FIRST ORDER PREDICATE LOGIC

Formal Reasoning

Wolfgang Schreiner and Wolfgang Windsteiger
Wolfgang.(Schreiner|Windsteiger)@risc.jku.at
Research Institute for Symbolic Computation (RISC)
Johannes Kepler University (JKU), Linz, Austria
http://www.risc.jku.at

What is Formal Reasoning?

- Problem: how to show that the statement $F_1, \ldots, F_n \models G$ is true?
  - □ Is formula G true in every model in which the $F_1, \ldots, F_n$ are true?
  - □ $F_1, \ldots, F_n$: the “axioms” that characterize the considered models.
  - □ G: a “conjecture” that might be also true in all these models.
  - □ If the conjecture is indeed true, then G is actually a “theorem”.
  - □ Since there are infinitely models, how to ensure that G is a theorem?
- Solution: derive the “sequent” $F_1, \ldots, F_n \vdash G$ by a proof calculus.
  - □ “Sequent” $F_1, \ldots, F_n \vdash G$: “assumptions” $F_1, \ldots, F_n$ and “goal” G.
  - □ Proof calculus: a set of “inference rules” that derive sequents.
  - □ “Soundness”: if a sequent $F_1, \ldots, F_n \vdash G$ can be derived, then $F_1, \ldots, F_n \models G$.
  - □ Inference rules only depend on syntactic structure of the formulas (not on semantics).
- Proof: a “proof” is a derivation of a sequent $F_1, \ldots, F_n \vdash G$.
  - □ Since inference rule are syntactic, correctness of a proof can be mechanically checked.

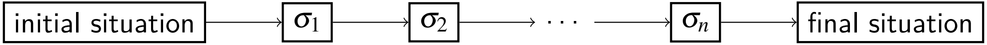
Formal reasoning is the demonstration of truth by checkable proofs.

What is a Proof?

We call sequents also proof situations.

Proof of G:

- Series of logical arguments that ascertain that G is true (under certain assumptions $F_1, \ldots, F_n$).
- Chain of proof situations.
- Proof situation captures “status” during a proof.
- Transition between situations corresponds to logical argument.



Formal Reasoning: How to Construct a Proof?

Forward interpretation:

A proof starts from trivial proof situations (obviously true),

Backward interpretation:

A proof starts from the goal to be proved,

progresses step-by-step

until it reaches the final situation, where the goal is proved.

until it reaches trivial proof situations (obviously true).

Formal Reasoning: How to Construct a Proof?

Individual proof steps are guided by inference rules, which are denoted as

$$\begin{array}{cl} S_1 \qquad \dots \qquad S_n & \longleftarrow \text{premises} \\ \hline S \quad \longleftarrow & \text{conclusion} \end{array}$$

where $S_1, \dots, S_n$ and S are sequents (proof situations).

Forward interpretation:

If $S_1, \dots, S_n$ can be proved,
then also S can be proved.

Backward interpretation:

In order to prove S,
we need to prove $S_1, \dots, S_n$.

Example (Inference Rules)

Let S_1 and S be the proof situations

$$\begin{aligned} S_1 &:= \text{“the assumption } A \text{ leads to a contradiction”} \\ S &:= \text{“the statement } \neg A \text{ holds”.} \end{aligned}$$

Rule: $\dfrac{S_1}{S}$

Forward interpretation: If the assumption A leads to a contradiction then $\neg A$ holds.

Backward interpretation: If we want to prove $\neg A$ then we assume A and derive a contradiction.

Example (Inference Rules)

Let S_1, S_2, and S be the proof situations

$$\begin{aligned} S_1 &:= \text{“the statement } A \text{ holds”} \\ S_2 &:= \text{“the statement } B \text{ holds”} \\ S &:= \text{“the statement } A \wedge B \text{ holds”.} \end{aligned}$$

Rule: $\dfrac{S_1 \qquad S_2}{S}$

Forward interpretation: If we can prove A and we can prove B then we can also prove $A \wedge B$.

Backward interpretation: If we want to prove $A \wedge B$ then we have to prove both A and B.

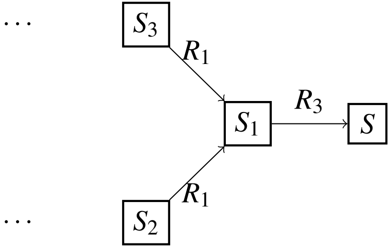
Example (Proof)

$S, S_1, \ldots, S_6$: sequents. Consider the following inference rules:

$$R_1\colon \frac{S_2 \qquad S_3}{S_1} \qquad R_2\colon \frac{}{S_4} \qquad R_3\colon \frac{S_1}{S} \qquad R_4\colon \frac{}{S_5}$$

$$R_5\colon \frac{S_4 \qquad S_5}{S_2} \qquad R_6\colon \frac{}{S_6} \qquad R_7\colon \frac{S_6}{S_3}$$

We want to prove S. Construct chain of sequents ending in S . . .



Example (Proof)

$S, S_1, \ldots, S_6$: sequents. Consider the following inference rules:

$$R_1\colon \frac{S_2 \qquad S_3}{S_1} \qquad R_2\colon \frac{}{S_4} \qquad R_3\colon \frac{S_1}{S} \qquad R_4\colon \frac{}{S_5}$$

$$R_5\colon \frac{S_4 \qquad S_5}{S_2} \qquad R_6\colon \frac{}{S_6} \qquad R_7\colon \frac{S_6}{S_3}$$

We want to prove S. Instead construct tree of sequents . . .

$$R_3\colon \dfrac{R_1\colon \dfrac{R_5\colon \dfrac{R_2\colon \dfrac{}{S_4} \qquad R_4\colon \dfrac{}{S_5}}{S_2} \qquad R_7\colon \dfrac{R_6\colon \dfrac{}{S_6}}{S_3}}{S_1}}{S}$$

Proof Trees

A formal proof can be seen as a finite tree, where

1. every node is a sequent,
2. if $S_1, \ldots, S_n$ are the children nodes of a node S, then there must be an inference rule of the form $\dfrac{S_1 \quad \ldots \quad S_n}{S}$.

Special case $n = 0$: A leaf has 0 children, hence

for every leaf S in the tree there must be a rule $\dfrac{}{S}$.

A formal proof of R is a proof tree with root R.

A Sketch of a Simple Proof Generation Procedure

Input: R **Output:** P such that P is a formal proof of R.

```
$P :=$ tree containing only the root node $R$
$Q := \{R\}$ // the leaves of the tree, i.e., the still “open” proof situations
while $Q \neq \emptyset$
    choose $T \in Q$ and choose a rule $\frac{S_1 \; \ldots \; S_n}{S}$ such that $S = T$
    replace $T$ in $Q$ by $S_1, \ldots, S_n$
    add $S_1, \ldots, S_n$ as children nodes of $T$ in $P$
return $P$
```

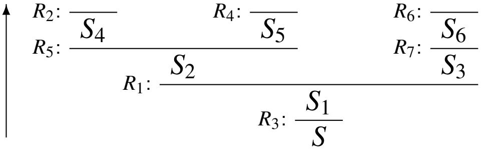
Depending on 1) the rules and 2) the choice of the rule in the loop, the procedure might not terminate.

Proving is the art of selecting the “right” rule applications to elaborate, from the desired goal as a root, a complete proof tree.

Proof Generation vs. Proof Presentation

Proof generation: start with sequent to be proved, then work backwards.

Read and apply rules from bottom to top.



Backward style proof presentation: In order to prove S, by R_3, we have to prove S_1. For this, by R_1, we have to

1. prove S_2: by R_5 we have to prove S_4 and S_5, which are guaranteed by R_2 and R_4, respectively. Now we still have to
2. prove S_3: by R_7 it is sufficient to prove S_6, which we know from R_6. QED (“quod erat demonstrandum”, “what was to be proved”).

Proof Generation vs. Proof Presentation

Proof presentation: often done in forward reasoning style, i.e. start with known facts and work forward until the sequent to be proved is reached.

Read and apply rules from top to bottom.

$$\Big\downarrow \qquad R_3\colon \dfrac{R_1\colon \dfrac{R_5\colon \dfrac{R_2\colon \dfrac{}{S_4} \qquad R_4\colon \dfrac{}{S_5}}{S_2} \qquad R_7\colon \dfrac{R_6\colon \dfrac{}{S_6}}{S_3}}{S_1}}{S}$$

Forward style proof presentation: We know S_4 and S_5 can be proved, hence by R_5, S_2 can be proved. Furthermore we know that S_6 can be proved, hence by R_7, also S_3 can be proved. Together with S_2, by R_1, we know that S_1 can be proved, and therefore, by R_3, also S. QED.

Proofs are always generated backwards (even if they are later presented in the forward style).

Inference Rules: Patterns and Matching

- Inference rules: schematic "patterns" to be "matched" against proof situations.

$$\text{name:}\quad \frac{K_1 \ldots \vdash G_1 \qquad \ldots \qquad K_n \ldots \vdash G_n}{K \ldots \vdash G}$$

  - □ n premises and one conclusion (special case $n = 0$: no premises).
- Premises/conclusions: "patterns" of sequents with schematic variables:
  - □ A "sequence variable" (e.g. $K\ldots$): an arbitrary sequence of formulas.
  - □ A "formula variable" (e.g. F, G, etc.): an arbitrary formula,
  - □ A "term variable" (e.g. t, u, etc.): an arbitrary term.
    - $K \ldots \vdash G_1 \wedge G_2$ matches a sequent whose goal is an arbitrary conjunction, i.e., a formula with "$\wedge$" as the outermost symbol and two subformulas G_1 and G_2.
- Order of assumptions: does not matter.
  - □ $K \ldots, F_1 \wedge F_2 \vdash G$ matches if a conjunction occurs among the assumptions.
- Multiple occurrences of the same variable: denote the same expression.
  - □ $K \ldots, F \wedge G \vdash G$ matches if among the assumptions a conjunction occurs whose second subformula is identical to the goal.

Example: Patterns and Matching

Let a be a constant, f, g unary function symbols, p, q, r, s unary predicate symbols, and t a schematic term variable.

Consider the following inference rules:

$$R_1\colon \frac{\vdash r(t) \qquad \vdash s(g(t))}{\vdash r(f(t))} \qquad R_2\colon \frac{}{\vdash p(a)} \qquad R_3\colon \frac{\vdash r(f(t))}{\vdash s(t)} \qquad R_4\colon \frac{}{\vdash q(a)}$$

$$R_5\colon \frac{\vdash p(t) \qquad \vdash q(t)}{\vdash r(t)} \qquad R_6\colon \frac{}{\vdash s(f(a))} \qquad R_7\colon \frac{\vdash s(f(t))}{\vdash s(g(t))}$$

We want to derive $\vdash s(a)$.

$$R_3\colon \dfrac{R_1\colon \dfrac{R_5\colon \dfrac{R_2\colon \dfrac{}{\vdash p(a)} \qquad R_4\colon \dfrac{}{\vdash q(a)}}{\vdash r(a)} \qquad R_7\colon \dfrac{R_6\colon \dfrac{}{\vdash s(f(a))}}{\vdash s(g(a))}}{\vdash r(f(a))}}{\vdash s(a)}$$

Proof Rules for Predicate Logic

One could give a (minimal) set of inference rules for first order predicate logic, which can be shown to be sound and complete, i.e.,

1. every formula, which has a formal proof, is also semantically true and
2. every semantically true formula has a formal proof.

$\leadsto$ e.g., sequent calculus, Gentzen calculus, natural deduction calculus, etc.

However, we give proof rules that help in practical proving of mathematical statements and checking of given proofs (differences lie in details only).

- propositional rules: closing rules, structural rules, connective rules.
- predicate logic rules: equality rules, quantifier rules.

For every logical connective and every quantifier, we give at least one rule, where the symbol occurs as the outermost symbol in the goal or one of the assumptions.

Closing Rules

For closing a proof, we need inference rules without premises.

- If the goal is among the assumptions, the goal can be proved.

$$\text{GoalAssum: } \frac{}{K \ldots, G \vdash G}$$

- If the assumptions are contradictory, any goal can be proved.

$$\text{ContrAssum: } \frac{}{K \ldots, A, \neg A \vdash G}$$

- If the assumptions include “false”, any goal can be proved.

$$\text{FalseAssum: } \frac{}{K \ldots, \bot \vdash G}$$

The leaves of a proof tree are constructed by application of these rules.

Structural Rules

- Any assumption may be dropped:

$$\text{Drop: } \frac{K \ldots \vdash G}{K \ldots, A \vdash G}$$

- Any assumption may be added, if it is also proved (the “cut rule”):

$$\text{Cut: } \frac{K \ldots \vdash A \qquad K \ldots, A \vdash G}{K \ldots \vdash G}$$

  We have to prove G. First we prove A: .... Now we prove G with the additional assumption A.

- Rather than proving G, we may assume $\neg G$ and derive a contradiction (an “indirect proof”):

$$\text{Indirect: } \frac{K \ldots, \neg G \vdash \bot}{K \ldots \vdash G}$$

  We have to prove G. Thus we may assume $\neg G$ and derive a contradiction.

Connective Rules: Negation

- Prove a negation as goal:

  $$\text{P-}\neg\text{:}\ \frac{K\ldots, G \vdash \bot}{K\ldots \vdash \neg G}$$

  We have to prove $\neg G$. Thus we may assume G and derive a contradiction.

- Use a negation as an assumption:

  $$\text{A-}\neg\text{:}\ \frac{K\ldots, \neg G \vdash A}{K\ldots, \neg A \vdash G}$$

  We know $\neg A$ and have to prove G. Thus we may assume $\neg G$ and prove A.

Example

$$\text{P-}\neg\text{:}\ \frac{\sqrt{2} \in \mathbb{Q}, \ldots \vdash \bot}{\ldots \vdash \sqrt{2} \notin \mathbb{Q}}$$

We have to prove that $\sqrt{2}$ is not rational. We do a proof by contradiction, hence, we assume that $\sqrt{2}$ was rational and derive a contradiction.

Connective Rules: Conjunction

- Prove a conjunction as a goal:

  $$\text{P-}\wedge\text{:}\ \frac{K \ldots \vdash F_1 \qquad K \ldots \vdash F_2}{K \ldots \vdash F_1 \wedge F_2}$$

  We have to prove $F_1 \wedge F_2$. First we prove F_1: .... Now we prove F_2: ....

- Use a conjunction as an assumption:

  $$\text{A-}\wedge\text{:}\ \frac{K \ldots, F_1, F_2 \vdash G}{K \ldots, F_1 \wedge F_2 \vdash G}$$

  We know $F_1 \wedge F_2$, therefore we know F_1 and we know F_2.

Connective Rules: Disjunction

- Prove a disjunction as a goal:

$$\text{P-}\lor\text{:}\ \frac{K\ldots,\neg F_1 \vdash F_2}{K\ldots \vdash F_1 \lor F_2} \qquad\qquad \text{P-}\lor\text{:}\ \frac{K\ldots,\neg F_2 \vdash F_1}{K\ldots \vdash F_1 \lor F_2}$$

  We have to prove $F_1 \lor F_2$. Thus we may assume $\neg F_1$ and prove $\neg F_2$. (or: Thus we may assume $\neg F_2$ and prove $\neg F_1$).

- Use a disjunction as an assumption ("proof by cases"):

$$\text{A-}\lor\text{:}\ \frac{K\ldots,F_1 \vdash G \qquad K\ldots,F_2 \vdash G}{K\ldots,F_1 \lor F_2 \vdash G}$$

  We know $F_1 \lor F_2$. We proceed by case distinction. Case F_1: .... Case F_2: ....

Example

$$\text{A-}\vee\text{:}\ \frac{\dfrac{P_1}{even(m) \vdash G} \qquad \dfrac{P_2}{odd(m) \vdash G}}{even(m) \vee odd(m) \vdash G}$$

We already know that m is even or m is odd. Thus, we can distinguish the two cases:

1. m is even: . . . (insert proof P_1 here)
2. m is odd: . . . (insert proof P_2 here)

Connective Rules: Implication

- Prove implication as a goal.

$$\text{P-}\!\rightarrow\!: \frac{K \ldots, F_1 \vdash F_2}{K \ldots \vdash F_1 \rightarrow F_2}$$

We have to prove $F_1 \rightarrow F_2$. Thus we may assume F_1 and prove F_2.

Connective Rules: Implication

- Use an implication as an assumption:

  $$\text{A-}\rightarrow\text{:}\ \frac{K \ldots \vdash F_1 \qquad K \ldots, F_2 \vdash G}{K \ldots, F_1 \rightarrow F_2 \vdash G}$$

  We know $F_1 \rightarrow F_2$. First we prove F_1: .... Now we know F_2.

- Often used instead: “modus ponens” and “modus tollens”

  $$\text{MP:}\ \frac{K \ldots, F_1, F_2 \vdash G}{K \ldots, F_1 \rightarrow F_2, F_1 \vdash G}$$

  We know $F_1 \rightarrow F_2$ and we know F_1. Therefore we know F_2.

  $$\text{MT:}\ \frac{K \ldots, \neg F_2, \neg F_1 \vdash G}{K \ldots, F_1 \rightarrow F_2, \neg F_2 \vdash G}$$

  We know $F_1 \rightarrow F_2$ and we know $\neg F_2$. Therefore we know $\neg F_1$.

Example

Prove $((A \to (B \vee C)) \wedge \neg C) \to (A \to B)$,

where A, B, and C are abbreviations for complex predicate logic formulas.

Develop proof tree top-down with root on top (convenient in practice).

$$
\begin{array}{rc}
 & \vdash ((A \to (B \vee C)) \wedge \neg C) \to (A \to B) \\
\text{P-}\!\to\!: & \overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad} \\
 & (A \to (B \vee C)) \wedge \neg C \vdash A \to B \\
\text{A-}\wedge\text{:} & \overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad} \\
 & A \to (B \vee C), \neg C \vdash A \to B \\
\text{P-}\!\to\!: & \overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad} \\
 & A \to (B \vee C), \neg C, A \vdash B \\
\text{MP:} & \overline{\qquad\qquad\qquad\qquad\qquad\qquad} \\
 & \neg C, A, B \vee C \vdash B \\
\text{A-}\vee\text{:} & \overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad} \\
 & \begin{array}{rcrc} & \ldots, B \vdash B & & \ldots, \neg C, C \vdash B \\ \text{GoalAssum:} & \overline{\qquad\qquad\qquad} & \text{ContrAssum:} & \overline{\qquad\qquad\qquad\qquad} \end{array}
\end{array}
\qquad \downarrow
$$

Connective Rules: Equivalence

- Prove equivalence as a goal:

$$\text{P-}\leftrightarrow\text{:}\ \frac{K\ldots \vdash F_1 \to F_2 \qquad K\ldots \vdash F_2 \to F_1}{K\ldots \vdash F_1 \leftrightarrow F_2}$$

  We prove $F_1 \leftrightarrow F_2$. First we prove $F_1 \to F_2$: ... Now we prove $F_2 \to F_1$: ...

- Use equivalence as an assumption (“substitution”):

$$\text{A-}\leftrightarrow\text{:}\ \frac{K\ldots[F_2/F_1], F_1 \leftrightarrow F_2 \vdash G}{K\ldots, F_1 \leftrightarrow F_2 \vdash G} \qquad \text{A-}\leftrightarrow\text{:}\ \frac{K\ldots, F_1 \leftrightarrow F_2 \vdash G[F_2/F_1]}{K\ldots, F_1 \leftrightarrow F_2 \vdash G}$$

$$\text{A-}\leftrightarrow\text{:}\ \frac{K\ldots[F_1/F_2], F_1 \leftrightarrow F_2 \vdash G}{K\ldots, F_1 \leftrightarrow F_2 \vdash G} \qquad \text{A-}\leftrightarrow\text{:}\ \frac{K\ldots, F_1 \leftrightarrow F_2 \vdash G[F_1/F_2]}{K\ldots, F_1 \leftrightarrow F_2 \vdash G}$$

  □ $\Gamma[F_2/F_1]$: replace in some formula(s) in Γ some occurrence of F_1 by F_2.

  We know $F_1 \leftrightarrow F_2$ and we know, e.g., $\neg F_2 \wedge F_3$. Therefore we know $\neg F_1 \wedge F_3$.

Equality Rules

- Prove an equality as a goal:

$$\text{P-=: } \frac{}{K\ldots\ \vdash t = t}$$

  We have to prove $t = t$ and are therefore done.

- Use an equality as assumption (“substitution”):

$$\text{A-=: } \frac{K\ldots[t_2/t_1], t_1 = t_2 \vdash G}{K\ldots, t_1 = t_2 \vdash G} \qquad \text{A-=: } \frac{K\ldots, t_1 = t_2 \vdash G[t_2/t_1]}{K\ldots, t_1 = t_2 \vdash G}$$

$$\text{A-=: } \frac{K\ldots[t_1/t_2], t_1 = t_2 \vdash G}{K\ldots, t_1 = t_2 \vdash G} \qquad \text{A-=: } \frac{K\ldots, t_1 = t_2 \vdash G[t_1/t_2]}{K\ldots, t_1 = t_2 \vdash G}$$

  - $\Gamma[t_2/t_1]$: replace in some formula(s) in Γ some occurrence of t_1 by t_2.

  We know $t_1 = t_2$ and we know (for example) $p(a, f(t_1))$. Therefore we know $p(a, f(t_2))$.

Example

$$\text{A-=:}\ \frac{\ldots, even(m), n = m^2 \vdash even(m^2)}{\ldots, even(m), n = m^2 \vdash even(n)}$$

We have to prove that n is even. Since we know $n = m^2$, it suffices to prove that m^2 is even.

Quantifier Rules: Universal Quantifier

- Prove universally quantified formula as a goal ("skolemization").

$$\text{P-}\forall\text{:}\ \frac{K \ldots \vdash F[\overline{x}/x]}{K \ldots \vdash (\forall x\colon F)} \qquad \text{where } \overline{x} \text{ does not occur in } K \ldots, F$$

  We prove $(\forall x\colon p(x, f(x)))$. We take arbitrary but fixed $\overline{x}$ and prove $p(\overline{x}, f(\overline{x}))$.

  - □ "fixed": "Skolem constant" $\overline{x}$ in contrast to variable x.
  - □ "arbitrary": $\overline{x}$ is a new constant about which nothing is known (it does not appear anywhere else in the proof situation).

Quantifier Rules: Universal Quantifier

- Use universally quantified formula as an assumption (“instantiation”):

$$\text{A-}\forall\text{:}\ \frac{K\ldots,(\forall x\colon F),F[t/x] \vdash G}{K\ldots,(\forall x\colon F) \vdash G}$$

where t is some term made up of symbols occuring in $K\ldots,F$

We know $(\forall x\colon p(x,f(x)))$. Thus we know (for $x := a$) $p(a,f(a))$ and (for $x := g(a)$) $p(g(a),f(g(a)))$.

  - □ $(\forall x\colon F)$ stays in the assumptions and can be instantiated again.
  - □ A “knowlege generating engine” that can be applied arbitrarily often.
  - □ The problem is to find suitable t that lets the proof make progress.
  - □ If an unsuitable t is chosen, the additional knowledge does not help.

Quantifier Rules: Existential Quantifier

- Prove an existentially quantified formula as a goal (“instantiation”):

  P-∃: $\dfrac{K\ldots \vdash F[t/x]}{K\ldots \vdash (\exists x\colon F)}$ where t is some term made up of symbols occuring in $K\ldots, F$

  We have to prove $(\exists x\colon p(x, f(x)))$. We prove (for $x := g(a)$) $p(g(a), f(g(a)))$.

  □ The problem is to find a “witness term” t that lets the proof succeed.

  □ If an unsuitable t is chosen, the proof fails.

Quantifier Rules: Existential Quantifier

- Use existentially quantified formula as an assumption ("skolemization"):

$$\text{A-}\exists\text{:}\ \frac{K\ldots, F[\bar{x}/x] \vdash G}{K\ldots, (\exists x\colon F) \vdash G} \quad \text{where } \bar{x} \text{ does not occur in } K\ldots, F, G$$

  We know $\exists x\colon p(x, f(x))$. Thus we know $p(\bar{x}, f(\bar{x}))$ for some $\bar{x}$.

  - □ $\bar{x}$ is an "arbitrary but fixed" Skolem constant.
  - □ $(\exists x\colon F)$ disappears from assumptions and cannot be skolemized again.
  - □ A "knowlege generating engine" that can be applied only once.

Example: A Quantifier Proof

$$\begin{array}{rcl}
 & \vdash (\exists x\colon \forall y\colon p(x,y)) \to (\forall y\colon \exists x\colon p(x,y)) & \\ \cline{2-2}
\text{P-}\to\text{:} & \exists x\colon \forall y\colon p(x,y) \vdash \forall y\colon \exists x\colon p(x,y) & \downarrow \\ \cline{2-2}
\text{P-}\forall\text{:} & \exists x\colon \forall y\colon p(x,y) \vdash \exists x\colon p(x,\bar{y}) & \\ \cline{2-2}
\text{A-}\exists\text{:} & \forall y\colon p(\bar{x},y) \vdash \exists x\colon p(x,\bar{y}) & \\ \cline{2-2}
\text{A-}\forall\text{:} & \forall y\colon p(\bar{x},y), p(\bar{x},\bar{y}) \vdash \exists x\colon p(x,\bar{y}) & \\ \cline{2-2}
\text{P-}\exists\text{:} & \forall y\colon p(\bar{x},y), p(\bar{x},\bar{y}) \vdash p(\bar{x},\bar{y}) & \\ \cline{2-2}
\text{GoalAssum:} & &
\end{array}$$

We prove

$$\exists x : \forall y : p(x,y)) \to (\forall y : \exists x : p(x,y)) \qquad \text{(a)}$$

We assume

$$\exists x : \forall y : p(x,y) \qquad \text{(1)}$$

and prove

$$\forall y : \exists x : p(x,y) \qquad \text{(b)}$$

We take arbitrary but fixed $\bar{y}$ and prove

$$\exists x : p(x,\bar{y}) \qquad \text{(c)}$$

From (1), we know $\forall y : p(\bar{x},y)$ for some $\bar{x}$, and from that, we know (for $y := \bar{y}$) (3) $p(\bar{x},\bar{y})$. In order to prove (c), let $x := \bar{x}$ and prove $p(\bar{x},\bar{y})$. QED (3).

Example: Another Quantifier Proof

$$\text{P-}\!\to\!:\ \dfrac{\vdash ((\exists x\colon p(x)) \land (\forall x\colon p(x) \to \exists y\colon q(x,y))) \to \exists x,y\colon q(x,y)}{\text{A-}\land:\ \dfrac{(\exists x\colon p(x)) \land (\forall x\colon p(x) \to \exists y\colon q(x,y)) \vdash \exists x,y\colon q(x,y)}{\text{A-}\exists:\ \dfrac{\exists x\colon p(x), \forall x\colon p(x) \to \exists y\colon q(x,y) \vdash \exists x,y\colon q(x,y)}{\text{A-}\forall\text{, Drop}:\ \dfrac{p(\bar{x}), \forall x\colon p(x) \to \exists y\colon q(x,y) \vdash \exists x,y\colon q(x,y)}{\text{MP,Drop}:\ \dfrac{p(\bar{x}), p(\bar{x}) \to \exists y\colon q(\bar{x},y) \vdash \exists x,y\colon q(x,y)}{\text{A-}\exists:\ \dfrac{\exists y\colon q(\bar{x},y) \vdash \exists x,y\colon q(x,y)}{\text{P-}\exists:\ \dfrac{q(\bar{x},\bar{y}) \vdash \exists x,y\colon q(x,y)}{\text{GoalAssum}:\ \dfrac{q(\bar{x},\bar{y}) \vdash q(\bar{x},\bar{y})}{}}}}}}}} \qquad \downarrow$$

In order to prove the goal we assume

$$(\exists x\colon p(x)) \land (\forall x\colon p(x) \to \exists y\colon q(x,y)) \qquad (1)$$

and show

$$\exists x,y\colon q(x,y) \qquad (b)$$

From (1), we know (2) $(\exists x\colon p(x))$ and (3) $(\forall x\colon p(x) \to \exists y\colon q(x,y))$. From (2), we know (4) $p(\bar{x})$ for some $\bar{x}$. From (3) we know (for $x := \bar{x}$) $(p(\bar{x}) \to \exists y\colon q(\bar{x},y))$ and together with (4), this gives $\exists y\colon q(\bar{x},y)$. From this, we know (7) $q(\bar{x},\bar{y})$ for some $\bar{y}$. In order to prove (b), let $x := \bar{x}$ and $y := \bar{y}$ and prove $q(\bar{x},\bar{y})$. QED (7).

Example: A Quantifier Proof with Branches

$$\text{P-}\!\to\text{: }\frac{\vdash \big((p(a)\lor q(b))\land(\forall x\colon p(x)\to r(x))\land(\forall x\colon q(x)\to r(f(x)))\big)\to\exists x\colon r(x)}{\text{A-}\land\text{: }\frac{(p(a)\lor q(b))\land(\forall x\colon p(x)\to r(x))\land(\forall x\colon q(x)\to r(f(x)))\vdash\exists x\colon r(x)}{\text{A-}\lor\text{, Drop: }\frac{p(a)\lor q(b),(\forall x\colon p(x)\to r(x)),(\forall x\colon q(x)\to r(f(x)))\vdash\exists x\colon r(x)}{\text{A-}\forall\text{, Drop: }\frac{p(a),(\forall x\colon p(x)\to r(x))\vdash\exists x\colon r(x)}{\text{MP: }\frac{p(a),p(a)\to r(a)\vdash\exists x\colon r(x)}{\text{P-}\exists\text{: }\frac{p(a),r(a)\vdash\exists x\colon r(x)}{\text{GoalAssum: }\frac{p(a),r(a)\vdash r(a)}{}}}}\qquad\text{A-}\forall\text{, Drop: }\frac{q(b),(\forall x\colon q(x)\to r(f(x)))\vdash\exists x\colon r(x)}{\text{MP: }\frac{q(b),q(b)\to r(f(b))\vdash\exists x\colon r(x)}{\text{P-}\exists\text{: }\frac{q(b),r(f(b))\vdash\exists x\colon r(x)}{\text{GoalAssum: }\frac{q(b),r(f(b))\vdash r(f(b))}{}}}}}}}\quad\downarrow$$

In order to prove the goal we assume

$$p(a)\lor q(b) \tag{1}$$

$$(\forall x\colon p(x)\to r(x)) \tag{2}$$

$$(\forall x\colon q(x)\to r(f(x))) \tag{3}$$

and prove

$$\exists x\colon r(x) \tag{b}$$

From (1), we have two cases:

1. Assume $p(a)$: From (2), we know $p(a)\to r(a)$ (for $x:=a$), and together with the case assumption, this gives $r(a)$ and therefore (b) for $x:=a$.
2. Assume $q(b)$: From (3), we know $q(b)\to r(f(b))$ (for $x:=b$), and together with the case assumption, this gives $r(f(b))$ and therefore (b) for $x:=f(b)$.

Proving Strategies

- Proving: partially “art” but mostly “craft”.
  - Most of a proof is guided by the structure of proof situations.
  - Only in a few places really “creativity” or “ingenuity” is required.
- First: apply only the “goal-oriented” rules.
  - Decompose the complex goal into one or more simpler goals.
  - Stop when goals become atomic or existentially quantified.
- Then: apply the “assumption-oriented” rules.
  - Decompose complex assumptions into simpler ones.
  - Skolemize existentially quantified assumptions.
  - Instantiate universally quantified assumptions.
- Ultimately: “close the gap” between assumptions and goal.
  - Derive atomic goal as an assumption.
  - Instantiate existentially quantified goal s.t. body of the formula is an assumption.

By considering proving as a “syntactic” process, already a major part of a proof can be elaborated (possibly even completed).

Mathematical Proofs

Mathematical proofs are typically written in a much more informal style.

- Do not mention all steps.
- Combine several steps into one.
- Reuse name of variable for name of Skolem constant.
- Use hidden assumptions. . . .

A mathematical proof is an easily readable "sketch" that just gives the essential information to reconstruct a corresponding formal proof.

Mathematical Proofs: An Example

Theorem: Suppose a divides b if and only if, for some $t \in \mathbb{N}$, $b = t \cdot a$. Then, if a divides b it also divides every multiple of b.

Proof: Assume $a, b, s \in \mathbb{N}$ arbitrary but fixed such that a divides b. We show that a divides $s \cdot b$, i.e. $\exists t \in \mathbb{N}$: $s \cdot b = t \cdot a$. Since a divides b, we know $b = \bar{t} \cdot a$ for some $\bar{t} \in \mathbb{N}$, thus, we have to find $t \in \mathbb{N}$ with $s \cdot \bar{t} \cdot a = t \cdot a$. Let now $t := s \cdot \bar{t} \in \mathbb{N}$, we have to show $s \cdot \bar{t} \cdot a = s \cdot \bar{t} \cdot a$. QED.

Every sentence in the proof is justified by one or more proof rules. Trivial steps (e.g. split conjunction in assumptions) are not mentioned explicitly.