Automated Reasoning

Lecture 4: Propositional Reasoning in Isabelle

Jacques Fleuriot
jdf@inf.ed.ac.uk
Recap

Last lecture:

- Completed the natural deduction system for propositional logic
- Started on proving propositions in Isabelle

Today:

- More details on proving propositions in Isabelle
- Alternative inference rules (L-system, a.k.a. “Sequent Calculus”)
- Why should we trust Isabelle?
The rule Method

To apply an inference rule, we use rule.

Consider the theorem disjI1

\[ ?P \implies ?P \lor ?Q \]

Using the command

apply (rule disjI1)

on the goal

\[ [A; B; C] \implies (A \land B) \lor D \]

yields the subgoal

\[ [A; B; C] \implies A \land B \]
General definition of method rule

When we apply the method rule \texttt{someRule} where

\[
\texttt{someRule} : \{P_1; \ldots; P_m\} \rightarrow Q
\]

to the goal

\[
\{A_1; \ldots; A_n\} \rightarrow C
\]

where \(Q\) and \(C\) can be unified, we generate the goals

\[
\{A'_1; \ldots; A'_n\} \rightarrow P'_1
\]

\[
\vdots
\]

\[
\{A'_1; \ldots; A'_n\} \rightarrow P'_m
\]

where \(A'_1, A'_2, \ldots, A'_n, P'_1, P'_2, \ldots, P'_m\) are the results of applying the substitution which unifies \(Q\) and \(C\) to \(A_1, A_2, \ldots, A_n, P_1, P_2, \ldots, P_m\).

We must now derive each of the rule’s assumptions using our goal’s assumptions.
A Problem with rule

Consider the disjE rule:

\[
\text{disjE: } [P \lor Q; P \Rightarrow R; Q \Rightarrow R] \Rightarrow R
\]

If we have the goal:

\[
[(A \land B) \lor C; D] \Rightarrow B \lor C
\]

Then applying rule disjE produces three new goals:

\[
[(A \land B) \lor C; D] \Rightarrow ?P \lor ?Q
\]
\[
[(A \land B) \lor C; D; ?P] \Rightarrow B \lor C
\]
\[
[(A \land B) \lor C; D; ?Q] \Rightarrow B \lor C
\]

We then solve the first subgoal by applying assumption.

This seems pointlessly roundabout... we often want to use one of our assumptions in our proof.
The erule Method

Used when the conclusion of theorem matches the conclusion of the current goal and the first premise of theorem matches a premise of the current goal.

Consider the theorem disjE

\[ [P \lor Q; P \implies R; Q \implies R] \implies R \]

Applying erule disjE to goal

\[ [(A \land B) \lor C; D] \implies B \lor C \]

yields the subgoals

\[ [D; (A \land B)] \implies B \lor C \quad [D; C] \implies B \lor C \]
General definition of method erule

When we apply the method erule someRule where

\[ \text{someRule}: [P_1; \ldots; P_m] \implies Q \]

to the goal

\[ [A_1; \ldots; A_n] \implies C \]

where \( P_1 \) and \( A_1 \) are unifiable and \( Q \) and \( C \) are unifiable, we generate the goals:

\[ [A'_2; \ldots; A'_n] \implies P'_2 \]

\[ \vdots \]

\[ [A'_2; \ldots; A'_n] \implies P'_m \]

where \( A'_2, \ldots, A'_n, P'_2, \ldots, P'_m \) are the results of applying the substitution which unifies \( P_1 \) to \( A_1 \) and \( Q \) to \( C \) to \( A_2, \ldots, A_n, P_2, \ldots, P_m \).

We eliminate an assumption from the rule and the goal, and must derive the rule’s other assumptions using our goal’s other assumptions.
General definition of method \texttt{drule}

When we apply the method \texttt{drule} \texttt{someRule} where

\[
\texttt{someRule: } [P_1; \ldots; P_m] \implies Q
\]

to the goal

\[
[A_1; \ldots; A_n] \implies C
\]

where \(P_1\) and \(A_1\) are unifiable, we generate the goals:

\[
[A'_2; \ldots; A'_n] \implies P'_2
\]

\[
\vdots
\]

\[
[A'_2; \ldots; A'_n] \implies P'_m
\]

\[
[Q'; A'_2; \ldots; A'_n] \implies C'
\]

where \(A'_2, A'_3, \ldots, A'_n, P'_2, P'_3 \ldots, P'_m, Q', C'\) are the results of applying the substitution which unifies \(P_1\) and \(A_1\) to \(A_2, A_3, \ldots, A_n, P_2, P_3, \ldots, P_m, Q, C\).

We delete an assumption, replacing it with the conclusion of the rule.
General definition of method \texttt{frule}

When we apply the method \texttt{frule} \texttt{someRule} where

\[
\texttt{someRule} : [P_1; \ldots; P_m] \rightarrow Q
\]

to the goal

\[
[A_1; \ldots; A_n] \rightarrow C
\]

where \(P_1\) and \(A_1\) are unifiable, we generate the goals:

\[
[A_1'; A_2'; \ldots; A_n'] \rightarrow P_2'
\]

\[
[A_1'; A_2'; \ldots; A_n'] \rightarrow P_2'
\]

\[
\vdots
\]

\[
[A_1'; A_2'; \ldots; A_n'] \rightarrow P_m'
\]

\[
[Q'; A_1'; A_2'; \ldots; A_n'] \rightarrow C'
\]

where \(A_1', A_2', \ldots, A_n', P_2', \ldots, P_m', Q', C'\) are the results of applying the substitution which unifies \(P_1\) and \(A_1\) to \(A_1, A_2, \ldots, A_n, P_2, \ldots, P_m, Q, C\).

This is like \texttt{drule} except the assumption in our goal is kept.
More Methods

▶ rule_tac, erule_tac, drule_tac and frule_tac are like their counterparts, but you can give substitutions for variables in the rule before they are applied.

Example

apply (erule_tac Q="B \land D" in conjE)

applied to the subgoal

\[[A \land B; C \land B \land D] \implies B \land D\]

generates the new goal

\[[A \land B; C; B \land D] \implies B \land D\]

▶ Isabelle also provides advanced tactics, like simp and auto which perform some automatic deduction.
The erule tactic points to another way of phrasing a system of inference rules in a system with sequents $\Gamma \vdash A$.

Instead of \textit{elimination} rules:

\[
\begin{align*}
\Gamma \vdash P \lor Q & \quad \Gamma, P \vdash R & \quad \Gamma, Q \vdash R \\
\hline
\Gamma \vdash R
\end{align*}
\]

\textit{(disjE)}

Have \textit{left introduction rules} (all the introduction rules in natural deduction introduce connectives on the right-hand side of the $\vdash$):

\[
\begin{align*}
\Gamma, P \vdash R & \quad \Gamma, Q \vdash R \\
\hline
\Gamma, P \lor Q \vdash R
\end{align*}
\]

This corresponds to applying rules using \texttt{erule} in Isabelle.

The \textit{left introduction rules} are often much easier to use in a backwards, goal-directed style.
The following L-System (a.k.a. Sequent Calculus) rules are an alternative sound and complete proof system for propositional logic:

\[
\frac{}{\Gamma, P \vdash P} \quad \text{(assumption)}
\]

\[
\frac{\Gamma \vdash P \quad \Gamma \vdash Q}{\Gamma \vdash P \land Q} \quad \text{(conjI)}
\]

\[
\frac{}{\Gamma, P, Q \vdash R} \quad \text{(e conjE)}
\]

\[
\frac{\Gamma \vdash P}{\Gamma \vdash P \lor Q} \quad \text{(disjI1)}
\]

\[
\frac{\Gamma \vdash Q}{\Gamma \vdash P \lor Q} \quad \text{(disjI2)}
\]

\[
\frac{}{\Gamma, P \lor Q \vdash R} \quad \text{(e disjE)}
\]

\[
\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \quad \text{(impI)}
\]

\[
\frac{\Gamma \vdash P \quad \Gamma, Q \vdash R}{\Gamma, P \rightarrow Q \vdash R} \quad \text{(e impE)}
\]

no right-intro rule for \( \bot \)

\[
\frac{\Gamma, P \vdash \bot}{\Gamma \vdash \neg P} \quad \text{(notI)}
\]

\[
\frac{\Gamma, P, \neg P \vdash R}{\Gamma \vdash \neg P \lor \neg \neg P} \quad \text{excluded_middle}
\]

\[
\frac{}{\Gamma, \bot \vdash P} \quad \text{(e FalseE)}
\]

Note: \texttt{e someRule} is short for \texttt{erule someRule}.

Note: in the above presentation left-hand-sides are \textit{sets} of formulas.
An Old Friend Revisited

\[
\begin{align*}
S, \neg S & \vdash R \quad \text{(e notE)} & & R, \neg S \vdash R \quad \text{(assumption)} \\
&S, \neg S \vdash R \quad \text{(e disjE)} & & \neg S \vdash R \quad \text{(e ConjE)} \\
&(S \lor R), \neg S \vdash R \quad \text{(impI)} & & (S \lor R) \land \neg S \vdash R
\end{align*}
\]
Re-using proofs: The Cut rule

So far, all proofs have been self-contained; they have only used the pre-existing rules of inference.

By the completeness theorem, this suffices to prove everything that is true, but can lead to extremely repetitive proofs.
Re-using proofs: The Cut rule

So far, all proofs have been self-contained; they have only used the pre-existing rules of inference.

By the completeness theorem, this suffices to prove everything that is true, but can lead to extremely repetitive proofs.

The cut rule: (we “cut” \( P \) into the proof)

\[
\Gamma \vdash P \quad \Gamma, P \vdash Q
\]

\[
\frac{}{\Gamma \vdash Q}
\]

allows the use of a lemma \( P \) in a proof of \( Q \). We can now reuse \( P \) multiple times in the proof of \( Q \).
Re-using proofs: The Cut rule

So far, all proofs have been self-contained; they have only used the pre-existing rules of inference.

By the completeness theorem, this suffices to prove everything that is true, but can lead to extremely repetitive proofs.

The cut rule: (we “cut” \( P \) into the proof)

\[
\frac{\Gamma \vdash P \quad \Gamma, P \vdash Q}{\Gamma \vdash Q}
\]

allows the use of a \textit{lemma} \( P \) in a proof of \( Q \). We can now reuse \( P \) multiple times in the proof of \( Q \).

In Isabelle:

- `cut_tac lemmaName` — adds the conclusion of `lemmaName` as a new assumption, and its assumptions as new subgoals
- `subgoal_tac P` — adds \( P \) as a new assumption, and introduces \( P \) as a new subgoal.
Why should you believe Isabelle?

When Isabelle says “No subgoals!” why should we believe that we have *really* proved something? Is Isabelle sound?

It is doing non-trivial work behind the scenes: unification, rewriting, maintaining a database of theorems+assumptions, automatic proof.

Isabelle uses two strategies to maintain soundness:

▶ A small trusted kernel: internally, every proof is broken down into primitive rule applications which are checked by a small piece of hand-verified code. This is the “LCF” model. So new proof procedures cannot introduce unsoundness.

▶ Encourages definitional extension of the logic: new concepts are introduced by definition rather than axiomatisation (more on this in Lecture 6). So new definitions cannot introduce unsoundness.

Threats to (practical) soundness still exist, including: Have we proved what we thought we proved? Are the formulas displayed on screen correctly? …

Why should you believe Isabelle?

When Isabelle says “No subgoals!” why should we believe that we have really proved something? Is Isabelle sound?

It is doing non-trivial work behind the scenes: unification, rewriting, maintaining a database of theorems+assumptions, automatic proof.

Isabelle uses two strategies to maintain soundness:

- **A small trusted kernel**: internally, every proof is broken down into primitive rule applications which are checked by a small piece of hand-verified code. This is the “LCF” model. So new proof procedures cannot introduce unsoundness.

- **Encourages definitional extension of the logic**: new concepts are introduced by definition rather than axiomatisation (more on this in Lecture 6). So new definitions cannot introduce unsoundness.
Why should you believe Isabelle?

When Isabelle says “No subgoals!” why should we believe that we have really proved something? Is Isabelle sound?

It is doing non-trivial work behind the scenes: unification, rewriting, maintaining a database of theorems+assumptions, automatic proof.

Isabelle uses two strategies to maintain soundness:

▶ **A small trusted kernel**: internally, every proof is broken down into primitive rule applications which are checked by a small piece of hand-verified code. This is the “LCF” model. So new proof procedures cannot introduce unsoundness.

▶ Encourages **definitional extension of the logic**: new concepts are introduced by definition rather than axiomatisation (more on this in Lecture 6). So new definitions cannot introduce unsoundness.

Threats to (practical) soundness still exist, including: Have we proved what we thought we proved? Are the formulas displayed on screen correctly? ...

Summary

- More tools for proving propositions in Isabelle
  - The erule, drule, frule methods
  - Their —_tac variants
  - $L$-systems, and Cut rules (cut_tac, subgoal_tac).
  - See the propositional logic exercises and examples:
    - Tutorial 1 and Additional Exercise on the AR webpage;
    - The Isabelle theory file Prop.thy;
    - Start using Isabelle (if you haven’t done so already).

- How Isabelle maintains soundness
  - Small trusted kernel
  - Definitional extension instead of axiomatic extension

- Next time:
  - First-Order Logic: $\forall x.P$ and $\exists x.P$