Automated Reasoning

Lecture 2: Propositional Logic and Natural Deduction

Jacques Fleuriot
jdf@inf.ed.ac.uk

Logic Puzzles

1. Tomorrow will be sunny or rainy.
   Tomorrow will not be sunny.
   What will the weather be tomorrow?
2. I like classical or pop music.
   If I like classical music, then I am sophisticated.
   I don't like pop music.
   Am I sophisticated?
3. Fred bought milk or Fred bought lemonade.
   Fred bought milk or Fred bought water.
   Fred did not buy both water and lemonade.
   What did Fred buy?

Syntax of Propositional Logic

Propositional Logic represents the problems we have just seen by using **symbols** to represent (atomic) **propositions**.

These can be combined using the following **connectives**:

Name	symbol	usage
not	$\neg$	$\neg P$
and	$\wedge$	$P \wedge Q$
or	$\vee$	$P \vee Q$
implies	$\rightarrow$	$P \rightarrow Q$
if and only if	$\leftrightarrow$	$P \leftrightarrow Q$

precedence $\rightarrow$

Treat all binary connectives as right associative (following Isabelle)

Example

1. $(\text{SunnyTomorrow} \vee \text{RainyTomorrow}) \wedge (\neg\text{SunnyTomorrow})$
2. $(\text{Class} \vee \text{Pop}) \wedge (\text{Class} \rightarrow \text{Soph}) \wedge \neg\text{Pop}$
3. $(\text{M} \vee \text{L}) \wedge (\text{M} \vee \text{W}) \wedge \neg(\text{L} \wedge \text{W})$

Syntax and Ambiguity

The meanings of some statements can (appear to) be ambiguous:

$$\text{Class} \lor \text{Pop} \land \text{Class} \to \text{Soph} \to \neg\text{Pop}$$

We can use brackets (parentheses) to disambiguate a statement:

$$(\text{Class} \lor \text{Pop}) \land (\text{Class} \to \text{Soph} \to \neg\text{Pop})$$

Note that, based on our choice of precedence (on the previous slide),

$$A \lor B \land C \quad \text{denotes} \quad A \lor (B \land C)$$

Also note that implication is right associative, so:

$$P \to Q \to R \quad \text{denotes} \quad P \to (Q \to R)$$

Formal Syntax

A syntactically correct formula is called a **well-formed formula (wff)**

Given a (possible infinite) alphabet of propositional symbols $\mathcal{L}$, the set of wffs is the smallest set such that

- any symbol $\mathrm{A} \in \mathcal{L}$ is a wff;
- if P and Q are wffs, so are $\neg P$, $P \vee Q$, $P \wedge Q$, $P \rightarrow Q$, and $P \leftrightarrow Q$;
- if P is a wff, then (P) is a wff.

When we are interested in *abstract* syntax (tree-structure of formulas) rather than *concrete* syntax, we forget the last clause.

Semantics

Each wff is assigned a meaning or **semantics**, **T** or **F**, depending on whether its constituent wffs are assigned **T** or **F**.

Truth tables are one way to assign truth values to wffs.

P	Q	$P \wedge Q$
T	**T**	**T**
T	**F**	**F**
F	**T**	**F**
F	**F**	**F**

P	Q	$P \vee Q$
T	**T**	**T**
T	**F**	**T**
F	**T**	**T**
F	**F**	**F**

P	$\neg P$
T	**F**
F	**T**

P	Q	$P \rightarrow Q$
T	**T**	**T**
T	**F**	**F**
F	**T**	**T**
F	**F**	**T**

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**			
T	**F**			
F	**T**			
F	**F**			

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**		
T	**F**			
F	**T**			
F	**F**			

Semantics of the weather problem

1. Tomorrow will be sunny or rainy
2. Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**		
T	**F**	**T**		
F	**T**			
F	**F**			

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**		
T	**F**	**T**		
F	**T**	**T**		
F	**F**			

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**		
T	**F**	**T**		
F	**T**	**T**		
F	**F**	**F**		

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**	**F**	
T	**F**	**T**		
F	**T**	**T**		
F	**F**	**F**		

Semantics of the weather problem

1. Tomorrow will be sunny or rainy
2. Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**	**F**	
T	**F**	**T**	**F**	
F	**T**	**T**		
F	**F**	**F**		

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**	**F**	
T	**F**	**T**	**F**	
F	**T**	**T**	**T**	
F	**F**	**F**		

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**	**F**	
T	**F**	**T**	**F**	
F	**T**	**T**	**T**	
F	**F**	**F**	**T**	

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**	**F**	**F**
T	**F**	**T**	**F**	
F	**T**	**T**	**T**	
F	**F**	**F**	**T**	

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**	**F**	**F**
T	**F**	**T**	**F**	**F**
F	**T**	**T**	**T**	
F	**F**	**F**	**T**	

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \vee R$	② $\neg S$	① $\wedge$ ② $(S \vee R) \wedge \neg S$
T	**T**	**T**	**F**	**F**
T	**F**	**T**	**F**	**F**
F	**T**	**T**	**T**	**T**
F	**F**	**F**	**T**	

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**	**F**	**F**
T	**F**	**T**	**F**	**F**
F	**T**	**T**	**T**	**T**
F	**F**	**F**	**T**	**F**

Semantics of the weather problem

① Tomorrow will be sunny or rainy

② Tomorrow will not be sunny

What will the weather be tomorrow?

SunnyTomorrow	RainyTomorrow	① $S \lor R$	② $\neg S$	① $\land$ ② $(S \lor R) \land \neg S$
T	**T**	**T**	**F**	**F**
T	**F**	**T**	**F**	**F**
F	**T**	**T**	**T**	**T**
F	**F**	**F**	**T**	**F**

So it will rain tomorrow.

Semantics: Definitions

Definition (Interpretation)

An *interpretation* (or *valuation*) is a **truth assignment** to the symbols in the alphabet $\mathcal{L}$: it is a function V from $\mathcal{L}$ to $\{\mathbf{T}, \mathbf{F}\}$.

Semantics: Definitions

Definition (Interpretation)

An *interpretation* (or *valuation*) is a **truth assignment** to the symbols in the alphabet $\mathcal{L}$: it is a function V from $\mathcal{L}$ to $\{\mathbf{T}, \mathbf{F}\}$.

An interpretation V of $\mathcal{L}$ is extended to an interpretation of a wff P by induction on its structure:

$$
\begin{array}{lcl}
[\![\mathrm{A}]\!]_V & = & V(\mathrm{A}) \\
[\![P \wedge Q]\!]_V & = & [\![P]\!]_V \textbf{ and } [\![Q]\!]_V \\
[\![P \vee Q]\!]_V & = & [\![P]\!]_V \textbf{ or } [\![Q]\!]_V
\end{array}
\qquad
\begin{array}{lcl}
[\![\neg P]\!]_V & = & \textbf{not } [\![P]\!]_V \\
[\![P \rightarrow Q]\!]_V & = & [\![P]\!]_V \textbf{ implies } [\![Q]\!]_V
\end{array}
$$

This is the Tarski definition of truth: the truth value of a compound sentence is established by breaking it down until we get to atomic propositions.

Semantics: Definitions

Definition (Satisfaction)

An interpretation V **satisfies** a wff P if $[\![P]\!]_V = \mathbf{T}$

Definition (Satisfiable)

A wff is **satisfiable** if there exists an interpretation that satisfies it.
A wff is unsatisfiable if it is not satisfiable.

Definition (Valid or Tautology)

A wff is **valid** or is a **tautology** if every interpretation satisfies it.

Example

$(\mathrm{S} \vee \mathrm{R}) \wedge \neg \mathrm{S}$ is **satisfiable**

(there is a state of affairs that makes it true)

$((\mathrm{S} \vee \mathrm{R}) \wedge \neg \mathrm{S}) \rightarrow \mathrm{R}$ is **valid**

(it is always true, no matter what the state of affairs)

Semantics: Definitions

Definition (Entailment)

The wffs $P_1, P_2, \ldots, P_n$ ***entail*** Q if for any interpretation which satisfies all of $P_1, P_2, \ldots, P_n$ also satisfies Q.

We then write $P_1, P_2, \ldots, P_n \models Q$.

Note If there is ***no*** interpretation which satisfies all of $P_1, P_2, \ldots, P_n$ then $P_1, P_2, \ldots, P_n \models Q$ for ***any*** Q. Contradictory assumptions entail everything!

Note Everything entails a tautology. If Q is a tautology, then $P_1, P_2, \ldots, P_n \models Q$ holds not matter what $P_1, P_2, \ldots P_n$ are.

We write $\models Q$ when Q is a tautology.

Example

Is $\neg P, Q \models Q \wedge (P \rightarrow Q)$ a valid entailment?

Proof, Inference Rules and Deductive Systems

For propositional logic, it is possible to reason on a computer directly using the semantics:

- Satisfiability, validity and entailment are decidable

So, in theory, we make conjectures and the computer checks them.

But this is not always possible:

- Propositional logic is not very expressive; but
- Expressive logics like FOL and HOL are not decidable; and
- Even for propositional logic, checking satisfiability, validity, entailment is not always feasible.

So we encode a notion of *proof*:

- A formal deductive system is a set of **valid inference rules** that tell us what conclusions we can draw from some premises.
- Inference rules can be applied manually or automatically.
- We will look at Natural Deduction, developed by Gentzen and Prawitz.

Inference Rules

An inference rule tells us how one wff can be **derived** in one step from zero, one, or more other wffs. We write

$$\frac{P_1 \qquad P_2 \qquad \ldots \qquad P_n}{Q}\ (\mathrm{R})$$

if wff Q is derived from wffs $P_1, P_2, \ldots, P_n$ using the rule R.

Example inference rules (with their corresponding Isabelle names):

$$\frac{P \qquad Q}{P \wedge Q}\ (\mathrm{conjI}) \qquad\qquad \frac{P \qquad P \longrightarrow Q}{Q}\ (\mathrm{mp})$$

The P and Q here are **meta-variables** (denoted by $?P$ and $?Q$ in Isabelle) and mp is the **modus ponens** rule of inference.

This **rule schema** characterises an infinite number of **rule instances**, obtained by substituting wffs for the P and Q. Example of an instance of mp is:

$$\frac{\overbrace{A \wedge B}^{P} \qquad \overbrace{(A \wedge B)}^{P} \longrightarrow \overbrace{C}^{Q}}{\underbrace{C}_{Q}}$$

Validity

- Inference rules must be **valid**. They must preserve truth.
- Formally, for all instances of

$$\frac{P_1 \qquad P_2 \qquad \dots \qquad P_n}{Q}\ (\mathrm{R})$$

  of the rule R, we must have $P_1, P_2, \dots, P_n \vDash Q$.
- Inference is **transitive**. If we can infer R from Q and we can infer Q from P, then we can infer R from P. This means we can chain deductions together to form a deduction *tree*.

Introduction and Elimination

In Natural Deduction (ND), rules are split into two groups:

Introduction and Elimination

In Natural Deduction (ND), rules are split into two groups:

Introduction rules : how to derive $P \wedge Q$ and $P \vee Q$:

$$\frac{P \qquad Q}{P \wedge Q}\ (\text{conjI}) \qquad \frac{P}{P \vee Q}\ (\text{disjI1}) \qquad \frac{Q}{P \vee Q}\ (\text{disjI2})$$

Introduction and Elimination

In Natural Deduction (ND), rules are split into two groups:

Introduction rules : how to derive $P \wedge Q$ and $P \vee Q$:

$$\frac{P \qquad Q}{P \wedge Q}\ (\text{conjI}) \qquad \frac{P}{P \vee Q}\ (\text{disjI1}) \qquad \frac{Q}{P \vee Q}\ (\text{disjI2})$$

Elimination rules : what can be derived from $P \wedge Q$ and from $P \vee Q$?

$$\frac{P \wedge Q}{P}\ (\text{conjunct1}) \qquad \frac{P \wedge Q}{Q}\ (\text{conjunct2})$$

$$\frac{P \vee Q \qquad \begin{array}{c}[P]\\ \vdots \\ R\end{array} \qquad \begin{array}{c}[Q]\\ \vdots \\ R\end{array}}{R}\ (\text{disjE})$$

A proof: distributivity of $\wedge$ and $\vee$

A proof that $P \wedge (Q \vee R) \models (P \wedge Q) \vee (P \wedge R)$.

$$\dfrac{\dfrac{P \wedge (Q \vee R)}{Q \vee R} \qquad \dfrac{\dfrac{\dfrac{P \wedge (Q \vee R)}{P} \quad [Q]}{P \wedge Q}}{(P \wedge Q) \vee (P \wedge R)} \qquad \dfrac{\dfrac{\dfrac{P \wedge (Q \vee R)}{P} \quad [R]}{P \wedge R}}{(P \wedge Q) \vee (P \wedge R)}}{(P \wedge Q) \vee (P \wedge R)}$$

Note Each proof step will normally be annotated with the name of its associated inference rule (e.g. disjE for the bottom most step).

Ways of applying rules

Inference rules are applied in two basic ways.

- **Forward proof** if we derive new wffs from existing wffs by applying rules top down.
- **Backward proof** if we conjecture some wff true and apply rules bottom-up to produce new wffs from which the original wff is derived.

In Isabelle,

- procedural proof very often proceeds backwards, from the goal. Forward proof is also possible, though.
- structured proof tends to be via forward reasoning.

Summary

- Propositional logic
  - Syntax (atomic propositions, wffs) (H&R 1.1+1.3)
  - Semantics (interpretations, satisfication, satisfiability, validity, entailment) (H&R 1.4.1, first part of 1.4.2)
- Natural deduction (H&R 1.2)
  - Introduction and Elimination rules;
  - Proofs are trees, with assumptions at the leaves.
- Next time
  - More on Natural Deduction ($\rightarrow$, $\leftrightarrow$, $\neg$);
  - Natural deduction in Isabelle/HOL.