Introduction to Logic

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:

Black is the colour of my true love's hair.

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

What is the colour of my true love's hair?

I wish my true love had hair.

Get a haircut!

Schematic Statements

The *meta-variables* X, Y, Z, ... 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 $Y \mapsto \mathsf{coal}$ *satisfies* the statement

> *Black is the colour of* Y*.*

but the interpretation $Y \mapsto \mathsf{strawberries}$ does not!

A statement A is *valid* if all interpretations satisfy A.

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:

$$\{X \text{ part of } Y,\ Y \text{ part of } Z,\ Z \text{ part of } X\}$$

$$\{n \text{ is a positive integer},\ n \neq 1,\ n \neq 2,\ \ldots\}$$

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

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

Use inferences to construct finite proofs!

Schematic Inference Rules

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

A valid inference:

$$\frac{\mathsf{spoke\ part\ of\ wheel} \qquad \mathsf{wheel\ part\ of\ bike}}{\mathsf{spoke\ part\ of\ bike}}$$

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

$$\frac{\mathsf{cow\ part\ of\ chair} \qquad \mathsf{chair\ part\ of\ ant}}{\mathsf{cow\ part\ of\ ant}}$$

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.

Syntax of Propositional Logic

$P, Q, R, \ldots$	propositional letter
$\mathbf{t}$	true
$\mathbf{f}$	false
$\neg A$	not A
$A \land B$	A and B
$A \lor B$	A or B
$A \to B$	if A then B
$A \leftrightarrow B$	A if and only if B

Semantics of Propositional Logic

$\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$ are *truth-functional*: functions of their operands

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$
t	**t**	**f**	**t**	**t**	**t**	**t**
t	**f**	**f**	**f**	**t**	**f**	**f**
f	**t**	**t**	**f**	**t**	**t**	**f**
f	**f**	**t**	**f**	**f**	**t**	**t**

Interpretations of Propositional Logic

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

Interpretation I *satisfies* a formula A if the formula evaluates to $\mathbf{t}$.
Write $\models_{\mathrm{I}} A$

A is *valid* (a *tautology*) if every interpretation satisfies A

S is *satisfiable* if some interpretation satisfies every formula in S

Implication, Entailment, Equivalence

$A \to B$ means simply $\neg A \vee B$

$A \models B$ means if $\models_I A$ then $\models_I B$ for every interpretation I

$A \models B$ if and only if $\models A \to B$

Equivalence

$A \simeq B$ means $A \models B$ and $B \models A$

$A \simeq B$ if and only if $\models A \leftrightarrow B$

Equivalences

$$A \wedge A \simeq A$$

$$A \wedge B \simeq B \wedge A$$

$$(A \wedge B) \wedge C \simeq A \wedge (B \wedge C)$$

$$A \vee (B \wedge C) \simeq (A \vee B) \wedge (A \vee C)$$

$$A \wedge \mathbf{f} \simeq \mathbf{f}$$

$$A \wedge \mathbf{t} \simeq A$$

$$A \wedge \neg A \simeq \mathbf{f}$$

Dual versions: exchange $\wedge$, $\vee$ and $\mathbf{t}$, $\mathbf{f}$ in any equivalence

Negation Normal Form

1. Get rid of $\leftrightarrow$ and $\to$, leaving just $\wedge$, $\vee$, $\neg$:

$$A \leftrightarrow B \simeq (A \to B) \wedge (B \to A)$$

$$A \to B \simeq \neg A \vee B$$

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

$$\neg\neg A \simeq A$$

$$\neg(A \wedge B) \simeq \neg A \vee \neg B$$

$$\neg(A \vee B) \simeq \neg A \wedge \neg B$$

From NNF to Conjunctive Normal Form

3. Push disjunctions in, using distributive laws:

$$A \vee (B \wedge C) \simeq (A \vee B) \wedge (A \vee C)$$

$$(B \wedge C) \vee A \simeq (B \vee A) \wedge (C \vee A)$$

4. Simplify:

- Delete any disjunction containing P and $\neg P$
- Delete any disjunction that's included in another
- Replace $(P \vee A) \wedge (\neg P \vee A)$ by A

Converting a Non-Tautology to CNF

$$P \vee Q \to Q \vee R$$

1. Elim $\to$: $\quad \neg(P \vee Q) \vee (Q \vee R)$
2. Push $\neg$ in: $\quad (\neg P \wedge \neg Q) \vee (Q \vee R)$
3. Push $\vee$ in: $\quad (\neg P \vee Q \vee R) \wedge (\neg Q \vee Q \vee R)$
4. Simplify: $\quad \neg P \vee Q \vee R$

Not a tautology: try $P \mapsto \mathbf{t}$, $Q \mapsto \mathbf{f}$, $R \mapsto \mathbf{f}$

Tautology checking using CNF

$$((P \to Q) \to P) \to P$$

1. Elim $\to$: $\neg[\neg(\neg P \vee Q) \vee P] \vee P$
2. Push $\neg$ in: $[\neg\neg(\neg P \vee Q) \wedge \neg P] \vee P$

   $[(\neg P \vee Q) \wedge \neg P] \vee P$
3. Push $\vee$ in: $(\neg P \vee Q \vee P) \wedge (\neg P \vee P)$
4. Simplify: $\mathbf{t} \wedge \mathbf{t}$

   $\mathbf{t}$ *It's a tautology!*

A Simple Proof System

Axiom Schemes

K $\quad A \to (B \to A)$

S $\quad (A \to (B \to C)) \to ((A \to B) \to (A \to C))$

DN $\quad \neg\neg A \to A$

Inference Rule: Modus Ponens

$$\frac{A \to B \qquad A}{B}$$

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

(1) $(A \to ((D \to A) \to A)) \to$
$((A \to (D \to A)) \to (A \to A))$ by S

(2) $A \to ((D \to A) \to A)$ by K

(3) $(A \to (D \to A)) \to (A \to A)$ by MP, (1), (2)

(4) $A \to (D \to A)$ by K

(5) $A \to A$ by MP, (3), (4)

Some Facts about Deducibility

A is *deducible from* the set S of 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 $S \vdash A \to B$.

Gentzen's Natural Deduction Systems

A varying context of *assumptions*

Each logical connective defined *independently*

Introduction rule for $\land$: how to deduce $A \land B$

$$\frac{A \qquad B}{A \land B}$$

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

$$\frac{A \land B}{A} \qquad \frac{A \land B}{B}$$

The Sequent Calculus

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

$$\text{if } A_1 \land \ldots \land A_m \text{ then } B_1 \lor \ldots \lor B_n$$

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

Γ and Δ are *sets* in $\Gamma \Rightarrow \Delta$

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

Sequent Calculus Rules

$$\frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta} \; (cut)$$

$$\frac{\Gamma \Rightarrow \Delta, A}{\neg A, \Gamma \Rightarrow \Delta} \; (\neg l) \qquad \frac{A, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg A} \; (\neg r)$$

$$\frac{A, B, \Gamma \Rightarrow \Delta}{A \land B, \Gamma \Rightarrow \Delta} \; (\land l) \qquad \frac{\Gamma \Rightarrow \Delta, A \qquad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \land B} \; (\land r)$$

More Sequent Calculus Rules

$$\frac{A, \Gamma \Rightarrow \Delta \qquad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta}\ (\vee l) \qquad \frac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}\ (\vee r)$$

$$\frac{\Gamma \Rightarrow \Delta, A \qquad B, \Gamma \Rightarrow \Delta}{A \rightarrow B, \Gamma \Rightarrow \Delta}\ (\rightarrow l) \qquad \frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \rightarrow B}\ (\rightarrow r)$$

Easy Sequent Calculus Proofs

$$
\dfrac{\dfrac{\overline{A, B \Rightarrow A}}{A \land B \Rightarrow A}\,(\land l)}{\Rightarrow A \land B \to A}\,(\to r)
$$

$$
\dfrac{\dfrac{\dfrac{\overline{A, B \Rightarrow B, A}}{A \Rightarrow B, B \to A}\,(\to r)}{\Rightarrow A \to B,\ B \to A}\,(\to r)}{\Rightarrow (A \to B) \lor (B \to A)}\,(\lor r)
$$

Part of a Distributive Law

$$
\begin{prooftree}
\AxiomC{}
\UnaryInfC{$A \Rightarrow A, B$}
\AxiomC{}
\UnaryInfC{$B, C \Rightarrow A, B$}
\RightLabel{$(\wedge l)$}
\UnaryInfC{$B \wedge C \Rightarrow A, B$}
\RightLabel{$(\vee l)$}
\BinaryInfC{$A \vee (B \wedge C) \Rightarrow A, B$}
\RightLabel{$(\vee r)$}
\UnaryInfC{$A \vee (B \wedge C) \Rightarrow A \vee B$}
\AxiomC{similar}
\RightLabel{$(\wedge r)$}
\BinaryInfC{$A \vee (B \wedge C) \Rightarrow (A \vee B) \wedge (A \vee C)$}
\end{prooftree}
$$

Second subtree proves $A \vee (B \wedge C) \Rightarrow A \vee C$ similarly

A Failed Proof

$$\dfrac{\dfrac{\dfrac{A \Rightarrow B, C \qquad \overline{B \Rightarrow B, C}}{A \lor B \Rightarrow B, C}\,(\lor l)}{A \lor B \Rightarrow B \lor C}\,(\lor r)}{\Rightarrow A \lor B \to B \lor C}\,(\to r)$$

$A \mapsto \mathbf{t},\ B \mapsto \mathbf{f},\ C \mapsto \mathbf{f}$ falsifies unproved sequent!

Ordered Binary Decision Diagrams

Normal form based upon the connective $\mathbf{if}(A, B, C)$:

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

Detects if a formula is tautologous ($\mathbf{t}$) or inconsistent ($\mathbf{f}$)

A **FAST** way of verifying digital circuits, . . .

The $\mathbf{if}$-Connective

$\mathbf{if}(A, B, C)$ means 'if A is true then B else C'

In symbols, $(A \to B) \land (\neg A \to C)$

$\mathbf{if}$ can express the other connectives:

$$A \land B \simeq \mathbf{if}(A, B, \mathbf{f}) \qquad \neg A \simeq \mathbf{if}(A, \mathbf{f}, \mathbf{t})$$

$$A \lor B \simeq \mathbf{if}(A, \mathbf{t}, B) \qquad A \to B \simeq \mathbf{if}(A, B, \mathbf{t})$$

$\mathbf{if}$-distribution law: $h(\mathbf{if}(A, B, C)) = \mathbf{if}(A, h(B), h(C))$

If Normal Form: Naïve Algorithm

1. Replace all other connectives by $\mathbf{if}$-forms

2. Simplify each test to a variable:

$$\mathbf{if}(\mathbf{t}, B, C) \simeq B \qquad \mathbf{if}(\mathbf{f}, B, C) \simeq C$$

$$\mathbf{if}(\mathbf{if}(A, B, C), D, E) \simeq \mathbf{if}(A, \mathbf{if}(B, D, E), \mathbf{if}(C, D, E))$$

3. Remove tested variables from subtrees; repeat step 2:

$$\mathbf{if}(P, A, B) \simeq \mathbf{if}(P,\ A[\mathbf{t}/P],\ B[\mathbf{f}/P])$$

If Normal Form, Continued

4. Erase redundant tests:

$$\mathbf{if}(P, A, A) \simeq A$$

5. Ensure every variable occurrence is tested:

$$P \simeq \mathbf{if}(P, \mathbf{t}, \mathbf{f})$$

Result is a *binary tree*:

Nodes carry variables; *leaves* are $\mathbf{t}$ or $\mathbf{f}$

Example: Normalising $P \vee Q \to Q \vee R$

1. $\mathbf{if}(\mathbf{if}(P, \mathbf{t}, Q), \mathbf{if}(Q, \mathbf{t}, R), \mathbf{t})$
2. $\mathbf{if}(P,\ \mathbf{if}(\mathbf{t}, \mathbf{if}(Q, \mathbf{t}, R), \mathbf{t}),\ \mathbf{if}(Q, \mathbf{if}(Q, \mathbf{t}, R), \mathbf{t}))$

   $\mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, R),\ \mathbf{if}(Q, \mathbf{if}(Q, \mathbf{t}, R),\ \mathbf{t}))$
3. $\mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, R),\ \mathbf{if}(Q, \mathbf{if}(\mathbf{t}, \mathbf{t}, R), \mathbf{t}))$

   $\mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, R),\ \mathbf{if}(Q, \mathbf{t}, \mathbf{t}))$
4. $\mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, R),\ \mathbf{t})$
5. $\mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, \mathbf{if}(R, \mathbf{t}, \mathbf{f})),\ \mathbf{t})$

FAST Canonical Form Algorithm

To do $\neg Z$, where Z is already canonical: swap $\mathbf{t} \rightleftharpoons \mathbf{f}$

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

Trivial if either is $\mathbf{t}$ *or* $\mathbf{f}$. *Treat* $\vee$, $\rightarrow$, $\leftrightarrow$ *similarly!*

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

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

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

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

Canonical Form of $((P \to Q) \to P) \to P$

$$
\begin{aligned}
P \to Q &: \mathbf{if}(P, \mathbf{t}, \mathbf{f}) \to \mathbf{if}(Q, \mathbf{t}, \mathbf{f}) \\
&: \mathbf{if}(P,\ \mathbf{t} \to \mathbf{if}(Q, \mathbf{t}, \mathbf{f}),\ \mathbf{f} \to \mathbf{if}(Q, \mathbf{t}, \mathbf{f})) \\
&: \mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, \mathbf{f}),\ \mathbf{t}) \\
(P \to Q) \to P &: \mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, \mathbf{f}),\ \mathbf{t}) \to \mathbf{if}(P, \mathbf{t}, \mathbf{f}) \\
&: \mathbf{if}(P,\ \mathbf{if}(Q, \mathbf{t}, \mathbf{f}) \to \mathbf{t},\ \mathbf{t} \to \mathbf{f}) \\
&: \mathbf{if}(P,\ \mathbf{t},\ \mathbf{f}) \\
((P \to Q) \to P) \to P &: \mathbf{if}(P,\ \mathbf{t},\ \mathbf{f}) \to \mathbf{if}(P,\ \mathbf{t},\ \mathbf{f}) \\
&: \mathbf{t}
\end{aligned}
$$

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 $\mathbf{if}(P, X, X) \simeq X$
- Fast processing of $X \wedge X$, $X \vee X$, $X \to X$, $\ldots$

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

Final Observations

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*

$$\frac{\mathsf{father}(\mathsf{father}(x)) = \mathsf{father}(\mathsf{father}(y))}{\mathsf{cousin}(x, y)}$$

Reasons about *all* and *some* individuals:

$$\frac{\text{All men are mortal} \qquad \text{Socrates is a man}}{\text{Socrates is mortal}}$$

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

Relation Symbols; Formulae

Each *relation symbol* stands for an n-place relation

Equality is the 2-place relation symbol $=$

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 $\neg$, $\wedge$, $\vee$, . . .

(Later we add *quantifiers*)

Power of Quantifier-Free FOL

Very expressive, given strong induction rules

Prove equivalence of mathematical functions:

$$p(z, 0) = 1 \qquad\qquad q(z, 1) = z$$

$$p(z, n + 1) = p(z, n) \times z \qquad\qquad q(z, 2 \times n) = q(z \times z, n)$$

$$q(z, 2 \times n + 1) = q(z \times z, n) \times z$$

Boyer/Moore Theorem Prover: checked Gödel's Theorem, . . .

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:

$\forall xyz\, A$	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:

$$\forall x\,(\mathsf{man}(x) \to \mathsf{mortal}(x))$$

All mothers are female:

$$\forall x\,\mathsf{female}(\mathsf{mother}(x))$$

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

$$\exists x\,[A(x) \land \forall y\,(A(y) \to y = x)]$$

How do we interpret $\mathsf{mortal}(\mathsf{Socrates})$?

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

D is a non-empty *universe*

I maps symbols to ‘real’ functions, relations

c a constant symbol	$\mathrm{I}[\mathrm{c}] \in \mathrm{D}$
f an n-place function symbol	$\mathrm{I}[\mathrm{f}] \in \mathrm{D}^n \to \mathrm{D}$
P an n-place relation symbol	$\mathrm{I}[\mathrm{P}] \subseteq \mathrm{D}^n$

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

A *valuation* supplies the values of free variables

It is a function V : variables $\to D$

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

$$\mathcal{I}_V[x] \stackrel{\text{def}}{=} V(x) \qquad \text{if } x \text{ is a variable}$$

$$\mathcal{I}_V[c] \stackrel{\text{def}}{=} I[c]$$

$$\mathcal{I}_V[f(t_1, \ldots, t_n)] \stackrel{\text{def}}{=} I[f](\mathcal{I}_V[t_1], \ldots, \mathcal{I}_V[t_n])$$

The Meaning of Truth — in FOL

For interpretation $\mathcal{I}$ and valuation V

$\models_{\mathcal{I},V} P(t)$	if $I[P](\mathcal{I}_V[t])$ holds
$\models_{\mathcal{I},V} t = u$	if $\mathcal{I}_V[t]$ equals $\mathcal{I}_V[u]$
$\models_{\mathcal{I},V} A \land B$	if $\models_{\mathcal{I},V} A$ and $\models_{\mathcal{I},V} B$
$\models_{\mathcal{I},V} \exists x\, A$	if $\models_{\mathcal{I},V\{m/x\}} A$ holds for some $m \in D$
$\models_{\mathcal{I}} A$	if $\models_{\mathcal{I},V} A$ holds for all V

A is *satisfiable* if $\models_{\mathcal{I}} A$ for some $\mathcal{I}$

Free v 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:

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

May *rename* bound variables:

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

Substitution for Free Variables

$A[t/x]$ means 'substitute t for x in A':

$$(B \land C)[t/x] \text{ is } B[t/x] \land C[t/x]$$

$$(\forall x\, B)[t/x] \text{ is } \forall x\, B$$

$$(\forall y\, B)[t/x] \text{ is } \forall y\, B[t/x] \qquad (x \neq y)$$

$$(P(u))[t/x] \text{ is } P(u[t/x])$$

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

$$\neg(\forall x\, A) \simeq \exists x\, \neg A$$

$$(\forall x\, A) \land B \simeq \forall x\, (A \land B)$$

$$(\forall x\, A) \lor B \simeq \forall x\, (A \lor B)$$

$$(\forall x\, A) \land (\forall x\, B) \simeq \forall x\, (A \land B)$$

$$(\forall x\, A) \to B \simeq \exists x\, (A \to B)$$

$$\forall x\, A \simeq \forall x\, A \land A[t/x]$$

Dual versions: exchange $\forall$, $\exists$ and $\land$, $\lor$

Reasoning by Equivalences

$$
\begin{aligned}
\exists x\,(x = a \land P(x)) &\simeq \exists x\,(x = a \land P(a)) \\
&\simeq \exists x\,(x = a) \land P(a) \\
&\simeq P(a)
\end{aligned}
$$

$$
\begin{aligned}
&\exists z\,(P(z) \to P(a) \land P(b)) \\
&\qquad \simeq \forall z\,P(z) \to P(a) \land P(b) \\
&\qquad \simeq \forall z\,P(z) \land P(a) \land P(b) \to P(a) \land P(b) \\
&\qquad \simeq \mathbf{t}
\end{aligned}
$$

Sequent Calculus Rules for $\forall$

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

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

$$\frac{\overline{P(y) \Rightarrow P(y)}}{P(y) \Rightarrow \forall y\, P(y)}\ (\forall r)$$

Examples of the $\forall$ Rules

$$
\dfrac{\dfrac{\dfrac{\dfrac{\overline{A \Rightarrow B, A} \qquad \overline{A, B \Rightarrow B}}{A,\ A \to B \Rightarrow B}\,(\to l)}{A,\ \forall x\,(A \to B) \Rightarrow B}\,(\forall l)}{A,\ \forall x\,(A \to B) \Rightarrow \forall x\, B}\,(\forall r)}{\forall x\,(A \to B) \Rightarrow A \to \forall x\, B}\,(\to r)
$$

$$
\dfrac{\dfrac{\overline{P(f(y)) \Rightarrow P(f(y))}}{\forall x\, P(x) \Rightarrow P(f(y))}\,(\forall l)}{\forall x\, P(x) \Rightarrow \forall y\, P(f(y))}\,(\forall r)
$$

x must not be free in A !

Sequent Calculus Rules for $\exists$

$$\frac{A, \Gamma \Rightarrow \Delta}{\exists x\, A, \Gamma \Rightarrow \Delta}\ (\exists l) \qquad \frac{\Gamma \Rightarrow \Delta, A[t/x]}{\Gamma \Rightarrow \Delta, \exists x\, A}\ (\exists r)$$

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

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

Say, to prove

$$\exists z\, (P(z) \to P(a) \land P(b))$$

Part of the $\exists$ Distributive Law

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \Rightarrow A, B}}{A \Rightarrow A \lor B}\,(\lor r)}{A \Rightarrow \exists x\,(A \lor B)}\,(\exists r)}{\exists x\,A \Rightarrow \exists x\,(A \lor B)}\,(\exists l) \qquad \dfrac{\text{similar}}{\exists x\,B \Rightarrow \exists x\,(A \lor B)}\,(\exists l)}{\exists x\,A \lor \exists x\,B \Rightarrow \exists x\,(A \lor B)}\,(\lor l)$$

Second subtree proves $\exists x\,B \Rightarrow \exists x\,(A \lor B)$ similarly

A Failed Proof

$$
\dfrac{\dfrac{\dfrac{\dfrac{A, B[y/x] \Rightarrow A \land B}{A, B[y/x] \Rightarrow \exists x\,(A \land B)}\;(\exists r)}{A, \exists x\,B \Rightarrow \exists x\,(A \land B)}\;(\exists l)}{\exists x\,A, \exists x\,B \Rightarrow \exists x\,(A \land B)}\;(\exists l)}{\exists x\,A \land \exists x\,B \Rightarrow \exists x\,(A \land B)}\;(\land l)
$$

Cannot use $(\exists l)$ twice with the same x

We can easily falsify the topmost sequent

Clause Form

Clause: a disjunction of *literals*

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

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

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

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

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 Decision Procedure

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

Davis-Putnam on a Non-Tautology

Consider $P \vee Q \to Q \vee R$

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

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

Clauses satisfiable by $P \mapsto \mathbf{t},\ Q \mapsto \mathbf{f},\ R \mapsto \mathbf{f}$

Example of a Case Split on P

$$\{\neg Q, R\} \quad \{\neg R, P\} \quad \{\neg R, Q\} \quad \{\neg P, Q, R\} \quad \{P, Q\} \quad \{\neg P, \neg Q\}$$

$\{\neg Q, R\}$	$\{\neg R, Q\}$	$\{Q, R\}$	$\{\neg Q\}$	if P is true
	$\{\neg R\}$	$\{R\}$		unit $\neg Q$
	$\square$			unit R
$\{\neg Q, R\}$	$\{\neg R\}$	$\{\neg R, Q\}$	$\{Q\}$	if P is false
$\{\neg Q\}$			$\{Q\}$	unit $\neg R$
			$\square$	unit $\neg Q$

The Resolution Rule

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

In set notation,

$$\frac{\{B, A_1, \ldots, A_m\} \qquad \{\neg B, C_1, \ldots, C_n\}}{\{A_1, \ldots, A_m, C_1, \ldots, C_n\}}$$

Some special cases:

$$\frac{\{B\} \qquad \{\neg B, C_1, \ldots, C_n\}}{\{C_1, \ldots, C_n\}} \qquad\qquad \frac{\{B\} \qquad \{\neg B\}}{\square}$$

Simple Example: Proving $P \land Q \to Q \land P$

0. Negate! $\neg[P \land Q \to Q \land P]$

1. Elim $\to$: $\neg[\neg(P \land Q) \lor Q \land P]$

2. Push $\neg$ in: $\neg\neg(P \land Q) \land \neg(Q \land P)$

   $(P \land Q) \land (\neg Q \lor \neg P)$

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

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

Resolve $\{Q\}$ and $\{\neg Q\}$ getting $\square$

Clauses from Assumptions

Proving $(P \vee Q) \wedge (P \vee R) \to P \vee (Q \wedge R)$

From the assumptions, get clauses $\{P, Q\}$ and $\{P, R\}$

From $\neg\,[P \vee (Q \wedge R)]$ 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 $\square$

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

Reducing FOL to Propositional Logic

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

$$\neg(\forall x\, A) \simeq \exists x\, \neg A$$

$$\neg(\exists x\, A) \simeq \forall x\, \neg A$$

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 $f(x)$
$\{P(x)\} \qquad \{\neg P(f(x))\}$	Final clauses

Correctness of Skolemization

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

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

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

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

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

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

$$H_{i+1} \stackrel{\text{def}}{=} H_i \cup \{f(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in H_i$$
$$\text{and } f \text{ is an } n\text{-place function symbol in } S\}$$

$$H \stackrel{\text{def}}{=} \bigcup_{i \geq 0} H_i \qquad \textit{Herbrand Universe}$$

$$HB \stackrel{\text{def}}{=} \{P(t_1, \ldots, t_n) \mid t_1, \ldots, t_n \in H_i$$
$$\text{and } P \text{ is an } n\text{-place predicate symbol in } S\}$$

Example of an Herbrand Model

$$\left.\begin{aligned} even(0) \\ \neg even(1) \\ even(X+Y) &\leftarrow even(X), even(Y) \\ even(X+Y) &\leftarrow \neg even(X), \neg even(Y) \end{aligned}\right\} \textit{clauses}$$

$$H = \{0, 1, 0+0, 0+1, 1+0, 1+1, 0+(0+0), \ldots\}$$

$$HB = \{even(0), even(1), even(0+0), even(0+1), \ldots\}$$

$$I[even] = \{even(0), even(0+0), even(1+1), \ldots\}$$

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**: and 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 $t_i \neq x_i$

$$f(t_1, \ldots, t_n)\theta = f(t_1\theta, \ldots, t_n\theta) \qquad \text{(terms)}$$

$$P(t_1, \ldots, t_n)\theta = P(t_1\theta, \ldots, t_n\theta) \qquad \text{(literals)}$$

$$\{L_1, \ldots, L_m\}\theta = \{L_1\theta, \ldots, L_m\theta\} \qquad \text{(clauses)}$$

Composing Substitutions

Composition of ϕ and θ, written $\phi \circ \theta$, satisfies for all terms t

$$t(\phi \circ \theta) = (t\phi)\theta$$

It is defined by (for all relevant x)

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

Consequences include $\theta \circ [] = \theta$, and *associativity*:

$$(\phi \circ \theta) \circ \sigma = \phi \circ (\theta \circ \sigma)$$

Most General Unifiers

θ is a *unifier* of terms t and u if $t\theta = u\theta$

θ is *more general* than ϕ if $\phi = \theta \circ \sigma$

θ is *most general* if it is more general than every other unifier

If θ 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 θ unifies t with u and θ' unifies $t'\theta$ with $u'\theta$

$$\begin{aligned}
(t, t')(\theta \circ \theta') &= (t, t')\theta\theta' \\
&= (t\theta\theta', t'\theta\theta') \\
&= (u\theta\theta', u'\theta\theta') \\
&= (u, u')\theta\theta' \\
&= (u, u')(\theta \circ \theta')
\end{aligned}$$

Examples of Unification

$f(x,b)$	$f(x,x)$	$f(x,x)$	$j(x,x,z)$
$f(a,y)$	$f(a,b)$	$f(y,g(y))$	$j(w,a,h(w))$
$f(a,b)$	$?$	$?$	$j(a,a,h(a))$
$[a/x, b/y]$	**FAIL**	**FAIL**	$[a/w, a/x, h(a)/z]$

We always get a **most general** unifier

Theorem-Proving Examples

$$(\exists y\, \forall x\, R(x,y)) \to (\forall x\, \exists y\, R(x,y))$$

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

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

$$(\forall x\, \exists y\, R(x,y)) \to (\exists y\, \forall x\, R(x,y))$$

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

$R(x,f(x))$ and $R(g(y),x)$ 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

$$\frac{\{B, A_1, \ldots, A_m\} \qquad \{\neg D, C_1, \ldots, C_n\}}{\{A_1, \ldots, A_m, C_1, \ldots, C_n\}\sigma} \qquad \textit{provided } B\sigma = D\sigma$$

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

$$\{P(x)\} \quad \text{and} \quad \{\neg P(g(x))\}$$

Always use a *most general* unifier (MGU)

Soundness? Same argument as for the propositional version

Factorisation

Collapsing similar literals *in one clause*:

$$\frac{\{B_1, \ldots, B_k, A_1, \ldots, A_m\}}{\{B_1, A_1, \ldots, A_m\}\sigma} \qquad \textit{provided } B_1\sigma = \cdots = B_k\sigma$$

Normally combined with resolution

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

The clauses are	$\{\neg P(y,a), \neg P(y,y)\}$	$\{P(y,y), P(y,a)\}$
Factoring yields	$\{\neg P(y,a)\}$	$\{P(y,a)\}$

Resolution yields the empty clause!

A Non-Trivial Example

$$\exists x\,[P \to Q(x)] \land \exists x\,[Q(x) \to P] \to \exists x\,[P \leftrightarrow Q(x)]$$

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

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

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

Resolve $\{P\}$ with $\{\neg P\}$ getting $\Box$

Implicit factoring: $\{P, P\} \mapsto \{P\}$ *Many other proofs!*

Prolog Clauses and Their Execution

At most one positive literal per clause!

Definite clause $\{\neg A_1, \ldots, \neg A_m, B\}$ or $B \leftarrow A_1, \ldots, A_m.$

Goal clause $\{\neg A_1, \ldots, \neg A_m\}$ 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,Y).
*                         :- parent(andrew,Y).
Y=beatrice                                :- □.
```

* = 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

$$A_1 \leftarrow \neg A_2, \ldots, \neg A_m$$

$$A_2 \leftarrow \neg A_3, \ldots, \neg A_m, A_1$$

$$\vdots$$

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

A brute force method: efficient but no subsumption, ...

A Survey of Automatic Theorem Provers

Hyper-resolution: Aura, Otter

Model Elimination: Prolog Technology Theorem Prover, SETHEO

Parallel ME: PARTHENON, PARTHEO

Higher-Order Logic: TPS

Tableau (sequent) based: HARP, LeanTAP

Approaches to Equality Reasoning

Equality is *reflexive*, *symmetric*, *transitive*

Equality is *substitutive* over functions, predicates

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

$$\frac{\{B[t], A_1, \ldots, A_m\} \qquad \{t = u, C_1, \ldots, C_n\}}{\{B[u], A_1, \ldots, A_m, C_1, \ldots, C_n\}}$$

Modal Operators

W: set of *possible worlds* (machine states, future times, . . .)

R: *accessibility relation* between worlds

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

$$\left.\begin{array}{l}\Box A \text{ means } A \text{ is } \textit{necessarily true} \\ \Diamond A \text{ means } A \text{ is } \textit{possibly true}\end{array}\right\} \text{— in all } \textbf{accessible} \text{ worlds}$$

$\neg\Diamond A \simeq \Box\neg A$ $\qquad$ 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

$w \Vdash A$ means A is true in world w $\qquad (w \in W)$

$$
\begin{aligned}
w \Vdash P &\iff w \in I(P) \\
w \Vdash A \land B &\iff w \Vdash A \text{ and } w \Vdash B \\
w \Vdash \Box A &\iff v \Vdash A \text{ for all } v \text{ such that } R(w, v) \\
w \Vdash \Diamond A &\iff v \Vdash A \text{ for some } v \text{ such that } R(w, v)
\end{aligned}
$$

Truth and Validity in Modal Logic

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

$w \Vdash A$ means A is true in world w

$\models_{W,R,I} A$ means $w \Vdash A$ for all w

$\models_{W,R} A$ means $w \Vdash A$ 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 $\mathbb{K}$

Extend your favourite propositional proof system with

$$\text{Dist} \quad \Box(A \to B) \to (\Box A \to \Box B)$$

Inference Rule: *Necessitation*

$$\frac{A}{\Box A}$$

Treat $\Diamond$ as a definition

$$\Diamond A \stackrel{\text{def}}{=} \neg\Box\neg A$$

Variant Modal Logics

Start with pure modal logic, K

Add axioms to constrain the accessibility relation:

T	$\Box A \to A$	(reflexive)	logic T
4	$\Box A \to \Box\Box A$	(transitive)	logic $S4$
B	$A \to \Box\Diamond A$	(symmetric)	logic $S5$

And countless others!

We shall mainly look at $S4$

Extra Sequent Calculus Rules for $S4$

$$\frac{A, \Gamma \Rightarrow \Delta}{\Box A, \Gamma \Rightarrow \Delta}\ (\Box l) \qquad \frac{\Gamma^* \Rightarrow \Delta^*, A}{\Gamma \Rightarrow \Delta, \Box A}\ (\Box r)$$

$$\frac{A, \Gamma^* \Rightarrow \Delta^*}{\Diamond A, \Gamma \Rightarrow \Delta}\ (\Diamond l) \qquad \frac{\Gamma \Rightarrow \Delta, A}{\Gamma \Rightarrow \Delta, \Diamond A}\ (\Diamond r)$$

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

$\Delta^* \stackrel{\text{def}}{=} \{\Diamond B \mid \Diamond B \in \Delta\}$ Erase non-$\Diamond$ goals

A Proof of the Distribution Axiom

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \Rightarrow A} \qquad \dfrac{}{B \Rightarrow B}}{A \to B, A \Rightarrow B}\;(\to l)}{A \to B, \Box A \Rightarrow B}\;(\Box l)}{\Box(A \to B), \Box A \Rightarrow B}\;(\Box l)}{\Box(A \to B), \Box A \Rightarrow \Box B}\;(\Box r)
$$

And thus $\Box(A \to B) \to (\Box A \to \Box B)$

Must apply $(\Box r)$ first!

Part of an Operator String Equivalence

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\Diamond A \Rightarrow \Diamond A}}{\Box\Diamond A \Rightarrow \Diamond A}(\Box l)}{\Diamond\Box\Diamond A \Rightarrow \Diamond A}(\Diamond l)}{\Box\Diamond\Box\Diamond A \Rightarrow \Diamond A}(\Box l)}{\Box\Diamond\Box\Diamond A \Rightarrow \Box\Diamond A}(\Box r)
$$

In fact, $\Box\Diamond\Box\Diamond A \simeq \Box\Diamond A$ also $\Box\Box A \simeq A$

The $S4$ operator strings are $\Box$ $\Diamond$ $\Box\Diamond$ $\Diamond\Box$ $\Box\Diamond\Box$ $\Diamond\Box\Diamond$

Two Failed Proofs

$$\cfrac{\cfrac{\Rightarrow A}{\Rightarrow \Diamond A}\ (\Diamond r)}{A \Rightarrow \Box\Diamond A}\ (\Box r)$$

$$\cfrac{\cfrac{B \Rightarrow A \land B}{B \Rightarrow \Diamond(A \land B)}\ (\Diamond r)}{\Diamond A, \Diamond B \Rightarrow \Diamond(A \land B)}\ (\Diamond l)$$

Can extract a countermodel from the proof attempt

Simplifying the Sequent Calculus

7 connectives *(or 9 for modal logic)*:

$$\neg \quad \wedge \quad \vee \quad \to \quad \leftrightarrow \quad \forall \quad \exists \qquad (\Box \quad \Diamond)$$

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

Idea! Work in **Negation Normal Form**

Fewer connectives: $\wedge \quad \vee \quad \forall \quad \exists \qquad (\Box \quad \Diamond)$

Sequents need *one side only!*

Simplified Calculus: Left-Only

$$\frac{}{\neg A, A, \Gamma \Rightarrow}\ (\mathrm{basic}) \qquad \frac{\neg A, \Gamma \Rightarrow \qquad A, \Gamma \Rightarrow}{\Gamma \Rightarrow}\ (\mathrm{cut})$$

$$\frac{A, B, \Gamma \Rightarrow}{A \land B, \Gamma \Rightarrow}\ (\land l) \qquad \frac{A, \Gamma \Rightarrow \qquad B, \Gamma \Rightarrow}{A \lor B, \Gamma \Rightarrow}\ (\lor l)$$

$$\frac{A[t/x], \Gamma \Rightarrow}{\forall x\, A, \Gamma \Rightarrow}\ (\forall l) \qquad \frac{A, \Gamma \Rightarrow}{\exists x\, A, \Gamma \Rightarrow}\ (\exists l)$$

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

Left-Only Sequent Rules for $S4$

$$\frac{A, \Gamma \Rightarrow}{\Box A, \Gamma \Rightarrow}\ (\Box l) \qquad \frac{A, \Gamma^* \Rightarrow}{\Diamond A, \Gamma \Rightarrow}\ (\Diamond l)$$

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

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

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

Right-only system is precisely dual

Proving $\forall x\,(A \to B) \Rightarrow A \to \forall x\, B$

Left-only, NNF version: $A \land \exists x\, \neg B,\ \forall x\,(\neg A \lor B) \Rightarrow$

(x not free in A)

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A, \neg B, \neg A \Rightarrow} \qquad \dfrac{}{A, \neg B, B \Rightarrow}}{A, \neg B, \neg A \lor B \Rightarrow}\,(\lor l)}{A, \neg B, \forall x\,(\neg A \lor B) \Rightarrow}\,(\forall l)}{A, \exists x\, \neg B, \forall x\,(\neg A \lor B) \Rightarrow}\,(\exists l)}{A \land \exists x\, \neg B, \forall x\,(\neg A \lor B) \Rightarrow}\,(\land l)
$$

Adding Unification

Special class of *unknowns* $?x, ?y, ?z, \ldots$

Rule $(\forall l)$ now inserts an unknown:

$$\frac{A[?z/x], \Gamma \Rightarrow}{\forall x\, A, \Gamma \Rightarrow} \; (\forall l)$$

In $\neg A, B, \Gamma \Rightarrow$ try unifying A with B to make basic sequent

Updating an unknown affects *entire proof tree*

What about rule $(\exists l)$?

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 (*miniscope*)

Proving $\exists x\, \forall y\, [P(x) \to P(y)]$

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

Skolemize: $\forall x\, [P(x) \land \neg P(f(x))]$

Miniscope: $\forall x\, P(x) \land \forall x\, \neg P(f(x))$

A Proof of $\exists x\, \forall y\, [P(x) \to P(y)]$

$$
\dfrac{\dfrac{\dfrac{\dfrac{?y \mapsto f(?z)}{P(?y),\ \neg P(f(?z)) \Rightarrow}\ \text{basic}}{P(?y),\ \forall x\, \neg P(f(x)) \Rightarrow}\ (\forall l)}{\forall x\, P(x),\ \forall x\, \neg P(f(x)) \Rightarrow}\ (\forall l)}{\forall x\, P(x)\ \land\ \forall x\, \neg P(f(x)) \Rightarrow}\ (\land l)
$$

Unification chooses the term for $(\forall l)$

Longer proof from $\forall x\, [P(x) \land \neg P(f(x))] \Rightarrow$

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

$$\dfrac{\dfrac{\dfrac{\dfrac{?y \mapsto a}{\neg P(a),\ \neg Q(b),\ P(?y) \Rightarrow} \qquad \dfrac{?y \mapsto b???}{\neg P(a),\ \neg Q(b),\ Q(?y) \Rightarrow}}{\neg P(a),\ \neg Q(b),\ P(?y) \lor Q(?y) \Rightarrow}\,(\lor l)}{\neg P(a),\ \neg Q(b),\ \forall x\,[P(x) \lor Q(x)] \Rightarrow}\,(\forall l)}{\neg P(a) \land \neg Q(b),\ \forall x\,[P(x) \lor Q(x)] \Rightarrow}\,(\land l)$$

The World's Smallest Theorem Prover?

```
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).
```