<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Propositional Logic</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Gentzen’s Logical Calculi</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>Ordered Binary Decision Diagrams</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>First-Order Logic</td>
<td>21</td>
</tr>
<tr>
<td>6</td>
<td>Formal Reasoning in First-Order Logic</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>Davis-Putnam &amp; Propositional Resolution</td>
<td>31</td>
</tr>
<tr>
<td>8</td>
<td>Skolem Functions and Herbrand’s Theorem</td>
<td>36</td>
</tr>
<tr>
<td>9</td>
<td>Unification</td>
<td>41</td>
</tr>
<tr>
<td>10</td>
<td>Resolution and Prolog</td>
<td>46</td>
</tr>
<tr>
<td>11</td>
<td>Modal Logics</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>Tableaux-Based Methods</td>
<td>56</td>
</tr>
</tbody>
</table>
Logic concerns statements in some language.

The language can be informal (e.g. English) or formal.

Some statements are true, others false or perhaps meaningless, . . .

Logic concerns relationships between statements: consistency, entailment, . . .

Logical proofs model human reasoning.

**Statements**

Statements are declarative assertions:

\[ \text{Black is the colour of my true love's hair.} \]

They are not greetings, questions, commands, . . .:

\[ \text{What is the colour of my true love's hair?} \]
\[ \text{I wish my true love had hair.} \]
\[ \text{Get a haircut!} \]
Schematic Statements

The *meta-variables* $X, Y, Z, \ldots$ range over ‘real’ objects

- Black is the colour of $X$’s hair.
- Black is the colour of $Y$.
- $Z$ is the colour of $Y$.

Schematic statements can express general statements, or questions:

*What things are black?*

Interpretations and Validity

An *interpretation* maps meta-variables to real objects

The interpretation $Y \mapsto$ coal *satisfies* the statement

- Black is the colour of $Y$.

but the interpretation $Y \mapsto$ strawberries does not!

A statement $\Lambda$ is *valid* if all interpretations satisfy $\Lambda$. 
Consistency, or Satisfiability

A set $S$ of statements is *consistent* if some interpretation satisfies all elements of $S$ at the same time. Otherwise $S$ is *inconsistent*.

Examples of inconsistent sets:

- $\{X \text{ part of } Y, \ Y \text{ part of } Z, \ X \text{ NOT part of } Z\}$
- $\{n \text{ is a positive integer, } n \neq 1, \ n \neq 2, \ldots\}$

satisfiable/unsatisfiable = consistent/inconsistent

Entailment, or Logical Consequence

A set $S$ of statements *entails* $A$ if every interpretation that satisfies all elements of $S$, also satisfies $A$. We write $S \models A$.

- $\{X \text{ part of } Y, \ Y \text{ part of } Z\} \models X \text{ part of } Z$
- $\{n \neq 1, \ n \neq 2, \ldots\} \models n \text{ is NOT a positive integer}$

$S \models A$ if and only if $\{\neg A\} \cup S$ is inconsistent

$\models A$ if and only if $A$ is valid
Inference

Want to check $A$ is valid

Checking all interpretations can be effective — but if there are infinitely many?

Let $\{A_1, \ldots, A_n\} \models B$. If $A_1, \ldots, A_n$ are true then $B$ must be true. Write this as the inference

$$
\frac{A_1 \ldots A_n}{B}
$$

Use inferences to construct finite proofs!

Schematic Inference Rules

$$
\frac{X \text{ part of } Y \quad Y \text{ part of } Z}{X \text{ part of } Z}
$$

A valid inference:

$$
\frac{\text{spoke part of wheel} \quad \text{wheel part of bike}}{\text{spoke part of bike}}
$$

An inference may be valid even if the premises are false!

$$
\frac{\text{cow part of chair} \quad \text{chair part of ant}}{\text{cow part of ant}}
$$
Survey of Formal Logics

propositional logic is traditional boolean algebra.

first-order logic can say for all and there exists.

higher-order logic reasons about sets and functions. It has been applied to hardware verification.

modal/temporal logics reason about what must, or may, happen.

type theories support constructive mathematics.

Why Should the Language be Formal?

Consider this ‘definition’:

The least integer not definable using eight words

Not equal to The number of atoms in the entire Universe

Also not equal to The least integer not definable using eight words

• A formal language prevents AMBIGUITY.
Syntax of Propositional Logic

P, Q, R, ... propositional letter

t true
f false
\neg A not A
A \land B A and B
A \lor B A or B
A \rightarrow B if A then B
A \leftrightarrow B A if and only if B

Semantics of Propositional Logic

\neg, \land, \lor, \rightarrow and \leftrightarrow are truth-functional: functions of their operands

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>\neg A</th>
<th>A \land B</th>
<th>A \lor B</th>
<th>A \rightarrow B</th>
<th>A \leftrightarrow B</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>t</td>
<td>f</td>
<td>f</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>f</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
</tbody>
</table>
Interpretations of Propositional Logic

An interpretation is a function from the propositional letters to \{t, f\}.

Interpretation I satisfies a formula \( \Lambda \) if the formula evaluates to \( t \).

Write \( \models I \Lambda \)

\( \Lambda \) is valid (a tautology) if every interpretation satisfies \( \Lambda \)

Write \( \models \Lambda \)

\( S \) is satisfiable if some interpretation satisfies every formula in \( S \)

Implication, Entailment, Equivalence

\( \Lambda \to B \) means simply \( \neg \Lambda \vee B \)

\( \Lambda \models B \) means if \( \models I \Lambda \) then \( \models I B \) for every interpretation \( I \)

\( \Lambda \models B \) if and only if \( \models \Lambda \to B \)

Equivalence

\( \Lambda \simeq B \) means \( \Lambda \models B \) and \( B \models \Lambda \)

\( \Lambda \simeq B \) if and only if \( \models \Lambda \leftrightarrow B \)
**Equivalences**

\[
\begin{align*}
A \land A &\simeq A \\
A \land B &\simeq B \land A \\
(A \land B) \land C &\simeq A \land (B \land C) \\
A \lor (B \land C) &\simeq (A \lor B) \land (A \lor C) \\
A \land \mathbf{f} &\simeq \mathbf{f} \\
A \land \mathbf{t} &\simeq A \\
A \land \neg A &\simeq \mathbf{f}
\end{align*}
\]

Dual versions: exchange $\land$, $\lor$ and $\mathbf{t}$, $\mathbf{f}$ in any equivalence.

**Negation Normal Form**

1. Get rid of $\leftrightarrow$ and $\rightarrow$, leaving just $\land$, $\lor$, $\neg$:

\[
\begin{align*}
A \leftrightarrow B &\simeq (A \rightarrow B) \land (B \rightarrow A) \\
A \rightarrow B &\simeq \neg A \lor B
\end{align*}
\]

2. Push negations in, using de Morgan’s laws:

\[
\begin{align*}
\neg \neg A &\simeq A \\
\neg (A \land B) &\simeq \neg A \lor \neg B \\
\neg (A \lor B) &\simeq \neg A \land \neg B
\end{align*}
\]
From NNF to Conjunctive Normal Form

3. Push disjunctions in, using distributive laws:

\[ A \lor (B \land C) \cong (A \lor B) \land (A \lor C) \]
\[ (B \land C) \lor A \cong (B \lor A) \land (C \lor A) \]

4. Simplify:
   - Delete any disjunction containing \( P \) and \( \neg P \)
   - Delete any disjunction that includes another
   - Replace \( (P \lor A) \land (\neg P \lor A) \) by \( A \)

Converting a Non-Tautology to CNF

\[ P \lor Q \rightarrow Q \lor R \]

1. Elim \( \rightarrow \): \( \neg(P \lor Q) \lor (Q \lor R) \)
2. Push \( \neg \) in: \( (\neg P \land \neg Q) \lor (Q \lor R) \)
3. Push \( \lor \) in: \( (\neg P \lor Q \lor R) \land (\neg Q \lor Q \lor R) \)
4. Simplify: \( \neg P \lor Q \lor R \)

Not a tautology: try \( P \mapsto t \), \( Q \mapsto f \), \( R \mapsto f \)
Tautology checking using CNF

\[(P \rightarrow Q) \rightarrow P\]

1. Elim $\rightarrow$: $\neg[\neg(-P \lor Q) \lor P] \lor P$

2. Push $\neg$ in: $[\neg(-P \lor Q) \land \neg P] \lor P$

3. Push $\lor$ in: $(-P \lor Q \lor P) \land (-P \lor P)$

4. Simplify: $t \land t$

\[t\]  *It's a tautology!*
A Simple Proof System

Axiom Schemes

K  \[ A \to (B \to A) \]
S  \[ (A \to (B \to C)) \to ((A \to B) \to (A \to C)) \]
DN \[ \neg \neg A \to A \]

Inference Rule: Modus Ponens

\[
\frac{A \to B \quad A}{B}
\]

A Simple (?) Proof of \(A \to A\)

\[
(A \to ((D \to A) \to A)) \to
\]

\[
((A \to (D \to A)) \to (A \to A)) \quad \text{by S}
\]

\[
A \to ((D \to A) \to A) \quad \text{by K}
\]

\[
(A \to (D \to A)) \to (A \to A) \quad \text{by MP, (1), (2)}
\]

\[
A \to (D \to A) \quad \text{by K}
\]

\[
A \to A \quad \text{by MP, (3), (4)}
\]
Some Facts about Deducibility

\( A \) is deducible from the set \( S \) if there is a finite proof of \( A \) starting from elements of \( S \). Write \( S \vdash A \).

**Soundness Theorem.** If \( S \vdash A \) then \( S \models A \).

**Completeness Theorem.** If \( S \models A \) then \( S \vdash A \).

**Deduction Theorem.** If \( S \cup \{A\} \vdash B \) then \( S \vdash A \rightarrow B \).

Gentzen’s Natural Deduction Systems

A varying context of assumptions

Each logical connective defined independently

*Introduction* rule for \( \land \): how to deduce \( A \land B \)

\[
\frac{A \quad B}{A \land B}
\]

*Elimination* rules for \( \land \): what to deduce from \( A \land B \)

\[
\frac{A \land B}{A} \quad \frac{A \land B}{B}
\]
The Sequent Calculus

Sequent $A_1, \ldots, A_m \Rightarrow B_1, \ldots, B_n$ means,

if $A_1 \land \ldots \land A_m$ then $B_1 \lor \ldots \lor B_n$

$A_1, \ldots, A_m$ are assumptions; $B_1, \ldots, B_n$ are goals

$\Gamma$ and $\Delta$ are sets in $\Gamma \Rightarrow \Delta$

$A, \Gamma \Rightarrow A, \Delta$ is trivially true (basic sequent)

Sequent Calculus Rules

\[
\begin{align*}
\frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta} & \quad (\text{cut}) \\
\frac{\Gamma \Rightarrow \Delta, A}{\neg A, \Gamma \Rightarrow \Delta} & \quad (\neg t) \\
\frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta, \neg A} & \quad (\neg r) \\
\frac{A, B, \Gamma \Rightarrow \Delta}{A \land B, \Gamma \Rightarrow \Delta} & \quad (\land l) \\
\frac{\Gamma \Rightarrow \Delta, A \quad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \land B} & \quad (\land r)
\end{align*}
\]
**More Sequent Calculus Rules**

\[
\frac{A, \Gamma \Rightarrow \Delta \quad B, \Gamma \Rightarrow \Delta}{A \lor B, \Gamma \Rightarrow \Delta} \quad (\lor l)
\]

\[
\frac{\Gamma \Rightarrow \Delta, A \land B}{\Gamma \Rightarrow \Delta} \quad (\lor r)
\]

\[
\frac{\Gamma \Rightarrow \Delta, A \quad B, \Gamma \Rightarrow \Delta}{A \Rightarrow B, \Gamma \Rightarrow \Delta} \quad (\rightarrow l)
\]

\[
\frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \rightarrow B} \quad (\rightarrow r)
\]

---

**Easy Sequent Calculus Proofs**

\[
\frac{A, B \Rightarrow A}{A \land B \Rightarrow A} \quad (\land l)
\]

\[
\frac{A \land B \Rightarrow A}{\Rightarrow A \land B \rightarrow A} \quad (\rightarrow r)
\]

\[
\frac{A, B \Rightarrow B, A}{A \Rightarrow B, B \rightarrow A} \quad (\rightarrow r)
\]

\[
\frac{A \Rightarrow B, B \rightarrow A}{\Rightarrow A \rightarrow B, B \rightarrow A} \quad (\lor r)
\]

\[
\frac{\Rightarrow (A \rightarrow B) \lor (B \rightarrow A)}{}
\]
Part of a Distributive Law

\[
\begin{align*}
A &\Rightarrow A, B \\ B, C &\Rightarrow A, B
\end{align*}
\]

\[
\begin{align*}
A &\Rightarrow A, B \\ B \land C &\Rightarrow A, B
\end{align*}
\]

\[
\begin{align*}
A &\lor (B \land C) \Rightarrow A, B \\ B \land C &\Rightarrow A, B
\end{align*}
\]

\[
\begin{align*}
A &\lor (B \land C) \Rightarrow A \lor B \\ A \lor (B \land C) &\Rightarrow (A \lor B) \land (A \lor C)
\end{align*}
\]

Similar

Second subtree proves \( A \lor (B \land C) \Rightarrow A \lor C \) similarly

A Failed Proof

\[
\begin{align*}
A &\Rightarrow B, C \\ B &\Rightarrow B, C
\end{align*}
\]

\[
A \lor B \Rightarrow B, C
\]

\[
\begin{align*}
A \lor B &\Rightarrow B \lor C \\ A \lor B &\Rightarrow B \lor C
\end{align*}
\]

\[
\Rightarrow A \lor B \rightarrow B \lor C
\]

\( A \leftrightarrow t, B \leftrightarrow f, C \leftrightarrow f \) falsifies unproved sequent!
Ordered Binary Decision Diagrams

Canonical form: essentially decision trees with sharing

- ordered propositional symbols ('variables')
- sharing of identical subtrees
- hashing and other optimisations

Detects if a formula is tautological (t) or inconsistent (f)

A FAST way of verifying digital circuits, ...
Converting a Decision Diagram to an OBDD

No duplicates

No redundant tests

Building OBDDs Efficiently

Do not construct full tree! (see Bryant, §3.1)

Do not expand $\rightarrow, \leftrightarrow, \oplus$ (exclusive OR) to other connectives

Treat $\neg Z$ as $Z \rightarrow \overline{f}$ or $Z \oplus t$

Recursively convert operands

Combine operand OBDDs — respecting ordering and sharing

Delete test if it proves to be redundant
**Canonical Form Algorithm**

To do $Z \land Z'$, where $Z$ and $Z'$ are already canonical:

*Trivial if either is $t$ or $f$. Treat $\lor, \rightarrow, \leftrightarrow$ similarly!*

Let $Z = \text{if}(P, X, Y)$ and $Z' = \text{if}(P', X', Y')$

If $P = P'$ then recursively do $\text{if}(P, X \land X', Y \land Y')$

If $P < P'$ then recursively do $\text{if}(P, X \land Z', Y \land Z')$

If $P > P'$ then recursively do $\text{if}(P', Z \land X', Z \land Y')$

---

**Canonical Form of $P \lor Q$**

![Diagram of $P \lor Q$]

Slide 406
**Canonical Form of** $P \lor Q \rightarrow Q \lor R$

Slide 407

**Optimisations Based On Hash Tables**

Never build the same OBDD twice: share pointers

- Pointer identity: $X = Y$ whenever $X \leftrightarrow Y$
- Fast removal of redundant tests by $\text{if}(P, X, X) \simeq X$
- Fast processing of $X \land X$, $X \lor X$, $X \rightarrow X$, …

Never process $X \land Y$ twice; keep table of canonical forms

Slide 408
The variable ordering is crucial. Consider

\[(P_1 \land Q_1) \lor \cdots \lor (P_n \land Q_n)\]

A good ordering is \(P_1 < Q_1 < \cdots < P_n < Q_n\)

A dreadful ordering is \(P_1 < \cdots < P_n < Q_1 < \cdots < Q_n\)

Many digital circuits have small OBDDs (not multiplication!)

OBDDs can solve problems in hundreds of variables

General case remains intractable!
Outline of First-Order Logic

Reasons about functions and relations over a set of individuals

\[
\text{father(father}(x)) = \text{father(father}(y)) \\
\text{cousin}(x, y)
\]

Reasons about all and some individuals:

All men are mortal  \\
Socrates is a man  \\
Socrates is mortal

Does not reason about all functions or all relations, . . .

Function Symbols; Terms

Each function symbol stands for an \( n \)-place function

A constant symbol is a 0-place function symbol

A variable ranges over all individuals

A term is a variable, constant or has the form

\[ f(t_1, \ldots, t_n) \]

where \( f \) is an \( n \)-place function symbol and \( t_1, \ldots, t_n \) are terms

We choose the language, adopting any desired function symbols
Each relation symbol stands for an \( n \)-place relation.

Equality is the 2-place relation symbol \( \equiv \).

An atomic formula has the form

\[
R(t_1, \ldots, t_n)
\]

where \( R \) is an \( n \)-place relation symbol and \( t_1, \ldots, t_n \) are terms.

A formula is built up from atomic formulæ using \( \neg, \land, \lor, \ldots \).

(Later we add quantifiers)

Very expressive, given strong induction rules.

Prove equivalence of mathematical functions:

\[
\begin{align*}
p(z, 0) &= 1 & q(z, 1) &= z \\
p(z, n + 1) &= p(z, n) \times z & q(z, 2 \times n) &= q(z \times z, n) \\
& & q(z, 2 \times n + 1) &= q(z \times z, n) \times z
\end{align*}
\]

Boyer/Moore Theorem Prover: checked Gödel's Theorem, \ldots

Many systems based on equational reasoning.
Universal and Existential Quantifiers

- $\forall x \ A$ for all $x$, $A$ holds
- $\exists x \ A$ there exists $x$ such that $A$ holds

Syntactic variations:

- $\forall x y z \ A$ abbreviates $\forall x \forall y \forall z \ A$
- $\forall z . A \land B$ is an alternative to $\forall z (A \land B)$

The variable $x$ is bound in $\forall x \ A$; compare with $\int f(x) \, dx$

Expressiveness of Quantifiers

All men are mortal:

$$\forall x \ (\text{man}(x) \rightarrow \text{mortal}(x))$$

All mothers are female:

$$\forall x \ \text{female}(\text{mother}(x))$$

There exists a unique $x$ such that $A$, written $\exists! x \ A$

$$\exists x \ [A(x) \land \forall y \ (A(y) \rightarrow y = x)]$$
How do we interpret \text{mortal}(\text{Socrates})?

Interpretation $\mathcal{I} = (D, I)$ of our first-order language

$D$ is a non-empty set, called the \textit{domain or universe}.

$I$ maps symbols to ‘real’ functions, relations

- $c$ a constant symbol: $I[c] \in D$
- $f$ an $n$-place function symbol: $I[f] \in D^n \rightarrow D$
- $P$ an $n$-place relation symbol: $I[P] \subseteq D^n$

How do we interpret \text{cousin}(\text{Charles}, y)??

A \textit{valuation} supplies the values of free variables.

It is a function $V: \text{variables} \rightarrow D$

$\mathcal{I}_V[t]$ extends $V$ to a term $t$ by the obvious recursion:

\[
\begin{align*}
\mathcal{I}_V[x] & \overset{\text{def}}{=} V(x) \quad \text{if } x \text{ is a variable} \\
\mathcal{I}_V[c] & \overset{\text{def}}{=} I[c] \\
\mathcal{I}_V[f(t_1, \ldots, t_n)] & \overset{\text{def}}{=} I[f](\mathcal{I}_V[t_1], \ldots, \mathcal{I}_V[t_n])
\end{align*}
\]
The Meaning of Truth — in FOL

For interpretation $I$ and valuation $V$

$$\models_I, V P(t) \quad \text{if } I_V[t] \in I[P] \text{ holds}$$

$$\models_I, V t = u \quad \text{if } I_V[t] \text{ equals } I_V[u]$$

$$\models_I, V A \land B \quad \text{if } \models_I, V A \text{ and } \models_I, V B$$

$$\models_I, V \exists x A \quad \text{if } \models_I, V[m/x] A \text{ holds for some } m \in D$$

$$\models_I A \quad \text{if } \models_I, V A \text{ holds for all } V$$

$A$ is satisfiable if $\models_I A$ for some $I$
**Free vs Bound Variables**

All occurrences of $x$ in $\forall x \ A$ and $\exists x \ A$ are **bound**

An occurrence of $x$ is **free** if it is not bound:

$$\forall y \exists z \ R(y, z, f(y, x))$$

May **rename** bound variables:

$$\forall w \exists z' \ R(w, z', f(w, x))$$

---

**Substitution for Free Variables**

$A[t/x]$ means substitute $t$ for $x$ in $A$:

- $(B \land C)[t/x]$ is $B[t/x] \land C[t/x]$
- $(\forall x \ B)[t/x]$ is $\forall x \ B$
- $(\forall y \ B)[t/x]$ is $\forall y B[t/x]$ (if $x \neq y$)
- $(P(u))[t/x]$ is $P(u[t/x])$

No variable in $t$ may be bound in $A$!

$(\forall y \ x = y)[y/x]$ is not $\forall y \ y = y$!
Some Equivalences for Quantifiers

\[\neg (\forall x A) \simeq \exists x \neg A\]
\[(\forall x A) \land B \simeq \forall x (A \land B)\]
\[(\forall x A) \lor B \simeq \forall x (A \lor B)\]
\[(\forall x A) \land (\forall x B) \simeq \forall x (A \land B)\]
\[(\forall x A) \rightarrow B \simeq \exists x (A \rightarrow B)\]
\[\forall x A \simeq \forall x A \land A[t/x]\]

Dual versions: exchange \(\forall, \exists\) and \(\land, \lor\)

Reasoning by Equivalences

\[\exists x (x = a \land P(x)) \simeq \exists x (x = a \land P(a))\]
\[\simeq \exists x (x = a) \land P(a)\]
\[\simeq P(a)\]
\[\exists z(P(z) \rightarrow P(a) \land P(b))\]
\[\simeq \forall z P(z) \rightarrow P(a) \land P(b)\]
\[\simeq \forall z P(z) \land P(a) \land P(b) \rightarrow P(a) \land P(b)\]
\[\simeq t\]
Sequent Calculus Rules for $\forall$

\[
\frac{A[t/x], \Gamma \Rightarrow \Delta}{\forall x A, \Gamma \Rightarrow \Delta} \quad (\forall l) \quad \frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta, \forall x A} \quad (\forall r)
\]

Rule $(\forall l)$ can create many instances of $\forall x A$.

Rule $(\forall r)$ holds provided $x$ is not free in the conclusion!

Not allowed to prove:

\[
\frac{P(y) \Rightarrow P(y)}{P(y) \Rightarrow \forall y P(y)} \quad (\forall r)
\]

Simple Example of the $\forall$ Rules

\[
\frac{P(f(y)) \Rightarrow P(f(y))}{\forall x P(x) \Rightarrow P(f(y))} \quad (\forall l)
\]

\[
\frac{\forall x P(x) \Rightarrow P(f(y))}{\forall x P(x) \Rightarrow \forall y P(f(y))} \quad (\forall r)
\]
Not-So-Simple Example of the $\forall$ Rules

\[
\frac{P \Rightarrow Q(y), P}{P, P \rightarrow Q(y) \Rightarrow Q(y)} \quad \frac{P, Q(y) \Rightarrow Q(y)}{P, P \rightarrow Q(y) \Rightarrow Q(y)} \quad (\rightarrow 1)
\]
\[
\frac{P, \forall x (P \Rightarrow Q(x)) \Rightarrow Q(y)}{P, \forall x (P \Rightarrow Q(x)) \Rightarrow \forall y Q(y)} \quad (\forall 1)
\]
\[
\frac{P, \forall x (P \Rightarrow Q(x)) \Rightarrow \forall y Q(y)}{\forall x (P \Rightarrow Q(x)) \Rightarrow P \rightarrow \forall y Q(y)} \quad (\rightarrow r)
\]

In $(\forall 1)$ we have replaced $x$ by $y$

Sequent Calculus Rules for $\exists$

\[
\frac{\Lambda, \Gamma \Rightarrow \Delta}{\exists x \Lambda, \Gamma \Rightarrow \Delta} \quad (\exists 1)
\]
\[
\frac{\Gamma \Rightarrow \Delta, \Lambda[t/x]}{\Gamma \Rightarrow \Delta, \exists x \Lambda} \quad (\exists r)
\]

Rule $(\exists 1)$ holds provided $x$ is not free in the conclusion!

Rule $(\exists r)$ can create many instances of $\exists x \Lambda$

Say, to prove

$\exists z (P(z) \rightarrow P(a) \land P(b))$
Part of the existential Distributive Law

\[
\begin{align*}
P(x) \Rightarrow P(x), Q(x) & \quad (\forall r) \\
P(x) \Rightarrow P(x) \lor Q(x) & \quad (\forall r) \\
P(x) \Rightarrow \exists y (P(y) \lor Q(y)) & \quad (\exists r) \\
\exists x P(x) \Rightarrow \exists y (P(y) \lor Q(y)) & \quad (\exists l) \\
\exists x P(x) \lor \exists x Q(x) & \Rightarrow \exists y (P(y) \lor Q(y)) \\
\end{align*}
\]

Second subtree proves \( \exists Q(x) \Rightarrow \exists y (P(y) \lor Q(y)) \) similarly.

In (\( \exists r \)) we have replaced \( y \) by \( x \).

---

A Failed Proof

\[
\begin{align*}
P(x), Q(y) \Rightarrow P(x) \land Q(x) & \quad (\exists r) \\
P(x), Q(y) \Rightarrow \exists z (P(z) \land Q(z)) & \quad (\exists l) \\
P(x), \exists x Q(x) \Rightarrow \exists z (P(z) \land Q(z)) & \quad (\exists l) \\
\exists x P(x), \exists x Q(x) \Rightarrow \exists z (P(z) \land Q(z)) & \quad (\exists l) \\
\exists x P(x) \land \exists x Q(x) \Rightarrow \exists z (P(z) \land Q(z)) & \quad (\land l)
\end{align*}
\]

We cannot use (\( \exists l \)) twice with the same variable.

We rename the bound variable in \( \exists Q(x) \) and get \( \exists y Q(y) \).
Clause Form

Clause: a disjunction of literals

$$\neg K_1 \lor \cdots \lor \neg K_m \lor L_1 \lor \cdots \lor L_n$$

Set notation: $$\{\neg K_1, \ldots, \neg K_m, L_1, \ldots, L_n\}$$

Kowalski notation: $$K_1, \cdots, K_m \rightarrow L_1, \cdots, L_n$$

$$L_1, \cdots, L_n \leftarrow K_1, \cdots, K_m$$

Empty clause: $$\square$$

EMPTY CLAUSE MEANS CONTRADICTION!

Outline of Clause Form Methods

To prove $$A$$, obtain a contradiction from $$\neg A$$:

1. Translate $$\neg A$$ into CNF as $$A_1 \land \cdots \land A_m$$
2. This is the set of clauses $$A_1, \ldots, A_m$$
3. Transform the clause set, preserving consistency

Empty clause refutes $$\neg A$$

Empty clause set means $$\neg A$$ is satisfiable
The Davis-Putnam-Logeman-Loveland Method

1. Delete tautological clauses: \{P, \neg P, \ldots \}

2. For each unit clause \{L\},
   - delete all clauses containing L
   - delete \neg L from all clauses

3. Delete all clauses containing pure literals

4. Perform a case split on some literal

It’s a decision procedure: it finds either a contradiction or a model.

---

Davis-Putnam on a Non-Tautology

Consider \( P \lor Q \rightarrow Q \lor R \)

Clauses are \{P, Q\} \{\neg Q\} \{\neg R\}

\{P, Q\} \{\neg Q\} \{\neg R\} initial clauses
\{P\} \{\neg R\} unit \neg Q
\{\neg R\} unit P (also pure)
unit \neg R (also pure)

Clauses satisfiable by \( P \leftrightarrow t, \ Q \leftrightarrow f, \ R \leftrightarrow f \)
**Example of a Case Split on $P$**

\[
\begin{align*}
\neg Q, R & \quad \neg R, P & \quad \neg R, Q & \quad \neg P, Q, R & \quad P, Q & \quad \neg P, \neg Q \\
\neg Q, R & \quad \neg R, Q & \quad Q, R & \quad \neg Q & \quad \text{if } P \text{ is true} \\
\neg R & \quad R & \quad \text{unit } \neg Q \\
\square & \quad \text{unit } R \\
\neg Q, R & \quad \neg R & \quad \neg R, Q & \quad Q & \quad \text{if } P \text{ is false} \\
\neg Q & \quad Q & \quad \text{unit } \neg R \\
\square & \quad \text{unit } \neg Q
\end{align*}
\]

**The Resolution Rule**

From $B \lor A$ and $\neg B \lor C$ infer $A \lor C$

In set notation,

\[
\frac{\{B, A_1, \ldots, A_m\} \quad \neg B, C_1, \ldots, C_n}{\{A_1, \ldots, A_m, C_1, \ldots, C_n\}}
\]

Some special cases:

\[
\frac{\{B\} \quad \neg B, C_1, \ldots, C_n}{\{C_1, \ldots, C_n\}} \quad \frac{\{B\} \quad \neg B}{\square}
\]
Simple Example: Proving \( P \wedge Q \rightarrow Q \wedge P \)

Hint: use \( \neg(A \rightarrow B) \cong A \wedge \neg B \)

1. Negate! \( \neg[P \wedge Q \rightarrow Q \wedge P] \)
2. Push \( \neg \) in: \( (P \wedge Q) \land \neg(Q \wedge P) \)
   \( (P \wedge Q) \land (\neg Q \lor \neg P) \)

Clauses: \{P\} \{Q\} \{\neg Q, \neg P\}

Resolve \{P\} and \{\neg Q, \neg P\} getting \{\neg Q\}

Resolve \{Q\} and \{\neg Q\} getting \Box

Another Example

Refute \( \neg[(P \lor Q) \land (P \lor R) \rightarrow P \lor (Q \land R)] \)

From \( (P \lor Q) \land (P \lor R) \), get clauses \{P, Q\} and \{P, R\}

From \( \neg[P \lor (Q \land R)] \) get clauses \{\neg P\} and \{\neg Q, \neg R\}

Resolve \{\neg P\} and \{P, Q\} getting \{Q\}

Resolve \{\neg P\} and \{P, R\} getting \{R\}

Resolve \{Q\} and \{\neg Q, \neg R\} getting \{\neg R\}

Resolve \{R\} and \{\neg R\} getting \Box
The Saturation Algorithm

At start, all clauses are passive. None are active.

1. Transfer a clause (current) from passive to active.
2. Form all resolvents between current and an active clause.
3. Use new clauses to simplify both passive and active.
4. Put the new clauses into passive.

Repeat until CONTRADICTION found or passive becomes empty.

Refinements of Resolution

Preprocessing: removing tautologies, symmetries . . .
Set of Support: working from the goal
Weighting: priority to the smallest clauses
Subsumption: deleting redundant clauses
Hyper-resolution: avoiding intermediate clauses
Indexing: data structures for speed
Finding Refutations in FOL

Prenex: Move quantifiers to the front
Skolemize: Remove quantifiers, preserving consistency
Herbrand models: Reduce the class of interpretations
Herbrand's Thm: Contradictions have finite, ground proofs
Unification: Automatically find the right instantiations

Finally, combine unification with resolution

Prenex Normal Form

Convert to Negation Normal Form using additionally

\[-(\forall x \ A) \simeq \exists x \neg A\]
\[-(\exists x \ A) \simeq \forall x \neg A\]

Then move quantifiers to the front using

\[(\forall x \ A) \land B \simeq \forall x (A \land B)\]
\[(\forall x \ A) \lor B \simeq \forall x (A \lor B)\]

and the similar rules for \(\exists\)
Skolemization

Take a formula of the form

$$\forall x_1 \forall x_2 \cdots \forall x_k \exists y \ A$$

Choose a new $k$-place function symbol, say $f$

Delete $\exists y$ and replace $y$ by $f(x_1, x_2, \ldots, x_k)$. We get

$$\forall x_1 \forall x_2 \cdots \forall x_k A[f(x_1, x_2, \ldots, x_k)/y]$$

Repeat until no $\exists$ quantifiers remain

Example of Conversion to Clauses

For proving $\exists x [P(x) \rightarrow \forall y \ P(y)]$

$$\neg [\exists x [P(x) \rightarrow \forall y \ P(y)]]$$ negated goal

$$\forall x [P(x) \land \exists y \neg P(y)]$$ conversion to NNF

$$\forall x \exists y [P(x) \land \neg P(y)]$$ pulling $\exists$ out

$$\forall x [P(x) \land \neg P(f(x))]$$ Skolem term $f(x)$

$$\{P(x)\} \quad \{\neg P(f(x))\}$$ Final clauses
Correctness of Skolemization

The formula $\forall x \exists y \ A$ is consistent

$\iff$ it holds in some interpretation $I = (D, I)$

$\iff$ for all $x \in D$ there is some $y \in D$ such that $A$ holds

$\iff$ some function $\hat{f}$ in $D \rightarrow D$ yields suitable values of $y$

$\iff$ $A[f(x)/y]$ holds in some $I'$ extending $I$ so that $f$ denotes $\hat{f}$

$\iff$ the formula $\forall x A[f(x)/y]$ is consistent.

Herbrand Interpretations for a set of clauses $S$

$H_0 \overset{\text{def}}{=} \text{the set of constants in } S$

$H_{i+1} \overset{\text{def}}{=} H_i \cup \{f(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in H_i \}$

and $f$ is an $n$-place function symbol in $S$

$H \overset{\text{def}}{=} \bigcup_{i \geq 0} H_i \quad \text{Herbrand Universe}$

$HB \overset{\text{def}}{=} \{P(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in H \}$

and $P$ is an $n$-place predicate symbol in $S$
Example of an Herbrand Model

\[
\begin{align*}
\neg & \text{even}(1) \\
& \text{even}(2) \\
& \text{even}(X \cdot Y) \leftarrow \text{even}(X), \text{even}(Y)
\end{align*}
\]

\[H = \{1, 2, 1 \cdot 1, 1 \cdot 2, 2 \cdot 1, 2 \cdot 2, 1 \cdot (1 \cdot 1), \ldots\}\]

\[\text{HB} = \{\text{even}(1), \text{even}(2), \text{even}(1 \cdot 1), \text{even}(1 \cdot 2), \ldots\}\]

\[\text{I}[\text{even}] = \{\text{even}(2), \text{even}(1 \cdot 2), \text{even}(2 \cdot 1), \text{even}(2 \cdot 2), \ldots\}\]

(for model where \( \cdot \) means product; could instead use sum!)

A Key Fact about Herbrand Interpretations

Let \( S \) be a set of clauses.

\( S \) is unsatisfiable \( \iff \) no Herbrand interpretation satisfies \( S \)

- Holds because some Herbrand model mimicks every 'real' model
- We must consider only a small class of models
- Herbrand models are syntactic, easily processed by computer
Herbrand’s Theorem

Let $S$ be a set of clauses.

$S$ is unsatisfiable $\iff$ there is a finite unsatisfiable set $S'$ of ground instances of clauses of $S$.

- **Finite**: we can compute it
- **Instance**: result of substituting for variables
- **Ground**: no variables remain—it's propositional!
**Unification**

Finding a common instance of two terms

- Logic programming (Prolog)
- Polymorphic type-checking (ML)
- Constraint satisfaction problems
- Resolution theorem proving for FOL
- Many other theorem proving methods

**Substitutions**

A finite set of replacements

\[ \theta = [t_1/x_1, \ldots, t_k/x_k] \]

where \( x_1, \ldots, x_k \) are distinct variables and \( t_i \neq x_i \)

- \( f(t, u)\theta = f(t\theta, u\theta) \) (terms)
- \( P(t, u)\theta = P(t\theta, u\theta) \) (literals)
- \( \{L_1, \ldots, L_m\}\theta = \{L_1\theta, \ldots, L_m\theta\} \) (clauses)
Composing Substitutions

Composition of \( \phi \) and \( \theta \), written \( \phi \circ \theta \), satisfies for all terms \( t \)

\[
t(\phi \circ \theta) = (t\phi)\theta
\]

It is defined by (for all relevant \( x \))

\[
\phi \circ \theta \overset{\text{def}}{=} [ (x\phi)\theta / x, \ldots ]
\]

Consequences include \( \theta \circ [] = \theta \), and associativity:

\[
(\phi \circ \theta) \circ \sigma = \phi \circ (\theta \circ \sigma)
\]

Most General Unifiers

\( \theta \) is a unifier of terms \( t \) and \( u \) if \( t\theta = u\theta \)

\( \theta \) is more general than \( \phi \) if \( \phi = \theta \circ \sigma \)

\( \theta \) is most general if it is more general than every other unifier

If \( \theta \) unifies \( t \) and \( u \) then so does \( \theta \circ \sigma \):

\[
t(\theta \circ \sigma) = t\theta\sigma = u\theta\sigma = u(\theta \circ \sigma)
\]

A most general unifier of \( f(a, x) \) and \( f(y, g(z)) \) is \([a/y, g(z)/x]\)

The common instance is \( f(a, g(z)) \)
Algorithm for Unifying Two Terms

Represent terms by binary trees

Each term is a Variable \( x, y, \ldots \), Constant \( a, b, \ldots \), or Pair \( (t, t') \)

Constants do not unify with different Constants

Constants do not unify with Pairs

Variable \( x \) and term \( t \): unifier is \([t/x] \) — unless \( x \) occurs in \( t \)

Cannot unify \( f(x) \) with \( x \! \)

Unifying Two Pairs

\( \theta \circ \theta' \) unifies \( (t, t') \) with \( (u, u') \)

if \( \theta \) unifies \( t \) with \( u \) and \( \theta' \) unifies \( t'\theta \) with \( u'\theta \)

\[
(t, t')(\theta \circ \theta') = (t, t')\theta\theta' \\
= (t\theta\theta', t'\theta\theta') \\
= (u\theta\theta', u'\theta\theta') \\
= (u, u')\theta\theta' \\
= (u, u')(\theta \circ \theta')
\]
Examples of Unification

<table>
<thead>
<tr>
<th></th>
<th>f(x, b)</th>
<th>f(x, x)</th>
<th>f(x, x)</th>
<th>j(x, x, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f(a, y)</td>
<td>f(a, b)</td>
<td>f(y, g(y))</td>
<td>j(w, a, h(w))</td>
</tr>
<tr>
<td></td>
<td>f(a, b)</td>
<td>?</td>
<td>?</td>
<td>j(a, a, h(a))</td>
</tr>
<tr>
<td>[a/x, b/y]</td>
<td>FAIL</td>
<td>FAIL</td>
<td>[a/w, a/x, h(a)/z]</td>
<td></td>
</tr>
</tbody>
</table>

We always get a most general unifier.

Theorem-Proving Examples

(\exists y \forall x \ R(x, y)) \rightarrow (\forall x \exists y \ R(x, y))

Clauses after negation are \{R(x, a)\} and \{-R(b, y)\}

R(x, a) and R(b, y) have unifier [b/x, a/y]: contradiction!

(\forall x \exists y \ R(x, y)) \rightarrow (\exists y \forall x \ R(x, y))

Clauses after negation are \{R(x, f(x))\} and \{-R(g(y), y)\}

R(x, f(x)) and R(g(y), y) are not unifiable: occurs check

Formula is not a theorem!
Variations on Unification

Efficient unification algorithms: near-linear time

Indexing & Discrimination networks: fast retrieval of a unifiable term

Order-sorted unification: type-checking in Haskell

Associative/commutative operators: problems in group theory

Higher-order unification: support λ-calculus

Boolean unification: reasoning about sets
**Binary Resolution**

\[
\frac{\{B, A_1, \ldots, A_m\} \quad \{\neg D, C_1, \ldots, C_n\}}{\{A_1, \ldots, A_m, C_1, \ldots, C_n\}\sigma} \quad \text{provided } B\sigma = D\sigma
\]

First *rename variables apart* in the clauses! — say, to resolve

\[\{P(x)\}\text{ and } \{\neg P(g(x))\}\]

Always use a *most general unifier* (MGU)

Soundness? Same argument as for the propositional version

---

**Factorisation**

Collapsing similar literals *in one clause*:

\[
\frac{\{B_1, \ldots, B_k, A_1, \ldots, A_m\}}{\{B_1, A_1, \ldots, A_m\}\sigma} \quad \text{provided } B_1\sigma = \cdots = B_k\sigma
\]

*Normally combined with resolution*

Prove \(\forall x \exists y \neg (P(y, x) \leftrightarrow \neg P(y, y))\)

The clauses are  \[\{\neg P(y, a), \neg P(y, y)\} \quad \{P(y, y), P(y, a)\}\]

Factoring yields  \[\{\neg P(a, a)\}\]

Resolution yields the empty clause!
### A Non-Trivial Example

\[ \exists x [P \rightarrow Q(x)] \land \exists x [Q(x) \rightarrow P] \rightarrow \exists x [P \leftrightarrow Q(x)] \]

Clauses are \( \{P, \neg Q(b)\}, \{P, Q(x)\}, \{-P, \neg Q(x)\}, \{-P, Q(a)\} \)

Resolve \( \{P, \neg Q(b)\} \) with \( \{P, Q(x)\} \) getting \( \{P\} \)

Resolve \( \{-P, \neg Q(x)\} \) with \( \{-P, Q(a)\} \) getting \( \{-P\} \)

Resolve \( \{P\} \) with \( \{-P\} \) getting \( \square \)

*Implicit factoring:* \( \{P, P\} \mapsto \{P\} \)

*Many other proofs!*

### Prolog Clauses and Their Execution

*At most one* positive literal per clause!

*Definite clause* \( \{-A_1, \ldots, \neg A_m, B\} \) or \( B \leftarrow A_1, \ldots, A_m \).

*Goal clause* \( \{-A_1, \ldots, \neg A_m\} \) or \( \leftarrow A_1, \ldots, A_m \).

*Linear* resolution: a program clause with last goal clause

*Left-to-right* through program clauses

*Left-to-right* through goal clause’s literals

*Depth-first search:* backtracks, but still incomplete

*Unification without occurs check:* fast, but unsound!
A (Pure) Prolog Program

parent(elizabeth, charles).
parent(elizabeth, andrew).

parent(charles, william).
parent(charles, henry).

parent(andrew, beatrice).
parent(andrew, eugenia).

grand(X, Z) :- parent(X, Y), parent(Y, Z).
cousin(X, Y) :- grand(Z, X), grand(Z, Y).

Prolog Execution

 :- cousin(X, Y).
 :- grand(Z1, X), grand(Z1, Y).
 :- parent(Z1, Y2), parent(Y2, X), grand(Z1, Y).

* :- parent(charles, X), grand(elizabeth, Y).
X=william
  :- grand(elizabeth, Y).
  :- parent(elizabeth, Y5), parent(Y5, X).
* Y=beatrice
  :- parent(andrew, Y).

* = backtracking choice point

16 solutions including cousin(william, william)
and cousin(william, henry)
The Method of Model Elimination

A Prolog-like method; complete for First-Order Logic

Contrapositives: treat clause \{A_1, \ldots, A_m\} as m clauses

\[
\begin{align*}
A_1 & \leftarrow \neg A_2, \ldots, \neg A_m \\
A_2 & \leftarrow \neg A_3, \ldots, \neg A_m, \neg A_1 \\
& \vdots
\end{align*}
\]

Extension rule: when proving goal \( P \), may assume \( \neg P \)

A brute force method: efficient but no refinements such as subsumption

A Survey of Automatic Theorem Provers

Hyper-resolution: Otter, Gandalf, SPASS, Vampire, \ldots

Model Elimination: Prolog Technology Theorem Prover, SETHEO

Parallel ME: PARTHENON, PARTHEO

Higher-Order Logic: TPS, LEO

Tableau (sequent) based: LeanTAP, 3TAP, \ldots
Equality is reflexive, symmetric, transitive

Equality is substitutive over functions, predicates

- Use specialized prover: Knuth-Bendix, . . .
- Assert axioms directly
- Paramodulation rule

\[
\frac{\{B[t], A_1, \ldots, A_m\} \quad \{t = u, C_1, \ldots, C_n\}}{\{B[u], A_1, \ldots, A_m, C_1, \ldots, C_n\}}
\]
Modal Operators

\( W \): set of possible worlds (machine states, future times, \ldots )

\( R \): accessibility relation between worlds

\((W, R)\) is called a modal frame

\( \Box A \) means \( A \) is necessarily true \( \Rightarrow \) in all accessible worlds

\( \Diamond A \) means \( A \) is possibly true

\( \neg \Diamond A \equiv \Box \neg A \)

\( A \) cannot be true \( \iff \) \( A \) must be false

Semantics of Propositional Modal Logic

For a particular frame \((W, R)\)

An interpretation \( I \) maps the propositional letters to subsets of \( W \)

\( w \models A \) means \( A \) is true in world \( w \)

\( w \models P \iff w \in I(P) \)

\( w \models A \land B \iff w \models A \) and \( w \models B \)

\( w \models \Box A \iff \forall v \in I(A) \) for all \( v \) such that \( R(w, v) \)

\( w \models \Diamond A \iff \exists v \in I(A) \) for some \( v \) such that \( R(w, v) \)
Truth and Validity in Modal Logic

For a particular frame \((W, R)\), and interpretation \(I\)

\[ w \models A \] means \(A\) is true in world \(w\)

\[ \models_{W, R, I} A \] means \(w \models A\) for all \(w\) in \(W\)

\[ \models_{W, R} A \] means \(w \models A\) for all \(w\) and all \(I\)

\[ \models A \] means \(\models_{W, R} A\) for all frames; \(A\) is universally valid

... but typically we constrain \(R\) to be, say, transitive

All tautologies are universally valid

A Hilbert-Style Proof System for \(K\)

Extend your favourite propositional proof system with

\[ \text{Dist} \quad \Box (A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B) \]

Inference Rule: Necessitation

\[ \frac{\text{A}}{\Box \text{A}} \]

Treat \(\Diamond\) as a definition

\[ \text{\quad} \Diamond \text{A} \overset{\text{def}}{=} \neg \neg \text{A} \]
**Variant Modal Logics**

Start with pure modal logic, which is called $K$

Add axioms to constrain the accessibility relation:

- $\mathbf{T}$: $\Box A \rightarrow A$ (*reflexive*) logic $\mathbf{T}$
- $\mathbf{4}$: $\Box A \rightarrow \Box \Box A$ (*transitive*) logic $\mathbf{S4}$
- $\mathbf{B}$: $A \rightarrow \Box \Diamond A$ (*symmetric*) logic $\mathbf{S5}$

And countless others!

We shall mainly look at $\mathbf{S4}$

---

**Extra Sequent Calculus Rules for S4**

\[
\begin{align*}
\frac{\Gamma, \Delta \Rightarrow A}{\Box A, \Gamma \Rightarrow \Delta} & \quad \text{(}\Box \text{l)} \\
\frac{\Gamma \Rightarrow \Delta, \Box A}{\Gamma^{*} \Rightarrow \Delta^{*}, A} & \quad \text{(}\Box \text{r)} \\
\frac{\Gamma^{*} \Rightarrow \Delta^{*}, \Delta}{\Diamond A, \Gamma \Rightarrow \Delta} & \quad \text{(}\Diamond \text{l)} \\
\frac{\Gamma \Rightarrow \Delta, \Diamond A}{\Gamma \Rightarrow \Delta, \Box A} & \quad \text{(}\Diamond \text{r)}
\end{align*}
\]

$\Gamma^{*} \overset{\text{def}}{=} \{ \Box B \mid \Box B \in \Gamma \}$  
Erase non-$\Box$ assumptions

$\Delta^{*} \overset{\text{def}}{=} \{ \Diamond B \mid \Diamond B \in \Delta \}$  
Erase non-$\Diamond$ goals!
A Proof of the Distribution Axiom

\[ \begin{align*}
A \Rightarrow B, A & \quad B, A \Rightarrow B \\
\overline{A \Rightarrow B, A} & \quad \overline{B, A \Rightarrow B} \\
A \Rightarrow B, \Box A \Rightarrow B & \\
\Box (A \Rightarrow B), \Box A \Rightarrow B & \\
\Box (A \Rightarrow B), \Box A \Rightarrow \Box B & \\
\Diamond (A \Rightarrow B), \Box A \Rightarrow \Box B & \\
\Box (A \Rightarrow B), \Box A \Rightarrow \Box B & \\
\end{align*} \]

And thus  \( \Box (A \Rightarrow B) \rightarrow (\Box A \Rightarrow \Box B) \)

Must apply (\( \Box r \)) first!

Part of an Operator String Equivalence

\[ \begin{align*}
\Diamond A \Rightarrow \Diamond A & \\
\Box \Diamond A \Rightarrow \Diamond A & \\
\Diamond \Box \Diamond A \Rightarrow \Diamond A & \\
\Box \Diamond \Box \Diamond A \Rightarrow \Diamond A & \\
\Box \Diamond \Box \Diamond A \Rightarrow \Box \Diamond A & \\
\end{align*} \]

In fact,  \( \Box \Box \Box \Diamond A \simeq \Diamond \Box A \) also \( \Box \Box A \simeq \Box A \)

The S4 operator strings are  \( \Box \quad \Diamond \quad \Box \Diamond \quad \Box \Box \Diamond \quad \Box \Diamond \)
Two Failed Proofs

\[ \Rightarrow A \]
\[ \Rightarrow \Diamond A \]
\[ A \Rightarrow \Box \Diamond A \]

\[ B \Rightarrow A \land B \]
\[ B \Rightarrow \Diamond (A \land B) \]
\[ \Diamond A, \Diamond B \Rightarrow \Diamond (A \land B) \]

Can extract a countermodel from the proof attempt
Simplifying the Sequent Calculus

7 connectives (or 9 for modal logic):

\[ \land \ \land \ \lor \ \rightarrow \ \leftrightarrow \ \forall \ \exists \ (\Box \ \Diamond) \]

Left and right: so 14 rules (or 18) plus basic sequent, cut

Idea! Work in **Negation Normal Form**

Fewer connectives: \( \land \ \lor \ \forall \ \exists \ (\Box \ \Diamond) \)

Sequents need *one side only!*

---

Simplified Calculus: Left-Only

- \( \neg A, A, \Gamma \Rightarrow \) *(basic)*
- \( \neg A, \Gamma \Rightarrow A, \Gamma \Rightarrow \) *(cut)*

\[
\begin{align*}
& A, B, \Gamma \Rightarrow \quad (\land 1) \\
& A \land B, \Gamma \Rightarrow \\
& A, \Gamma \Rightarrow B, \Gamma \Rightarrow \\
& A \lor B, \Gamma \Rightarrow \quad (\lor 1)
\end{align*}
\]

\[
\begin{align*}
& A[t/x], \Gamma \Rightarrow \quad (\forall 1) \\
& \forall x A, \Gamma \Rightarrow \\
& A, \Gamma \Rightarrow \exists x A, \Gamma \Rightarrow \quad (\exists 1)
\end{align*}
\]

Rule \( \exists 1 \) holds *provided* \( x \) is not free in the conclusion!
Left-Only Sequent Rules for S4

\[
\begin{align*}
\frac{\Lambda, \Gamma \Rightarrow}{\square \Lambda, \Gamma \Rightarrow} & \quad (\square l) \\
\frac{\Lambda, \Gamma \Rightarrow}{\Diamond \Lambda, \Gamma \Rightarrow} & \quad (\Diamond l)
\end{align*}
\]

\(\Gamma^* \overset{\text{def}}{=} \{ \square B \mid \square B \in \Gamma \} \)

Erase non-\(\square\) assumptions

From 14 (or 18) rules to 4 (or 6)

Left-side only system uses **proof by contradiction**

Right-side only system is an exact **dual**

---

Proving \(\forall x (P \to Q(x)) \Rightarrow P \to \forall y Q(y)\)

Move the right-side formula to the left and convert to NNF:

\(P \land \exists y \neg Q(y), \forall x (\neg P \lor Q(x)) \Rightarrow\)

\[
\frac{P, \neg Q(y), \neg P}{P, \neg Q(y), \neg P \lor Q(y) \Rightarrow} \quad (\lor 1) \\
\frac{P, \neg Q(y), Q(y) \Rightarrow}{P, \neg Q(y), \neg P \lor Q(y) \Rightarrow} \quad (\lor 1) \\
\frac{P, \neg Q(y), \forall x (\neg P \lor Q(x)) \Rightarrow}{P, \exists y \neg Q(y), \forall x (\neg P \lor Q(x)) \Rightarrow} \quad (\exists 1) \\
\frac{P \land \exists y \neg Q(y), \forall x (\neg P \lor Q(x)) \Rightarrow}{P \land \exists y \neg Q(y), \forall x (\neg P \lor Q(x)) \Rightarrow} \quad (\land 1)
\]
Adding Unification

Rule (∀l) now inserts a new free variable:

\[ \frac{A[z/x], \Gamma \Rightarrow (∀l)}{∀x A, \Gamma \Rightarrow} \]

Let unification instantiate any free variable

In \(\neg A, B, \Gamma \Rightarrow\) try unifying \(A\) with \(B\) to make a basic sequent

Updating a variable affects entire proof tree

What about rule (∃l)? Skolemize!

Skolemization from NNF

Follow tree structure; don’t pull out quantifiers!

\([\forall y \exists z Q(y, z)] \land \exists x P(x)\) to \([\forall y Q(y, f(y))] \land P(a)\)

Better to push quantifiers in (called miniscoping)

Proving \(\exists x \forall y [P(x) \rightarrow P(y)]\)

Negate; convert to NNF: \(\forall x \exists y [P(x) \land \neg P(y)]\)

Push in the \(\exists y:\) \(\forall x [P(x) \land \exists y \neg P(y)]\)

Push in the \(\forall x:\) \(\forall x P(x) \land \exists y \neg P(y)\)

Skolemize: \(\forall x P(x) \land \neg P(a)\)
A Proof of $\exists x \forall y [P(x) \rightarrow P(y)]$

$y \mapsto f(z)$

\[
\begin{align*}
\frac{y \mapsto f(z)}{P(y), \neg P(f(y)), P(z), \neg P(f(z)) \Rightarrow} & \quad \text{(basic)} \\
\frac{P(y), \neg P(f(y)), P(z) \land \neg P(f(z)) \Rightarrow} {P(y),\neg P(f(y)), \forall x [P(x) \land \neg P(f(x))] \Rightarrow} & \quad \text{(\forall 1)} \\
\frac{P(y) \land \neg P(f(y)), \forall x [P(x) \land \neg P(f(x))] \Rightarrow} {\forall x [P(x) \land \neg P(f(x))] \Rightarrow} & \quad \text{(\forall 1)} \\
\end{align*}
\]

Unification chooses the term for $\forall 1$

---

A Failed Proof

Try to prove $\forall x [P(x) \lor Q(x)] \Rightarrow \forall x P(x) \lor \forall x Q(x)$

NNF: $\exists x \neg P(x) \land \exists x \neg Q(x), \forall x [P(x) \lor Q(x)] \Rightarrow$

Skolemize: $\neg P(a) \land \neg Q(b), \forall x [P(x) \lor Q(x)] \Rightarrow$

$y \mapsto a$

\[
\frac{\neg P(a), \neg Q(b), P(y) \Rightarrow}{\neg P(a), \neg Q(b), Q(y) \Rightarrow} \quad \text{(\forall 1)}
\]

$y \mapsto b$??

\[
\frac{\neg P(a), \neg Q(b), P(y) \Rightarrow}{\neg P(a), \neg Q(b), Q(y) \Rightarrow} \quad \text{(\forall 1)}
\]

\[
\frac{\neg P(a), \neg Q(b), \forall x [P(x) \lor Q(x)] \Rightarrow}{\neg P(a) \land \neg Q(b), \forall x [P(x) \lor Q(x)] \Rightarrow} \quad \text{(\forall 1)}
\]
prove((A,B),UnExp,Lits,FreeV,VarLim) :- !,
    prove(A, [B|UnExp], Lits, FreeV, VarLim).
prove((A;B), UnExp, Lits, FreeV, VarLim) :- !,
    prove(A, UnExp, Lits, FreeV, VarLim),
    prove(B, UnExp, Lits, FreeV, VarLim).
prove(all(X,Fml), UnExp, Lits, FreeV, VarLim) :- !,
    \+ length(FreeV, VarLim),
    copy_term((X,Fml,FreeV),(X1,Fml1,FreeV)),
    append(UnExp, [all(X,Fml)], UnExp1),
    prove(Fml1, UnExp1, Lits, [X1|FreeV], VarLim).
prove(Lit,_,[L|Lits],_,_):-
    (Lit = -Neg; -Lit = Neg) ->
    (unify(Neg,L); prove(Lit, [], Lits, _, _)).
prove(Lit, [Next|UnExp], Lits, FreeV, VarLim) :-
    prove(Next, UnExp, [Lit|Lits], FreeV, VarLim).