Automated Reasoning

Lecture 3: Natural Deduction and Starting with Isabelle

Jacques Fleuriot
jdf@inf.ed.ac.uk

Recap

- Last time I introduced **natural deduction**
- We saw the rules for $\wedge$ and $\vee$:

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

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

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

But what about the other connectives $\rightarrow$, $\leftrightarrow$ and $\neg$?

Rules for Implication

$$\dfrac{\begin{matrix}[P] \\ \vdots \\ Q\end{matrix}}{P \longrightarrow Q}\ (\text{impI})$$

impI forward: If on the assumption that P is true, Q can be shown to hold, then we can conclude $P \longrightarrow Q$.

impI backward: To prove $P \longrightarrow Q$, assume P is true and prove that Q follows.

$$\dfrac{P \longrightarrow Q \qquad P}{Q}\ (\text{mp})$$

The **modus ponens** rule.

$$\dfrac{P \longrightarrow Q \qquad P \qquad \begin{matrix}[Q] \\ \vdots \\ R\end{matrix}}{R}\ (\text{impE})$$

Another possible implication rule is this one. Note: this is not necessarily a standard ND rule but may be useful in mechanized proofs.

Rules for $\leftrightarrow$

$$\dfrac{\begin{matrix}[Q] \\ \vdots \\ P\end{matrix} \qquad \begin{matrix}[P] \\ \vdots \\ Q\end{matrix}}{P \leftrightarrow Q}\ (\text{iffI}) \qquad \dfrac{P \leftrightarrow Q \qquad P}{Q}\ (\text{iffD1})$$

$$\dfrac{P \leftrightarrow Q \qquad Q}{P}\ (\text{iffD2})$$

These rules are derivable from the rules for $\wedge$ and $\longrightarrow$, using the abbreviation $P \leftrightarrow Q \equiv (P \longrightarrow Q) \wedge (Q \longrightarrow P)$.

Note: In Isabelle, the $\leftrightarrow$ is also denoted by $=$

Rules for False and Negation

It is convenient to introduce a 0-ary connective $\bot$ to represent false. The connective $\bot$ has the rules:

no introduction rule for $\bot$

$$\frac{\bot}{P} \text{ (FalseE)}$$

Note $\bot$ is written `False` in Isabelle.

$$\frac{\begin{array}{c} P \\ \vdots \\ \bot \end{array}}{\neg P} \text{ (notI)} \qquad \frac{P \qquad \neg P}{\bot} \text{ (notE)}$$

Note: we could *define* $\neg P$ to be $P \rightarrow \bot$

Note: In Isabelle, notE is different:

$$\frac{P \qquad \neg P}{R} \text{ (notE)}$$

In this course, you can use either version in your proofs.

Proof

Recall the logic problems from lecture 2: we can now prove

$$((\text{Sunny} \lor \text{Rainy}) \land \neg\text{Sunny}) \to \text{Rainy}$$

which we previously knew only by semantic means.

$$\dfrac{\dfrac{\dfrac{[(S \lor R) \land \neg S]_1}{S \lor R} \qquad \dfrac{[S]_2 \qquad \dfrac{[(S \lor R) \land \neg S]_1}{\neg S}}{R} \qquad \dfrac{[R]_2}{R}}{R}\ (\text{disjE}_2)}{((S \lor R) \land \neg S) \to R}\ (\text{impI}_1)$$

The subscripts $[\cdot]_1$ and $[\cdot]_2$ on the assumptions refer to the rule instances (also with subscripts) where they are discharged. This makes the proof easier to follow.

Note: For a full proof, the names of *all* the ND rules being used should be given (i.e. not just impI and disjE as in the above).

Soundness and Completeness

Theorem (Soundness)

If Q is provable from assumptions $P_1, \ldots, P_n$, then $P_1, \ldots, P_n \models Q$.

This follows because all our rules are *valid*.

Is the converse true?

Can't prove Pierce's law: $((A \rightarrow B) \rightarrow A) \rightarrow A$

Can prove it using the *law of excluded middle*: $P \vee \neg P$.

So far, our proof system is sound and complete for Intuitionistic Logic. Intuitionistic logic rejects the law of excluded middle.

Rules for classical reasoning

$$\frac{}{\neg P \vee P}\ \text{(excluded\_middle)} \qquad \frac{\begin{array}{c}[\neg P]\\ \vdots \\ \bot\end{array}}{P}\ \text{(ccontr)}$$

Either one suffices.

Theorem (Completeness)

If $P_1, \ldots, P_n \models Q$, then Q is provable from the assumptions $P_1, \ldots, P_n$.

Proof: more complicated, see H&R 1.4.4.

Sequents

We have been representing proofs with assumptions like so:

$$\dfrac{\begin{array}{cccc} & P_2 & & \\ P_1 & \vdots & & P_n \\ \vdots & \vdots & \cdots & \vdots \end{array}}{Q}$$

Another notation is sequent-style or Fitch-style:

$$P_1, P_2, \ldots, P_n \vdash Q$$

The assumptions are usually collectively referred to using Γ:

$$\Gamma \vdash Q$$

This style is fiddlier on paper, but easier to prove meta-theoretic properties for, and easier to represent on a computer.

Natural Deduction Sequents

New rule: $$\frac{P \in \Gamma}{\Gamma \vdash P} \text{ (assumption)}$$

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

$$\frac{\Gamma \vdash P}{\Gamma \vdash P \vee Q} \text{ (disjI1)} \qquad \frac{\Gamma \vdash Q}{\Gamma \vdash P \vee Q} \text{ (disjI2)} \qquad \frac{\Gamma \vdash P \vee Q \qquad \Gamma, P \vdash R \qquad \Gamma, Q \vdash R}{\Gamma \vdash R} \text{ (disjE)}$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \longrightarrow B} \text{ (impI)} \qquad \frac{\Gamma \vdash A \longrightarrow B \qquad \Gamma \vdash A}{\Gamma \vdash B} \text{ (mp)}$$

No introduction rule for $\bot$

$$\frac{\Gamma \vdash \bot}{\Gamma \vdash P} \text{ (FalseE)}$$

$$\frac{\Gamma, P \vdash \bot}{\Gamma \vdash \neg P} \text{ (notI)} \qquad \frac{\Gamma \vdash P \qquad \Gamma \vdash \neg P}{\Gamma \vdash \bot} \text{ (notE)} \qquad \frac{}{\Gamma \vdash P \vee \neg P} \text{ (excluded\_middle)}$$

Natural Deduction in Isabelle/HOL

Isabelle represents the sequent $P_1, P_2, \ldots, P_n \vdash Q$ with the following notation:

$$P_1 \Longrightarrow (P_2 \Longrightarrow \ldots \Longrightarrow (P_n \Longrightarrow Q) \ldots)$$

which is also written as: $[\![P_1; P_2; \ldots; P_n]\!] \Longrightarrow Q$

Note: To enable the bracket notation for sequents in Isabelle, select: `Plugins` $\rightarrow$ `Plugin Options` in the Isabelle JEdit menu bar. Then select `Isabelle` $\rightarrow$ `General` and enter *brackets* in the `Print Mode` box.

The symbol $\Longrightarrow$ is *meta-implication*.

Meta-implication is used to represent the relationship between premises and conclusions of rules.

$$\dfrac{\begin{matrix}[P] \\ \vdots \\ Q\end{matrix}}{P \rightarrow Q} \quad \text{is written as} \quad (?P \Longrightarrow ?Q) \Longrightarrow (?P \rightarrow ?Q)$$

Natural Deduction Rules in Isabelle

A selection of natural deduction rules in Isabelle notation:

$$\frac{P \qquad Q}{P \wedge Q}\ (\text{conjI}) \qquad \llbracket ?P; ?Q \rrbracket \Longrightarrow ?P \wedge ?Q$$

$$\frac{P \wedge Q}{P}\ (\text{conjunct1}) \qquad \llbracket ?P \wedge ?Q \rrbracket \Longrightarrow ?P$$

$$\frac{P}{P \vee Q}\ (\text{disjI1}) \qquad \llbracket ?P \rrbracket \Longrightarrow ?P \vee ?Q$$

$$\frac{P \vee Q \qquad \begin{matrix}[P]\\ \vdots \\ R\end{matrix} \qquad \begin{matrix}[Q]\\ \vdots \\ R\end{matrix}}{R}\ (\text{disjE}) \qquad \llbracket ?P \vee ?Q; ?P \Longrightarrow ?R; ?Q \Longrightarrow ?R \rrbracket \Longrightarrow ?R$$

Doing Proofs in Isabelle: Theory Set-up

Syntax:

`theory` *MyTh*
`imports` $T_1 \ldots T_n$
`begin`
(definitions, theorems, proofs, ...)*
`end`

MyTh: name of theory. Must live in file *MyTh*`.thy`

T_i: names of *imported* theories. Import is transitive.

Often: `imports Main`

Doing Proofs in Isabelle

A declaration like so enters proof mode:

theorem K: "$A \longrightarrow B \longrightarrow A$"

Isabelle responds:

proof (prove)

goal (1 subgoal):
1. $A \longrightarrow B \longrightarrow A$

We now apply proof methods (tactics) that affect the subgoals. Either:

- generate new subgoal(s), breaking the problem down; or
- solve the subgoal

When there are no more subgoals, then the proof is complete.

The `assumption` Method

Given a subgoal of the form:

$$[\![A; B]\!] \Longrightarrow A$$

This subgoal is solvable because we want to prove A under the assumption that A is true.

We can solve this subgoal using the `assumption` method:

```
apply assumption
```

The `rule` Method

To apply an inference rule backward, we use `rule`.

Consider the theorem `disjI1`

$$?P \Longrightarrow ?P \vee ?Q$$

Using the command

apply (rule disjI1)

on the goal

$$[\![A; B; C]\!] \Longrightarrow (A \wedge B) \vee D$$

yields the subgoal

$$[\![A; B; C]\!] \Longrightarrow A \wedge B$$

Using `rule` can be viewed as a way of breaking down the problem into subproblems.

Matching and Unification

In applying rule (with the ? in front of variables omitted)

$$P \Longrightarrow P \vee Q$$

to goal

$$[\![A; B; C]\!] \Longrightarrow (A \wedge B) \vee D$$

The pattern $P \vee Q$ is **matched** with the target $(A \wedge B) \vee D$ to yield the instantiations $P \mapsto A \wedge B,\ \ Q \mapsto D$ which make the pattern and target the same. The following goal results

$$[\![A; B; C]\!] \Longrightarrow A \wedge B$$

In general, if the goal conclusion contains schematic variables, the rule and goal conclusions are **unified** i.e. both are instantiated so as to make them the same.

More on **unification** later!

Summary

- More natural deduction (H&R 1.2, 1.4)
  - The rules for $\rightarrow$, $\leftrightarrow$ and $\neg$
  - Rules for classical reasoning
  - Soundness and completeness properties
  - Sequent-style presentation
- Starting with proofs in Isabelle
- Next time:
  - More on using Isabelle to do proofs
  - N-style vs. L-style proof systems