a result similar to this to prove he was the Pope, when questioned by a student to derive the statement ”I am the Pope”, from the premise ”1=0”
[10] This is one half of de Morgan’s laws for propositional logic.

Proving the other direction for equivalence follows the same logic: use $\wedge$-Elimination to get two separate statements to prove, then combine them together using $\wedge$-Introduction and finally use $\rightarrow$-Definition.

2. Subproblem 2:
   This statement can be disproved. Let formula p hold and q and r not hold, then (p$\wedge$q)$\rightarrow$r holds, but (p$\rightarrow$r) does not hold.

3. Subproblem 3:

Statement	**Justification**
$\mathcal{F} \vdash \neg p \vee \neg q$	Premise
$\mathcal{F} \cup \{p \wedge q\} \vdash (p \wedge q)$	Assumption
$\mathcal{F} \cup \{p \wedge q\} \vdash p$	$\wedge$-Elimination
$\mathcal{F} \cup \{p \wedge q\} \vdash q$	$\wedge$-Elimination
$\mathcal{F} \cup \{\neg p\} \vdash \neg(p \wedge q)$	Contrapositive
$\mathcal{F} \cup \{\neg q\} \vdash \neg(p \wedge q)$	Contrapositive
$\mathcal{F} \vdash \neg(p \wedge q)$	$\vee$-Elimination

Statement	**Justification**
$\mathcal{F} \vdash \neg(p \wedge q)$	Premise
$\mathcal{F} \cup \{p\} \cup \{q\} \vdash p \wedge q$	Assumption applied twice and $\wedge$-Introduction
$\mathcal{F} \cup \{q\} \cup \{\neg(p \wedge q)\} \vdash \neg p$	Contrapositive
$\mathcal{F} \vdash \neg(p \wedge q) \vee \neg p$	$\vee$-Introduction
$\mathcal{F} \cup \{p \wedge q\} \vdash \neg p$	$\rightarrow$-Definition
$\mathcal{F} \cup \{p \wedge q\} \cup \{q\} \vdash \neg p$	Monotonicity
$\mathcal{F} \cup \{q\} \vdash \neg p$	Proof by cases
$\mathcal{F} \vdash \neg q \vee \neg p$	$\rightarrow$-Definition

Coming back to the completeness of FOL:
The completeness theorem states that a set of sentences $T \models \psi$ iff $T \vdash \psi$ (ψ can be derived from T by applying the above rules a finite number of times). To prove that $T \vdash \psi$ implies $T \models \psi$, it suffices to observe that all the rules we consider are *sound*. If $T \vdash \psi$, then there is a finite formal proof utilizing the rules we have derived above, concluding with $T \vdash \psi$. Each line contains a statement of the form $\mathcal{X} \vdash \mathcal{Y}$. We need to show for every line of the proof, if $\mathcal{X} \vdash \mathcal{Y}$, then $\mathcal{X} \models \mathcal{Y}$. This can be verified by proving that each rule is *sound*, either by constructing truth tables for simple rules such as $\vee$-introduction or by considering the semantics of the symbols like we did above for $\exists$-Introduction. For the other direction, suppose $T \models \psi$. That means that every model M of T models ψ. Thus, $T \cup \{\neg\psi\}$ has no model, and is thus not a theory. That is, $T \cup \{\neg\psi\} \vdash \bot$ for some contradiction $\bot$. By the contradiction rule, we have

$T \cup \{\neg\psi\} \vdash \psi$. By assumption, we have $T \cup \{\psi\} \vdash \psi$. By proof by cases, we have $T \vdash \psi$.

Theorem 3. ***Compactness*** *Let T be a set of sentences. Every finite subset of T is satisfiable if and only if T is satisfiable*[11].

Proof: Any model of T is a model of a finite subset of T, hence every finite subset of a satisfiable set is satisfiable. Suppose that T is unsat. This means we can derive a contradiction $\perp$ from T. Since formal derivations are finite, only a finite number of the formulas in T actually feature in this derivation. Hence, the set containing all these formulas which feature in the derivation is unsatisfiable.[12]

Example: Is it possible to generate a FOL theory that holds for connected graphs and connected graphs only?

Solution: No. Let $d_z(x, y)$ be the FOL sentence that states that vertices x and y in the graph are z apart $(\exists x_1 \cdots x_{z-1} \wedge R(x, x_1) \wedge R(x_{z-1}, y) \wedge \bigwedge_{i=1}^{z-1}(x \neq x_i) \wedge \bigwedge_{i=1}^{z-1}(y \neq x_i) \wedge \bigwedge_{i \neq j}(x_i \neq x_j))$.[13] If there was such a theory, add to it the sentences $d_i(a, b)$ for all natural numbers i and some constants a, b. Any finite subset of this is satisfiable (let $d_n(a, b)$ be the largest of the d_i's to be present in the subset, then this subset can be modelled by the chain $\{1, \cdots n+1\}$ with edges between adjacent numbers, a interpreted as 1 and b interpreted as n+1). Hence, by completeness, the theory is satisfiable, however, this means that a and b are not connected by any finite path, violating the connectedness of the graph.

Example: Define a sentence in the vocabulary of graphs that has arbitrarily large finite models, and all of its finite models are of even size.

Solution: The vocabulary of graphs has a relation R which we specify as anti-reflexive $(\forall x \neg R(x, x))$ and symmetric $(\forall x \forall y R(x, y) \rightarrow R(y, x))$. We then specify that each element has a unique neighbour it is paired with $\forall x \exists y (R(x, y) \wedge \forall z (R(x, z) \rightarrow z = y))$. Every finite model is a perfect matching and hence has

[11] For an example of a language that is not compact, consider the natural language English and the infinite set of sentences {"There are finitely many objects in this set.", "There are more than 1 objects in this set.", "There are more than 2 objects in this set.", $\cdots$}. Any finite subset of these sentences has a model, but the complete set cannot. Due to this, we have the result that FOL cannot describe a theory which has finite models of arbitrary size and finite models only.

[12] Compactness of propositional logic can be proved in another way: Using the Well-Ordering Principle, we enumerate all the sentences in an infinite set all of whose finite subsets are satisfiable. Since each of the first β sentences taken together is satisfiable where β is finite, we can have a sequence of assignments $< A_n >$ such that each A_i models the first i sentences. If we also enumerate all the atomic formulas in T by their order of appearance in $< A_n >$ (i.e all atomic formulas that appear in A_1, followed by those appearing in A_2 but not in A_1, etc.), and create the sequence $< w_n >$ of binary strings, where the i^{th} has 1 at all positions where A_i interprets an atomic formula as true and 0 at the other positions. It is an easily proven result that there exists an infinite binary string w such that each of its prefixes is also a prefix of infinitely many of these $< w_n >$. Let us then create the infinite assignment which interprets the i^{th} atomic formula as true if the i^{th} digit of w is 1 and false otherwise. Each of the infinitely many formulas has to hold under this assignment (Proof left to the reader).)

[13] Generate a recursive definition for these sentences using FOL and by defining either $d_0(x, y)$ or $d_1(x, y)$.

even size, and the conjunction of the above sentences is a sentence with arbitrarily large finite models.
Example: Does the above sentence stipulate that all models are of even size?
Solution: No. If there was a theory which had arbitrarily large even models and only even models, then we could add to it the sentences "This model has more than 0 elements", "This model has more than 2 elements", "This model has more than 4 elements" and so on (Each of these sentences is of the form $\exists x_1 \cdots \exists x_{2n+1} \bigwedge_{i \neq j} x_i \neq x_j$). Any finite subset of this new theory is satisfiable, hence the new theory is satisfiable. This means that the original theory could not have been modeled by models of even size and no other models.

5 Structures and Embeddings

Definition 5.1. Let $\mathcal{V}$ be a vocabulary and M and N be $\mathcal{V}$-structures. A function f:M→N preserves the formula ψ if, for each tuple $\widetilde{a}$, $M \models \psi(\widetilde{a})$ implies $N \models \psi(f(\widetilde{a}))$.

Definition 5.2. An embedding is a function $f : M \to N$ which preserves all atomic formulas and their negations.

Definition 5.3. An elementary embedding is a function $f : M \to N$ which preserves all formulas.

Note that a literal embedding is equivalent to the graph-theoretic concept of an induced subgraph. An elementary embedding means that two structures cannot be distinguished by First Order Logic alone. For instance, the rational numbers can be elementarily embedded into the real numbers, rendering them indistinguishable.

Corollary 3.1. *Embeddings preserve existential formulas (formulas of the form $\exists \widetilde{a} \psi(\widetilde{a})$ where $\widetilde{a}$ is a tuple and ψ does not contain any universal or existential quantifiers.)*

Corollary 3.2. *Let M and N be $\mathcal{V}$-structures. If the function $f : M \to N$ is onto, then it is a literal embedding iff it is an elementary embedding.*

Proof: Observe that f^{-1} is a one-one function from N onto M. We show that both f and f^{-1} preserve all $\mathcal{V}$-formulas. We prove this by induction on the complexity of ψ. If ψ is atomic, if $M \models \psi(\widetilde{a})$, $N \models \psi(f(\widetilde{a}))$ since f preserves literals (atomics and negations of atomics) and if $N \models \psi(f(\widetilde{a}))$, $M \models \psi(\widetilde{a})$ since f preserves $\neg\psi$. Similar arguments show that f and f^{-1} preserve $\neg\psi$ and $\psi \wedge \phi$ if f and f^{-1} preserve ψ and ϕ. It remains to be shown that $\exists x \phi$ is preserved if ϕ is preserved.
Let us consider $\psi(\widetilde{x}) = \exists \widetilde{x} \phi(\widetilde{x}, \widetilde{y})$. First, we show that f preserves ψ. Suppose that $M \models \phi(\widetilde{a}, \widetilde{b})$, then $N \models \phi(f(\widetilde{a}), f(\widetilde{b}))$. Conversely, let $N \models \phi(\widetilde{a}, \widetilde{b})$. Then, since f is onto, $f^{-1}(\widetilde{a}, \widetilde{b})$ is defined. Since $\neg\phi$ is preserved by f, $M \models \phi(f^{-1}(\widetilde{a}, \widetilde{b})$. Hence, by induction on complexity, we have proven the claim.

Corollary 3.3. *If M and N are two $\mathcal{V}$-structures which are isomorphic, they are elementarily equivalent.*

Proof: The isomorphism function $f : M \to N$ and its inverse preserve every $\mathcal{V}$-formula.

Remark. The converse need not hold. The structures $\mathbf{Q}_<$ defined as the rational numbers with the $\leq$ operation and $\mathbf{R}_<$ defined as the real numbers with the $\leq$ operations are elementarily equivalent, but one is countable and the other is not.[14]

Definition 5.4. M is a substructure of N, denoted by $M \subset N$, if they both have the same vocabulary, $U_M \subset U_N$ and M interprets the vocabulary the same way N does.

For instance, the structure with its Universe being the natural numbers and interpreting the $\leq$ operator ($\mathbf{N}_<$) is a substructure of the structure with integers as its Universe ($\mathbf{Z}_<$), which is itself a substructure of the rationals ($\mathbf{Q}_<$), which is a substructure of the model of real numbers ($\mathbf{R}_<$).

Example: Let N be the structure ($\mathcal{N}$—s) that has the natural numbers as its Universe and interprets s as the unary successor function. Only those subsets of $\mathcal{N}$ that are closed under s can serve as Universes of the substructures. Thus, the Universal sets of any substructure of N contain a lowest number a and all natural numbers which are its successor, successor's successor and so on. Hence, there are countably many substructures of N. If we add to the vocabulary the constant c, which N interprets as a number n, we see that any substructure must contain n and all natural numbers greater than n, hence we are limited to n substructures.

Definition 5.5. M is an elementary substructure of N if the identity function is an elementary embedding of M into N.

Definition 5.6. In the example above, for N=($\mathcal{N}$—s), the only elementary substructure is itself. From all of the other substructures of N, we can construct elementary embeddings to N, but the identity function is not this elementary embedding. For instance, let n be the lowest number in M. Let $\psi(x)$ state $\exists y(s(y) = x)$. N satisfies $\psi(n)$ but M does not, hence the identity is not an elementary embedding.

[14] We can also give examples of two structures of the same cardinality which are elementarily equivalent but not isomorphic. Consider the theory which states that a model is a dense linear order with endpoints, both of which are present in the Universe of the model. It can be proven that this theory is complete, since it is countably categorical and has only infinite models. If T were not complete, it would have models which are not elementarily equivalent. By the Lowenhiem-Skolem theorems, these models would both have elementary substructures of countable cardinality. But then these could not be isomporphic, violating the condition that T is countably categorical. Define a model $\mathbf{H}_{[0,2]}$ whose Universe is all the rational numbers in [0,1] and all the reals in [1,2]. This structure models the theory of dense linear orders, but is not isomorphic to the unit interval [0,1], which is also a model of this complete theory since there exist distinct numbers in $\mathbf{H}_{[0,2]}$ with only countably many elements between them, which is not the case in $\mathbf{R}_{[0,1]}$.

Exercise: How many substructures does the complete graph on n vertices K_n have? How many does it have up to isomorphism? How many are elementary?

Exercise: Let N be the structure ($\mathcal{N}$—S) defined on natural numbers with the successor relation. Prove that N has uncountably many non-isomorphic substructures.

Solution: First, observe that any non-empty subset of the natural numbers is a valid Universe for a substructure of N. Secondly, note that these Universal sets can be defined up to isomorphism by the lengths of their consecutive sequences of numbers present in the Universe (since the Successor relation is capable of defining the width of an interval all of whose numbers are in the Universe, but is incapable of expressing in FOL the distance between two numbers if there are numbers not part of the Universe in between them). The number of ways of generating distinct sequences of lengths is uncountable (easy to prove).

Exercise: Using the result that existential formulas are preserved under extensions, derive the result that universal formulas are preserved under substructures.

Exercise: Let $M \subset N$ denote M is a substructure of N and $M \prec N$ denote M is an elementary substructure of N. Given $A \subset B \subset C$, either prove or give counterexamples for:

1. If $A \prec B$ and $B \prec C$, then $A \prec C$.
2. If $A \prec C$ and $B \prec C$, then $A \prec B$.
3. If $A \prec B$ and $A \prec C$, then $B \prec C$.

Solution: Consider the Theory of Dense Linear Orders. It is a complete theory, since it is countably categorical and has only infinite models. Thus, any two models of this theory are elementarily equivalent. Let C be $\mathbf{R}_<$ and let A and B be any suitable substructures of $\mathbf{Q}_<$ which satisfy the theory but neither of whose Universe is a subset of the other's. For instance, the Universes can be the sets of the dyadic and tryadic rationals, or $Q/\{0\}$ and $Q/\{1\}$, etc.

Theorem 4. ***Tarski-Vaught Criterion***
Let M and N be $\mathcal{V}$-structures with $N \subset M$. Suppose that for any $\mathcal{V}$-formula $\psi(\widetilde{x}, y)$ and any tuple $\widetilde{a}$ of elements from U_N, the following holds:

$$M \models \exists y \psi(\widetilde{a}, y) implies N \models \exists y \psi(\widetilde{a}, y)$$

Then $N \prec M$.

Proof: We prove that for every $\mathcal{V}$ formula ψ and tuple $\widetilde{a}$ in U_N,

$$N \models \psi(\widetilde{a}) iff M \models \psi(\widetilde{a})$$

This is achieved by induction on the complexity of ψ. It is true for atomic ψ by the definition of substructures, and quite obviously the proposition is true

for the negation and conjunction of formulas to which the proposition applies. Now, suppose $\psi(\widetilde{x})$ has the form $\exists y \phi(\widetilde{x}, y)$.
Suppose that $N \models \exists y \phi(\widetilde{a}, y)$. Then for some b, $N \models \phi(\widetilde{a}, b)$. $M \models \phi(\widetilde{a}, b)$. $M \models \exists y \phi(\widetilde{a}, y)$
The reverse direction of this claim is simply the condition stipulated in the proposition. Thus, we are done with the proof.[15]

6 Lowenhiem-Skolem

Theorem 5. ***Upward Lowenhiem-Skolem Theorem***
If a theory has infinite models, it has infinite models of arbitrarily large cardinality.

Proof: Let M be an infinite model of T. Let k be any cardinal. Let us expand the vocabulary $\mathcal{V}$ of M to a new vocabulary $\mathcal{V}'$ by including a set C of k constants which do not occur in $\mathcal{V}$. Let $\mathcal{T}$ be the set of all $\mathcal{V}'$ sentences having the form $\neg(a = b)$, where a and b are constants from C. Any $\mathcal{V}'$-structure modelling $\mathcal{T}$ has to have at least k elements in its Universe.
We note that $T \cup \mathcal{T}$ is satisfiable because any finite subset of $T \cup \mathcal{T}$ will contain only finitely many constants from C, and is modeled by the infinite structure M', which interprets every element of C as distinct elements of U_M and otherwise interprets $\mathcal{V}$ as M does. By compactness, $T \cup \mathcal{T}$ is satisfiable and its models have a size greater than or equal to the arbitrary cardinal k.

Theorem 6. ***Downward Lowenhiem-Skolem***
Let M be a structure having vocabulary $\mathcal{V}$ and Universe U_M. For any $X \subset U_M$, there exists an elementary substructure N of M such that:

1. *X is a subset of U_N*
2. *$|N| \leq |X| + ||\mathcal{V}||$, where $||\mathcal{V}||$ is the cardinality of the number of sentences in the vocabulary $\mathcal{V}$.*

Proof: We define a sequence $X_1 \subset X_2 \subset \cdots$ of subsets of U_M. Let $X_1 = X$. Now suppose X_m has been defined. Suppose that $|X_m| \leq |X| + ||\mathcal{V}||$. Let $\mathcal{V}_{\Updownarrow}$ be the expansion of $\mathcal{V}$ obtained by adding new constants for each element of X_m. Let M_m be the natural expansion of M to this vocabulary. Let E_m be the set of all V_m sentences of the form $\exists x \psi(x)$ which are modeled by M_m.
By the Well-Ordering Principle, E_m can be enumerated as $\{\exists x \phi_\beta(x) | \beta < \alpha\}$ for some ordinal α. For each $\beta < \alpha$, there exists an element a_β such that $M \models \phi_\beta(a_\beta)$. If we enumerate all these elements in a set A, $|A| \leq |E_m| \leq ||\mathcal{V}||$. Let $X_{m+1} = X_m \cup A$. We have $|X_{m+1} \leq |X_m| + |A| \leq (|X| + ||\mathcal{V}||) + ||\mathcal{V}|| =$

[15]This condition can be strengthened, and the strengthened version is what is used in the proof of the Downward Lowenhiem-Skolem Theorem below: If U is a subset of U_M for a $\mathcal{V}$-structure M, and for any $\mathcal{V}$-formula $\psi(\widetilde{x}, y$ and any tuple $\widetilde{a}$ of elements from U, if $M \models \exists y \psi(\widetilde{a}, y)$, then $M \models \psi(\widetilde{a}, b)$ for some $b \in U$. Then, U is the underlying set of an elementary subset of M.

$|X| + ||\mathcal{V}||$.
Now, let $U' = \bigcup_{i=1}^{\infty} X_i$.
$|U'| \leq |N| \cdot (|X| + ||\mathcal{V}||) = |X| + ||\mathcal{V}||$.
Furthermore, if $M \models \exists y \psi(\widetilde{a}, y)$ for some tuple $\widetilde{a}$ in U', then $M \models \psi(\widetilde{a}, b)$ for some b in U'. This can be proven by considering the fact that there must be some finite m such that X_m contains all of the finite number of elements in the tuple $\widetilde{a}$. It follows that $\exists y \psi(\widetilde{a}, y)$ is in E_m. By the definition of X_{m+1}, $M \models \psi(\widetilde{a}, b)$ for some b in X_{m+1}, and $X_{m+1} \subset U'$. By the Tarski-Vaught criterion, we have shown that U' is the underlying set of an elementary substructure N of M.
Exercise: Let T be a theory with an infinite model. Prove that T has a model of size k for each infinite cardinal k with $k \geq |T|$.
Solution: Apply the Upward Lowenhiem-Skolem to get that T has a model of cardinality $\geq k$. Let M be this model Now consider an arbitrary subset of U_M of cardinality k. By the Downward Lowenhiem-Skolem theorem, there exists a model N of this theory with size at most $k + ||\mathcal{V}||$. Since each sentence in T is finite, there can be at most $|T|$ symbols in $\mathcal{V}$ if $|T|$ is infinite, and a finite number of symbols otherwise. $||\mathcal{V}||$ can have at most $|\mathcal{V}|$ sentences since it is a countable union of sets of size $|\mathcal{V}|$ if $|\mathcal{V}|$, and is a countable union of finite sets otherwise.
Exercise: Let T be an incomplete countable theory. For each of the following, either prove or give a counterexample:

1. If T has an uncountable model, then T has a countable model.
2. If T has arbitrarily large finite models, then T has a countable infinite model.
3. If T has finite models and a countably infinite model, then T has arbitrarily large finite models.

Solution: The first result arises directly from applying the Downward Lowenhiem-Skolem theorem. The second result arises from compactness, which states that T has an infinite model, followed by the Downward Lowenhiem-Skolem theorem, which would imply T has a countably infinite model. A counter-theory can be proposed for the third result: one example would be the theory with the following sentences:

$$\forall x \forall y \forall z (x > y) \wedge (y > x) \to (x > z)$$

$$\exists x \forall y (y > x) \wedge (y > y \to y = x)$$

$$\forall x \forall y (x \neq y) \wedge (x > y) \to \neg(y > x)$$

$$\forall x \forall y (x \neq y) \wedge (y > x) \to \neg(x > y)$$

$$\forall x (f(x) > x)$$

$$\exists x (x = x)$$

Exercise: Let T_1 be a complete V_1-theory and T_2 be a complete V_2 theory. Show that $T_1 \cup T_2$ is consistent iff $\psi_1 \wedge \psi_2$ is satisfiable for every pair $\psi_1 \in T_1$ and $\psi_2 \in T_2$.
Hint: Consider the compactness theorem and the fact that complete theories are closed under conjunction[17].
Exercise: Let ψ be a FOL sentence that is not contained in any complete theory. Show that $\{\psi\} \vdash \neg\psi$.
Solution: Let us assume $\{\psi\}$ is satisfiable. Then, it has a model M. We have proven earlier that the theory of a model $Th(M)$ is complete. Thus, $Th(M) \models \neg\psi$. Thus, $M \models \neg\psi$ since every model of $Th(M)$ models $\neg\psi$. Hence, every model of $\{\psi\}$ models $\neg\psi$. By definition of $\models$, $\{\psi\} \models \neg\psi$. By the completeness theorem, $\{\psi\} \vdash \neg\psi$.
Exercise: The four color theorem states that any planar graph is 4-colorable[18]. Here, 4-colorable means that every vertex can be assigned a color such that no two adjacent vertices have the same color. Assuming this result is true for finite graphs, show that it is true for infinite graphs.
Exercise: Show that any partial order on a set A (finite or infinite) can be extended to a linear order on A[19].
Exercise: Let T be the set of all sentences in the vocabulary $\mathcal{V}_< = \{<\}$ that hold in every well-ordered set. Show that there exists a model M of T that does not interpret ¡ as a well-ordering on its Universe.
Exercise: Let $\mathcal{V}$ be a vocabulary that contains only constants and no functions or relations. Let M and N be two infinite $\mathcal{V}$-structures. Using the Tarski-Vaught Criterion, show that if $M \subset N$, $M \prec N$.

7 Proof Theory (Optional)

Checking for the satisfiability, validity or unsatisfiability of statements is a primary goal of Logic. Resolution (to be explained later) is a very useful method of checking for the satisfiability of sentences. It is used in languages like Prolog.

Definition 7.1. Prenex Normal Form: A formula ψ is in Prenex Normal Form (PNF) if it has the form $Q_1 x_1 \cdots Q_n x_n \phi$ where ϕ is a quantifier-free formula and all Q_is are quantifiers ($\exists$ or $\forall$).

[16]WLOG let us denote the element which has (x¿x) as 0. If f maps zero to itself, we can have a model with only one element (0) or 0 mapping to 0 and an infinite increasing chain of elements. If f maps zero to any other elements, we have a infinite chain of elements (maybe with one branch). Either way, it has finite models and no arbitrarily large finite models despite having a countable model.

[17]In fact, they are deductively closed, a stronger result than closure under conjunction alone.

[18]A planar graph is one that can be drawn on the Euclidean Plane with no intersecting edges

[19]A partial order defines ¡ as transitive, and stipulates that for all a, b in A, at most one of the following hold: a ¡ b, b ¡ a or a=b

Definition 7.2. Conjunctive Prenex Normal Form: A formula ψ is in CPNF if it is PNF and the quantifier-free formula ϕ is the conjunction of disjunction of literals.

We call each set of disjunctions between conjunctions a clause.
We define an algorithm to generate an equivalent formula in CPNF for any FOL formula[20]. First, we show that there is an algorithm to convert an arbitrary formula ψ to a PNF formula ψ', and then that any quantifier-free formula can be converted to CNF.
If ψ is atomic, we let $\psi' = \psi$.
Now, let ω and ϕ be formulas which are equivalent to the PNF formulas ω' and ϕ'. Suppose $\psi = \neg\phi$. Since ϕ' is in PNF, it is of the form $Q_1 x_1 \cdots Q_n x_n \phi_0$. Set $\psi' = \widetilde{Q}_1 x_1 \cdots \widetilde{Q}_n x_n \neg\phi_0$ where $\widetilde{Q}$ is $\exists$ if Q is $\forall$ and vice versa. This formula is in PNF.
Next, suppose $\psi = \phi \wedge \omega$. Let $\phi' = Q_1 x_1 \cdots Q_n x_n \phi_0(x_1, \cdots, x_n)$ and $\omega' = Q_1 x_1 \cdots Q_m x_m \omega_0(x_1, \cdots, x_m)$. Since replacing the x_i's with new variables does not affect the formula, let us replace $(x_1, \cdots, x_n)$ with $(y_1, \cdots, y_n)$ and $(x_1, \cdots, x_m)$ with $(z_1, \cdots, z_m)$, where the y_i's and z_i's do not occur in ϕ_0 and ω_0.
Now, set $\psi' = Q_1 y_1 \cdots Q_n y_n Q_1 z_1 \cdots Q_n z_m (\phi_0(y_1, \cdots, y_n) \wedge \omega_0(z_1, \cdots, z_m))$.
Lastly, if $\psi = \exists x \phi$, set $\psi' = \exists x \phi'$ and we are done.
Now, we prove that every quantifier-free formula is equivalent to a formula in CNF and a formula in DNF[21]. Again, this is by induction on the complexity of the formula. The result is trivial for atomic formulas. Let ϕ and ω have ϕ_1 and ω_1' as equivalent CNF formulas and ϕ_2 and ω_2 as equivalent DNF formulas.
If $\psi = \neg\phi$, $\psi \equiv \neg\phi_2$. The negation of a formula in DNF is a formula in CNF (easily proved). Similarly, $\psi \equiv \neg\phi_1$ and the negation of a formula in CNF is a formula in DNF. Hence, we can get an equivalent formula for ψ in CNF and DNF.
Suppose $\psi = \phi \wedge \omega$. $\phi_1 \wedge \omega_1$ is an equivalent formula in CNF. It only remains to be shown that there is an equivalent formula in DNF. Since each of ψ_2, ω_2 is in DNF, they can be written as

$$\psi_2 = \bigvee_i C_{1i}, \omega_2 = \bigvee_i C_{2i}$$

for clauses from the sets $< C_1 i >$ and $< C_2 i >$. We have

$$\psi = \left(\bigvee_i C_{1i} \right) \wedge \left(\bigvee_j C_{2j} \right)$$

It is easy to show this is actually

$$\psi = \bigvee_i \left(\bigvee_j (C_{1i} \wedge C_{2j}) \right)$$

[20]Note that this can include exponential blowup.
[21]Disjunctive Normal Form, where each clause is a conjunction of literals and each formula is a disjunction of clauses

using the distributivity of $\wedge$ over $\vee$.

7.1 Skolem Normal Form

Definition 7.3. A sentence is in Skolem Normal Form (SNF) if it is Universal and in CPNF.

Given any sentence in CPNF, we can convert it to SNF.
If the formula in CPNF is already universal or quantifier-free, we have nothing to do.
If not, let $\psi = \forall x_1 \cdots \forall x_{i-1} \exists x_i Q_{i+1} x_{i+1} \cdots Q_n x_n \phi_0(x_1, \cdots, x_n)$. Let ψ' be the sentence $\psi' = \forall x_1 \cdots \forall x_{i-1} Q_{i+1} x_{i+1} \cdots Q_n x_n \phi_0(x_1, \cdots, x_{i-1}, f(x_1, \cdots, x_{i-1}), x_{i+1}, \cdots, x_n)$ where f is an (i-1)-ary arbitrary function (constant function if i=1) which does not occur in ϕ_0.
Since ψ' has fewer existential quantifiers than ψ, by repeated application of this algorithm, we'll finally get a Universal sentence with is CPNF.

7.2 Herbrand Theory

We reduce sentences in FOL to sentences in Propositional Logic, and then apply a simple method called resolution on these sentences to check their unsatisfiability.

Definition 7.4. Let $\mathcal{V}$ be a vocabulary. The Herbrand universe H of $\mathcal{V}$ is the set of all variable-free $\mathcal{V}$ terms.

For instance, if $\mathcal{V}$ contains a constant c and a unary function f, its Herbrand Universe is $\{c, f(c), f(f(c)), \cdots\}$.

Definition 7.5. Any $\mathcal{V}$-structure that has H as its Universe and interprets the constants in $\mathcal{V}$ as the elements they correspond to in H, and similarly interprets the functions in $\mathcal{V}$ just as H does, is a Herbrand structure of $\mathcal{V}$.

Note that the Herbrand Universe is empty if $\mathcal{V}$ has no constants, is finite if $\mathcal{V}$ has no functions and is unique if $\mathcal{V}$ has no relations.

Definition 7.6. Let T be a set of sentences. The Herbrand vocabulary of T, denoted $\mathcal{V}_T$ is defined as follows: Let $\mathcal{V}_T$ be the set of all functions, relations and constants in T. If it has no constants, add an arbitrary constant c to the vocabulary. Create the corresponding Herbrand universe $H(T)$ and get a model M which has universe $H(T)$ and models T.

Consider the sentence $\forall x((f(x) \neq x) \wedge (f(f(x)) = x))$. The Herbrand Universe is $\{c, f(c), f(f(c)), \cdots\}$. In any Herbrand model, $f(f(c))$ and c are thus distinct elements of the Universe and the sentence is unsatisfiable. But in reality, the FOL sentence is satisfied.

Theorem 7. *Given any set T of sentences in SNF which do not contain the equality operator $=$, the set is a theory*[22] *iff it has a Herbrand model.*

Proof: The forward direction is trivial.
Conversely, suppose T is satisfiable. Let $\mathcal{V}_T$ be the Herbrand vocabulary, let N be a $\mathcal{V}_T$-structure that models T, and let M' be an arbitrary Herbrand $\mathcal{V}_T$-structure.
We define a $\mathcal{V}_T$-structure which has the Universe H_T and interprets functions and constants the same way as N does. i.e $M \models R(t_1, \cdots, t_n)$ iff $N \models R(t_1, \cdots, t_n)$ for all t_i in the Universe of M. This can be defined since each t_i is a variable-free $\mathcal{V}_T$ term, and hence is interpreted by N as a particular element of U_N (N interprets all constants and functions of terms as elements of its Universe).
We make and prove two claims:

1. For any sentence which is both quantifier-free and equality-free, M models the sentence iff N models it.

2. For any SNF sentence that is equality-free, M models the sentence if N models it.

If claim 2 is true, then M must model T. That is because each sentence in T is SNF and equality-free, and is modeled by N. Hence, M would become a Herbrand model of T.
Proof of Claim 1:
Let ψ be quantifier-free. If ψ is atomic, since it does not use "="and is a sentence, it must be of the form $R(t_1, \cdots, t_n)$ where each term is variable-free. That means that each term is in H by definition of the Herbrand universe H. By the definition of M, $M \models \psi$ iff $N \models \psi$. The rest of the proof is trivial by induction on the complexity of ψ.
Proof of Claim 2: We prove this by induction on the number of quantifiers in ψ. If ψ has no quantifiers, then by Claim 1, $M \models \psi$ iff $N \models \psi$.
Suppose ψ is $\forall x_1 \cdots \forall x_m \phi(x_1, \cdots, x_m)$. Our induction hypothesis holds for the formula $\psi' = \forall x_2 \cdots \forall x_m \phi(x_2, \cdots, x_m)$ obtained by removing the first quantifier from ψ. Let t be any variable-free V-term. Then, $N \models \psi$ implies $N \models \psi'(t)$ (because ψ is a Universal formula in terms of x_1). By our induction hypothesis, $M \models \psi'(t)$ since ψ' has n-1 universal quantifiers.
But t was an arbitrary variable-free $\mathcal{V}$-term (i.e a term present in H). Since H is the Universe of M, $M \models \forall x_1 \psi'(x_1)$, hence $M \models \psi$. Our proof is thus complete.

7.3 Dealing with Eauality

Now suppose ψ has the symbol "=". We replace every instance of $t_1 = t_2$ with $E(t_1, t_2)$ and proceed to add to ψ the following conditions on E. First, we define

[22] Recall that all theories are satisfiable and have models

E to be equivalence:

$$\forall x E(x,x) \wedge \forall x \forall y (E(x,y) \rightarrow E(y,x)) \wedge \forall x \forall y \forall z (E(x,y) \wedge E(y,z) \rightarrow E(x,z))$$

For each relation R in the vocabulary of ψ (note that for finite ψ, its vocabulary is obviously finite), we have the sentence:

$$\forall x_1 \cdots \forall x_n \forall y_1 \cdots \forall y_n (\left(\bigwedge_{i=1}^{n} E(x_i, y_i) \wedge R(x_1, \cdots x_n) \right) \rightarrow R(y_1, \cdots y_n))$$

Likewise, for each function in the vocabulary, we have the sentence:

$$\forall x_1 \cdots \forall x_n \forall y_1 \cdots \forall y_n (\left(\bigwedge_{i=1}^{n} E(x_i, y_i) \right) \rightarrow E(f(x_1, \cdots x_n), f(y_1, \cdots y_n)))$$

Now let us take the conjunction of all these formulas and the original sentence to get ψ_E.
Claim: For any formula ψ in SNF, ψ is satisfiable iff ψ_E is satisfiable.
Proof: If $M \models \psi$ we can get a model for ψ_E by interpreting E as equality. Conversely, suppose ψ_E has a model N. Then E is an equivalent relation on the Universal set U of N. We define a $\mathcal{V}$-structure N_E having the set of all E-equivalence sets as an underlying class, where $\mathcal{V}$ is the vocabulary of N. For each a, denote the E-equivalence class containing a as $[a]$.
For any n-tuple of Equivalence classes and an n-ary relation R, $N_E \models R([a_1], \cdots, [a_n])$ iff $N \models R(a_1, \cdots, a_n)$.
For any n-ary function f, $N_E \models f([a_1], \cdots, [a_n]) = [b]$ iff $N \models f(a_1, \cdots, a_n) = b$.
It can be trivially shown that N_E is a model of ψ_E by induction on the complexity of formulas.

7.4 The Herbrand Method

Now, we formally describe how to algorithmically find if a sentence of FOL is unsatisfiable. We assume the sentence is equality-free and in SNF. Let the sentence thus be $\forall x_1 \cdots \forall x_n \psi(x_1, \cdots x_n)$.
Let E be the set

$$\{psi(t_1, \cdots t_n)\} | t_1, \cdots t_n \in H$$

So, E is the set obtained by substituting terms from the Herbrand Universe H for the variables of ψ in every possible way. If M models the original sentence, M should model each ψ_i in E and hence E would be satisfiable.
Thus, if E is unsatisfiable, we have concluded that the original FOL sentence is unsatisfiable. Since each sentence in E has no quantifiers or equality, it can be viewed as a set of sentences of propositional logic, wherein we will define Resolution as a method of determining unsatisfiability.

8 The Unification Algorithm (Optional)

Definition 8.1. 'Resolution' refers to the argument $(\psi \vee \phi) \wedge (\neg\phi \vee \omega) \equiv (\psi \vee \omega)$

Resolution can in finite time tell us if a sentence is unsatisfiable, but need not necessarily tell us anything about whether it is satisfiable. Nevertheless, it is a very useful method of proof.

8.0.1 Proving Completeness of Resolution for Propositional Logic

Propositional logic does not have quantifiers, constants or formulas.
We show that resolution can be used to determine whether or not a given propositional logic formula is satisfiable. WLOG, assume the formula ψ is in CNF. Let $Res^0(\psi)$ be $\{C|C$ is a clause of $\psi\}$. For each $n > 0$, let $Res^n(F) = Res^{n-1}(F) \cup \{R|R$ is a resolvent of two clauses in $Res^{n-1}(F)\}$. Since $Res^0(\psi)$ is a finite set, only finitely many clauses can be derived from $Res^0(\psi)$ using resolvents. In fact, if F has n atomic formulas, there are at most 2^{2n} clauses possible using these atomic formulas, hence the resolution process is finite. Eventually, we will find some m such that $Res^m(F) = Res^{m+1}(F)$. Let $Res^*(F)$ denote $Res^m(F)$.

Corollary 7.1. *F is unsatisfiable iff $\emptyset \in Res^*(F)$ for a propositional logic formula F.*

Proof: If $\emptyset \in Res^*(F)$, then $\emptyset \in Res^n(F)$ for some n. Since $\emptyset \notin Res^0(F)$, there exists some n such that $\{A\} \in Res^n(F)$ and $\{\neg A\} \in Res^n(F)$ for some atomic formula A. Thus, both A and $\neg A$ are consequences of F[23]. $F \vdash (A \wedge \neg A)$. Hence, F is unsatisfiable.
Now, to prove the converse. We will prove this proposition by induction on the number of atomic formulas that occur in F.
Let us consider the base case n=1. Let A be the only atomic formula in F. Then each clause in F is either $\{A\}$ or $\{\neg A\}$[24]. If F is unsatisfiable, F is $\{A\} \wedge \{\neg A\}$, whose resolvent is $\emptyset$.
Now suppose F has atomic formulas $A_1, \cdots, A_{n+1}$. Suppose further that $\emptyset \in Res^*(G)$ for any formula G that uses only $A_1, \cdots, A_n$.
Let $\widetilde{F_0}$ be the conjunction of all clauses in F that do not contain $\neg A_{n+1}$.
Let $\widetilde{F_1}$ be the conjunction of all clauses in F that do not contain A_{n+1}.
Obviously, all clauses in F are contained in either or both of these formulas. Let $F_0 = \{C_i - \{A_{n+1}|C_i \in \widetilde{F_0}\}\}$ and $F_1 = \{C_i - \{\neg A_{n+1}|C_i \in \widetilde{F_1}\}\}$. That is, F_0 is formed by assuming A_{n+1} is a contradiction and F_1 is formed by assuming A_{n+1} is a tautology. Since A_{n+1} must either have a truth value of 1 or 0, it follows

[23] To prove this, consider proof by cases. If $F \vdash (\psi \vee \phi) \wedge (\neg\psi \vee \omega)$, we want to prove $F \vdash (\phi \vee \omega)$. By assumption and $\vee$-Modus Ponens, we can prove $F \cup \{\neg\psi\} \vdash \phi$ and $F \cup \{\psi\} \vdash \omega$. By $\vee$-Introduction and proof by cases, the desired result follows.

[24] We can assume WLOG that no clause is a tautology, like $A \vee \neg A$, because we can drop those clauses.

that F is equisatisfiable with $F_0 \vee F_1$. If F is unsatisfiable, F_0 and F_1 are both unsatisfiable. By our induction hypothesis, $\emptyset \in Res^*(F_0)$ and $\emptyset \in Res^*(F_1)$.
Now, since F_0 was formed from $\widetilde{F_0}$ by dropping A_{n+1} from each clause, we can derive either $\emptyset$ or $\{A_{n+1}\}$ from $\widetilde{F_0}$. Similarly, we can derive either $\emptyset$ or $\{\neg A_{n+1}\}$ from $\widetilde{F_1}$. If we can derive $\emptyset$ from either of the two, we are done. If not, we can still derive $\emptyset$ by resolution of A from $\widetilde{F_0} \cup \widetilde{F_1}$. Since $F = \widetilde{F_0} \cup \widetilde{F_1}$, we conclude that $\emptyset \in Res^*(F)$.
The completeness of resolution for infinite sets of sentences holds by the compactness of FOL. If the infinite set is unsatisfiable, then some finite subset must be unsatisfiable, and hence $\emptyset$ must be derivable from it using resolution.

8.1 Unification for FOL

Let $\mathcal{L}$ be a set of literals. We say $\mathcal{L}$ is unifiable if there exist variables $x_1, \cdots, x_m$ and terms $t_1, \cdots, t_m$ such that substituting t_i for x_i makes each literal in $\mathcal{L}$ look the same. For any sentence ψ in SNF, we denote the result of applying this substitution to be ψsub.
For example, if $sub = (x/w, y/f(a), z/f(w))$ and $\psi = \{\neg Q(x,y), R(a,w,z)\}$, then $\psi sub = \{\neg Q(w, f(a)), R(a, w, f(w))\}$[25].
Let $\mathcal{L}$ be a set of literals, then $\mathcal{L}sub$ denotes the set of all $L_i sub$ such that $L_i \in \mathcal{L}$. So $\mathcal{L}$ is unifiable iff there exists a substitution sub such that $\mathcal{L}sub$ has only one literal. In such a case, we call sub a unifier for $\mathcal{L}$ and say sub unifies $\mathcal{L}$.

Definition 8.2. Let $\mathcal{L}$ be a set of literals. We call sub a most general unifier for $\mathcal{L}$ if it unifies $\mathcal{L}$ and for any other unifier sub', we have $subsub' = sub'$.

Considering the most general unifier is often most useful for practical applications.
Example: Let $\mathcal{L} = \{P(f(x), y), P(f(a), w)\}$. Let $sub_1 = (x/a, y/w)$ and $sub_2 = (x/a, y/a, w/a)$ $\mathcal{L}sub_1 = \{P(f(a), w)\}$ and $\mathcal{L}sub_2 = \{P(f(a), a)\}$. If we have $sub_3 = (w/a)$, note that $\mathcal{L}sub_1 sub_3 = \mathcal{L}sub_2$.

Unification Algorithm:
Given a finite set of literals $\mathcal{L}$.
Let $\mathcal{L}_0$=$\mathcal{L}$ and $sub_0 = \emptyset$. Suppose we know $\mathcal{L}_k$ and sub_k. If $\mathcal{L}_k$ contains only one literal, $sub_0 sub_1 \cdots sub_k$ is the most general unifier for $\mathcal{L}$.
Otherwise there exists L_i and L_j in $\mathcal{L}$ such that their n$^{\text{th}}$ symbol is the first symbol at which they differ. If the n$^{\text{th}}$ symbol of L_i is a variable v and the n$^{\text{th}}$ symbol of L_j is the first symbol of a term t that does not contain v or vice versa, then let $sub_{k+1} = (v/t)$ and $\mathcal{L}_{k+1} = \mathcal{L}_k sub_{k+1}$. If any of the hypotheses above does not hold, we conclude that $\mathcal{L}$ is not unifiable.

Proof that this Algorithm works:

[25] Note that all substitutions take place simultaneously, not sequentially

If the algorithm concludes that $\mathcal{L}$ is not unifiable, it is for one of two reasons. One is when you have a discrepancy between two literals where the discrepancy does not involve variables. For instance $\{f(a), f(b)\}$ is not unifiable, since constants cannot be substituted out. The other possibility is when a variable and term are involved, and the variable is included in the term. For instance, $\{x, f(x)\}$. No matter what you substitute for x, the second literal will have one more instance of f than the first. Hence, both reasons for concluding unsatisfiability are valid reasons.
Now, when applied on a finite $\mathcal{L}$, only finitely many variables occur in $\mathcal{L}$. Since each variable can be substituted once, this algorithm terminates in a finite number of steps. Now, we have to show that the output is actually a most general unifier. Now, let sub' be any other unifier for $\mathcal{L}$. We know that $sub_0 sub' = sub'$ since sub_0 is empty. Now, suppose $sub_0 \cdots sub_m sub' = sub'$ for some m, $0 \leq m < k$, where sub_k is the last unifier in the output. Then $\mathcal{L} sub_s ub_0 \cdots sub_m sub' = \mathcal{L} sub' = \{L\}$. Suppose sub_{m+1} in the Unification algorithm is (x/t). Since sub' takes $\mathcal{L}_m$ to $\{L\}$, it must be able to deal with x and t. That is, $xsub' = tsub'$. It follows that $sub_{m+1} sub' = (x/t) sub'$. By induction, we have $sub_0 \cdots sub_{m+1} sub' = sub'$ for all m ¡ k. In particular, the result is provably the most general unifier for $\mathcal{L}$.
Resolution for FOL:
Once a sentence is converted to Skolem Normal form, ψ is of the form $\forall x_1 \cdots \forall x_m \phi(x_1, \cdots, x_m)$. Let $C(\psi)$ denote $C(\phi)$, where $C(\phi) = \{C_1, \cdots C_n\}$, where each C_i is the set of all literals occuring in the i^{th} disjunction.
Let C_1 and C_2 be two clauses with no two variables in common. Let $L_1, \cdots, L_m \in C_1$ $L'_1, \cdots, L'_n \in C_2$ be such that $\mathcal{L} = \{\neg L_1 \cdots \neg L_m, L'_1 \cdots L'_n\}$ is unifiable. Let sub be the most general unifier given by the Unification algorithm.
Then $R = [(C_1 - \{L_1, \cdots L_m\}) \cup (C_2 - \{L'_1, \cdots L'_n\})]sub$ is a resolvent of C_1 and C_2.
Let ψ be a sentence in SNF. Then $\psi = \{C_1, \cdots C_n\}$ for some clauses whose conjunction it is. Let $Res(\phi)$ be the set of resolvents of all pairs of clauses in ϕ if possible. Let $Res^0(\psi) = \psi$ and $Res^{n+1}(\psi) = Res(Res^n(\psi))$. Let $Res'(\psi) = \bigcup_n Res^n(\psi)$[26].
Example: Prove that the sentence $\forall x(P(x) \vee Q(x)) \to \exists x P(x) \vee \forall x Q(x)$ is a tautology.
Solution: We prove that the negation of the same is unsatisfiable. Expressed as clauses after Skolemization, the negation of this statement is $\{P(x), Q(x)\}, \{\neg P(x)\}, \{\neg Q(c)\}$. We apply Resolution on clauses 1 and 2 without any substitutions to get $\{Q(x)\}, \{\neg Q(c)\}$. Now, run the substitution (x/c) and apply resolution again to get $\emptyset$. Hence, the negation of this sentence is unsatisfiable, and the sentence is valid.
Exercise: Derive a contradiction from the following SNF sentence: $\forall x_1 \forall x_2 \forall x_3 \forall x_4 f(x_1, x_3, x_2) \neq f(g(x_2), j(x_4), h(x_3, a))$.
Hint: Let $sub_1 = (x_1/g(x_2))$, $sub_2 = (x_3/j(x_4))$, $sub_3 = (x_2/h(j(x_4), a))$.
Exercise: Find most general unifiers of the following:

[26] Note that unlike in Propositional Logic, $Res'(\psi)$ can be infinite.

1. $f(g(x_1), h(x_2), x_4)$ and $f(g(k(x_2, x_3)), x_3, h(x_1))$
2. $f(x, y, z)$ and $f(y, z, x)$
3. $f(g(x), x)$ and $f(y, g(y))$

Exercise: Give unifiers for the first two problems from the preceding exercise which are not most general unifiers.
Solution:

1. $(x_1/k(a, h(a)), x_3/h(a), x_4/h(k(a, h(a)))), (x_2/a)$
2. $(x/a, y/a, z/a)$

We conclude this document with two results that make Resolution much simpler to run algorithmically, and together comprise the SLD algorithm:

8.2 L and N Resolution

Definition 8.3. N-resolution requires that one of the two parents contains only negative literals.

Definition 8.4. Linear resolution requires that one of the parents be the resolvent from the previous step, for all steps except the first.

Theorem 8. *Let ψ be a sentence in SNF. It is unsatisfiable iff $\emptyset$ can be derived from ψ using N-resolution.*

Theorem 9. *Let ψ be a sentence in SNF. It is unsatisfiable iff $\emptyset$ can be derived from ψ using linear resolution. In fact, if $\mathcal{F}$ is a set of sentences in propositional logic that are in CNF such that any strict subset is satisfiable but $\mathcal{F}$ is unsatisfiable[27], then for any clause in $\mathcal{F}$, $\emptyset$ can be derived from $\mathcal{F}$ starting with that clause as one parent and using only linear resolution.*

[27] Note that compactness implies the finiteness of $\mathcal{F}$