Auto2 prover

Bohua Zhan

March 17, 2025

Abstract

Auto2 is a saturation-based heuristic prover for higher-order logic, implemented as a tactic in Isabelle.

This entry contains the instantiation of auto2 for Isabelle/HOL, along with two basic examples: solutions to some of the Pelletier's problems, and elementary number theory of primes.

Contents

1 Introduction

Auto2 [2] is a proof automation tool implemented in Isabelle. It uses a saturation-based approach to proof search: starting with a list of initial assumptions, it iteratively adds facts that can be derived from these assumptions, with the aim of ultimately deriving a contradiction. Users can add their own proof procedures to auto2 in the form of *proof steps*, in order to implement domain-specific knowledge. Auto2 can be instantiated to both Isabelle/HOL (for ordinary usage) and Isabelle/FOL (for formalization of mathematics based on set theory).

This AFP entry contains the instantiation of auto2 to Isabelle/HOL, and two basic applications:

- Pelletier's problems: solutions to some of the problems in Pelletier's collection of problems for testing automatic theorem provers [1]. Auto2 is not intended to compete with ATPs. In our examples, we merely show how to use the prover to solve some of the problems, sometimes with hints.
- Elementary number theory: theory of prime numbers up to the infinitude of primes and unique factorization. This example follows the development in HOL/Computational_Algebra/Primes.thy in the Isabelle distribution.

2 Pelletier's problems

theory *Pelletier*
imports *Logic-Thms*
begin

theorem *p1*: $(p \longrightarrow q) \longleftrightarrow (\neg q \longrightarrow \neg p)$ **by** *auto2*

theorem *p2*: $(\neg\neg p) \longleftrightarrow p$ **by** *auto2*

theorem *p3*: $\neg(p \longrightarrow q) \Longrightarrow q \longrightarrow p$ **by** *auto2*

theorem *p4*: $(\neg p \longrightarrow q) \longleftrightarrow (\neg q \longrightarrow p)$ **by** *auto2*

theorem *p5*: $(p \lor q) \longrightarrow (p \lor r) \Longrightarrow p \lor (q \longrightarrow r)$ **by** *auto2*

theorem *p6*: $p \lor \neg p$ **by** *auto2*

theorem *p7*: $p \lor \neg\neg\neg p$ **by** *auto2*

theorem *p8*: $((p \longrightarrow q) \longrightarrow p) \Longrightarrow p$ **by** *auto2*

theorem *p9*: $(p \vee q) \wedge (\neg p \vee q) \wedge (p \vee \neg q) \Longrightarrow \neg(\neg p \vee \neg q)$ **by** *auto2*

theorem *p10*: $q \longrightarrow r \Longrightarrow r \longrightarrow p \wedge q \Longrightarrow p \longrightarrow q \vee r \Longrightarrow p \longleftrightarrow q$ **by** *auto2*

theorem *p11*: $p \longleftrightarrow p$ **by** *auto2*

theorem *p12*: $((p \longleftrightarrow q) \longleftrightarrow r) \longleftrightarrow (p \longleftrightarrow (q \longleftrightarrow r))$
@**proof**
  @**case** p
  @**case** q
@**qed**

theorem *p13*: $p \vee (q \wedge r) \longleftrightarrow (p \vee q) \wedge (p \vee r)$ **by** *auto2*

theorem *p14*: $(p \longleftrightarrow q) \longleftrightarrow ((q \vee \neg p) \wedge (\neg q \vee p))$ **by** *auto2*

theorem *p15*: $(p \longrightarrow q) \longleftrightarrow (\neg p \vee q)$ **by** *auto2*

theorem *p16*: $(p \longrightarrow q) \vee (q \longrightarrow p)$ **by** *auto2*

theorem *p17*: $(p \wedge (q \longrightarrow r) \longrightarrow s) \longleftrightarrow (\neg p \vee q \vee s) \wedge (\neg p \vee \neg r \vee s)$ **by** *auto2*

theorem *p18*: $\exists\, y{::}'a.\ \forall\, x.\ F(y) \longrightarrow F(x)$
@**proof**
  @**case** $\forall\, x.\ F(x)$ @*with*
    @**obtain** $y{::}'a$ **where** $y = y$ @**have** $\forall\, x.\ F(y) \longrightarrow F(x)$
  @**end**
  @**obtain** y **where** $\neg F(y)$ @**have** $\forall\, x.\ F(y) \longrightarrow F(x)$
@**qed**

theorem *p19*: $\exists\, x{::}'a.\ \forall\, y\ z.\ (P(y) \longrightarrow Q(z)) \longrightarrow (P(x) \longrightarrow Q(x))$
@**proof**
  @**case** $\exists\, x.\ P(x) \longrightarrow Q(x)$ @*with*
    @**obtain** x **where** $P(x) \longrightarrow Q(x)$
    @**have** $\forall\, y\ z.\ (P(y) \longrightarrow Q(z)) \longrightarrow (P(x) \longrightarrow Q(x))$
  @**end**
  @**obtain** $x{::}'a$ **where** $x = x$
  @**have** $\forall\, y\ z.\ (P(y) \longrightarrow Q(z)) \longrightarrow (P(x) \longrightarrow Q(x))$
@**qed**

theorem *p20*: $\forall\, x\ y.\ \exists\, z.\ \forall\, w.\ P(x) \wedge Q(y) \longrightarrow R(z) \wedge S(w) \Longrightarrow$
  $\exists\, x\ y.\ P(x) \wedge Q(y) \Longrightarrow \exists\, z.\ R(z)$
@**proof**
  @**obtain** $x\ y$ **where** $P(x)\ Q(y)$
  @**obtain** z **where** $\forall\, w.\ P(x) \wedge Q(y) \longrightarrow R(z) \wedge S(w)$
@**qed**

theorem *p21*: $\exists\, x.\ p \longrightarrow F(x) \Longrightarrow \exists\, x.\ F(x) \longrightarrow p \Longrightarrow \exists\, x.\ p \longleftrightarrow F(x)$
@**proof**

@**case** p @*with* @**obtain** x **where** $F(x)$ @**have** $p \longleftrightarrow F(x)$ @**end**
@**case** $\neg p$ @*with* @**obtain** x **where** $\neg F(x)$ @**have** $p \longleftrightarrow F(x)$ @**end**
@**qed**

theorem *p22*: $\forall\, x{::}'a.\ p \longleftrightarrow F(x) \Longrightarrow p \longleftrightarrow (\forall\, x.\ F(x))$
@**proof**
@**case** p @**obtain** $x{::}'a$ **where** $x = x$
@**qed**

theorem *p23*: $(\forall\, x{::}'a.\ p \vee F(x)) \longleftrightarrow (p \vee (\forall\, x.\ F(x)))$ **by** *auto2*

theorem *p29*: $\exists\, x.\ F(x) \Longrightarrow \exists\, x.\ G(x) \Longrightarrow$
$((\forall\, x.\ F(x) \longrightarrow H(x)) \wedge (\forall\, x.\ G(x) \longrightarrow J(x))) \longleftrightarrow$
$(\forall\, x\ y.\ F(x) \wedge G(y) \longrightarrow H(x) \wedge J(y))$
@**proof**
@**obtain** $a\ b$ **where** $F(a)\ \ G(b)$
@**case** $\forall\, x\ y.\ F(x) \wedge G(y) \longrightarrow H(x) \wedge J(y)$ @*with*
@**have** $\forall\, x.\ F(x) \longrightarrow H(x)$ @*with* @**have** $F(x) \wedge G(b)$ @**end**
@**have** $\forall\, y.\ G(y) \longrightarrow J(y)$ @*with* @**have** $F(a) \wedge G(y)$ @**end**
@**end**
@**qed**

theorem *p30*: $\forall\, x.\ F(x) \vee G(x) \longrightarrow \neg H(x) \Longrightarrow$
$\forall\, x.\ (G(x) \longrightarrow \neg I(x)) \longrightarrow F(x) \wedge H(x) \Longrightarrow \forall\, x.\ I(x)$
@**proof**
@**have** $\forall\, x.\ I(x)$ @*with* @**case** $F(x)$ @**end**
@**qed**

theorem *p31*: $\neg(\exists\, x.\ F(x) \wedge (G(x) \vee H(x))) \Longrightarrow \exists\, x.\ I(x) \wedge F(x) \Longrightarrow \forall\, x.\ \neg H(x)$
$\longrightarrow J(x) \Longrightarrow$
$\exists\, x.\ I(x) \wedge J(x)$ **by** *auto2*

theorem *p32*: $\forall\, x.\ (F(x) \wedge (G(x) \vee H(x))) \longrightarrow I(x) \Longrightarrow \forall\, x.\ I(x) \wedge H(x) \longrightarrow$
$J(x) \Longrightarrow$
$\forall\, x.\ K(x) \longrightarrow H(x) \Longrightarrow \forall\, x.\ F(x) \wedge K(x) \longrightarrow J(x)$ **by** *auto2*

theorem *p33*: $(\forall\, x.\ p(a) \wedge (p(x) \longrightarrow p(b)) \longrightarrow p(c)) \longleftrightarrow$
$(\forall\, x.\ (\neg p(a) \vee p(x) \vee p(c)) \wedge (\neg p(a) \vee \neg p(b) \vee p(c)))$ **by** *auto2*

theorem *p35*: $\exists\, (x{::}'a)\ (y{::}'b).\ P(x,y) \longrightarrow (\forall\, x\ y.\ P(x,y))$ **by** *auto2*

theorem *p39*: $\neg(\exists\, x.\ \forall\, y.\ F(y,x) \longleftrightarrow \neg F(y,y))$
@**proof**
@**contradiction**
@**obtain** x **where** $\forall\, y.\ F(y,x) \longleftrightarrow \neg F(y,y)$
@**case** $F(x,x)$
@**qed**

theorem *p40*: $\exists\, y.\ \forall\, x.\ F(x,y) \longleftrightarrow F(x,x) \Longrightarrow \neg(\forall\, x.\ \exists\, y.\ \forall\, z.\ F(z,y) \longleftrightarrow \neg F(z,x))$
@**proof**
  @**obtain** A **where** $\forall\, x.\ F(x,A) \longleftrightarrow F(x,x)$
  @**have** $\neg(\exists\, y.\ \forall\, z.\ F(z,y) \longleftrightarrow \neg F(z,A))$ @*with*
    @**have** (@*rule*) $\forall\, y.\ \neg(\forall\, z.\ F(z,y) \longleftrightarrow \neg F(z,A))$ @*with*
      @**have** $\neg(F(y,y) \longleftrightarrow \neg F(y,A))$ @*with* @**case** $F(y,y)$ @**end**
    @**end**
  @**end**
@**qed**

theorem *p42*: $\neg(\exists\, y.\ \forall\, x.\ F(x,y) \longleftrightarrow \neg(\exists\, z.\ F(x,z) \wedge F(z,x)))$
@**proof**
  @**contradiction**
  @**obtain** y **where** $\forall\, x.\ F(x,y) \longleftrightarrow \neg(\exists\, z.\ F(x,z) \wedge F(z,x))$
  @**case** $F(y,y)$
@**qed**

theorem *p43*: $\forall\, x\ y.\ Q(x,y) \longleftrightarrow (\forall\, z.\ F(z,x) \longleftrightarrow F(z,y)) \Longrightarrow$
  $\forall\, x\ y.\ Q(x,y) \longleftrightarrow Q(y,x)$ **by** *auto2*

theorem *p47*:
  $(\forall\, x.\ P1(x) \longrightarrow P0(x)) \wedge (\exists\, x.\ P1(x)) \Longrightarrow$
   $(\forall\, x.\ P2(x) \longrightarrow P0(x)) \wedge (\exists\, x.\ P2(x)) \Longrightarrow$
   $(\forall\, x.\ P3(x) \longrightarrow P0(x)) \wedge (\exists\, x.\ P3(x)) \Longrightarrow$
   $(\forall\, x.\ P4(x) \longrightarrow P0(x)) \wedge (\exists\, x.\ P4(x)) \Longrightarrow$
   $(\forall\, x.\ P5(x) \longrightarrow P0(x)) \wedge (\exists\, x.\ P5(x)) \Longrightarrow$
   $(\exists\, x.\ Q1(x)) \wedge (\forall\, x.\ Q1(x) \longrightarrow Q0(x)) \Longrightarrow$
   $\forall\, x.\ P0(x) \longrightarrow ((\forall\, y.\ Q0(y) \longrightarrow R(x,y)) \vee$
                 $(\forall\, y.\ P0(y) \wedge S(y,x) \wedge (\exists\, z.\ Q0(z) \wedge R(y,z)) \longrightarrow R(x,y))) \Longrightarrow$
   $\forall\, x\ y.\ P3(y) \wedge (P5(x) \vee P4(x)) \longrightarrow S(x,y) \Longrightarrow$
   $\forall\, x\ y.\ P3(x) \wedge P2(y) \longrightarrow S(x,y) \Longrightarrow$
   $\forall\, x\ y.\ P2(x) \wedge P1(y) \longrightarrow S(x,y) \Longrightarrow$
   $\forall\, x\ y.\ P1(x) \wedge (P2(y) \vee Q1(y)) \longrightarrow \neg R(x,y) \Longrightarrow$
   $\forall\, x\ y.\ P3(x) \wedge P4(y) \longrightarrow R(x,y) \Longrightarrow$
   $\forall\, x\ y.\ P3(x) \wedge P5(y) \longrightarrow \neg R(x,y) \Longrightarrow$
   $\forall\, x.\ P4(x) \vee P5(x) \longrightarrow (\exists\, y.\ Q0(y) \wedge R(x,y)) \Longrightarrow$
   $\exists\, x\ y.\ P0(x) \wedge P0(y) \wedge (\exists\, z.\ Q1(z) \wedge R(y,z) \wedge R(x,y))$
@**proof**
  @**obtain** $x1\ x2\ x3\ x4\ x5$ **where** $P1(x1)\ P2(x2)\ P3(x3)\ P4(x4)\ P5(x5)$
  @**have** $S(x3,x2)$ @**have** $S(x2,x1)$ @**have** $R(x3,x4)$ @**have** $\neg R(x3,x5)$
@**qed**

theorem *p48*: $a = b \vee c = d \Longrightarrow a = c \vee b = d \Longrightarrow a = d \vee b = c$ **by** *auto2*

theorem *p49*: $\exists\, x\ y.\ \forall\, (z::'a).\ z = x \vee z = y \Longrightarrow P(a) \wedge P(b) \Longrightarrow (a::'a) \neq b \Longrightarrow$
$\forall\, x.\ P(x)$
@**proof**
  @**obtain** $x\ y$ **where** $\forall\, (z::'a).\ z = x \vee z = y$
  @**have** $x = a \vee x = b$

@**have** $\forall\, c.\ P(c)$ @*with* @**have** $c = a \vee c = b$ @**end**
@**qed**

theorem *p50*: $\forall\, x.\ F(a,x) \vee (\forall\, y.\ F(x,y)) \Longrightarrow \exists\, x.\ \forall\, y.\ F(x,y)$
@**proof**
@**case** $\forall\, y.\ F(a,y)$
@**obtain** y **where** $\neg F(a,y)$
@**have** (@*rule*) $\forall\, z.\ F(a,y) \vee F(y,z)$
@**qed**

theorem *p51*: $\exists\, z\ w.\ \forall\, x\ y.\ F(x,y) \longleftrightarrow x = z \wedge y = w \Longrightarrow$
$\exists\, z.\ \forall\, x.\ (\exists\, w.\ \forall\, y.\ F(x,y) \longleftrightarrow y = w) \longleftrightarrow x = z$
@**proof**
@**obtain** $z\ w$ **where** $\forall\, x\ y.\ F(x,y) \longleftrightarrow x = z \wedge y = w$
@**have** $\forall\, x.\ (\exists\, w.\ \forall\, y.\ F(x,y) \longleftrightarrow y = w) \longleftrightarrow x = z$ @*with*
@**case** $x = z$ @*with* @**have** $\forall\, y.\ F(x,y) \longleftrightarrow y = w$ @**end**
@**end**
@**qed**

theorem *p52*: $\exists\, z\ w.\ \forall\, x\ y.\ F(x,y) \longleftrightarrow x = z \wedge y = w \Longrightarrow$
$\exists\, w.\ \forall\, y.\ (\exists\, z.\ \forall\, x.\ F(x,y) \longleftrightarrow x = z) \longleftrightarrow y = w$
@**proof**
@**obtain** $z\ w$ **where** $\forall\, x\ y.\ F(x,y) \longleftrightarrow x = z \wedge y = w$
@**have**$\forall\, y.\ (\exists\, z.\ \forall\, x.\ F(x,y) \longleftrightarrow x = z) \longleftrightarrow y = w$ @*with*
@**case** $y = w$ @*with* @**have** $\forall\, x.\ F(x,y) \longleftrightarrow x = z$ @**end**
@**end**
@**qed**

theorem *p55*:
$\exists\, x.\ L(x) \wedge K(x,a) \Longrightarrow$
$L(a) \wedge L(b) \wedge L(c) \Longrightarrow$
$\forall\, x.\ L(x) \longrightarrow x = a \vee x = b \vee x = c \Longrightarrow$
$\forall\, y\ x.\ K(x,y) \longrightarrow H(x,y) \Longrightarrow$
$\forall\, x\ y.\ K(x,y) \longrightarrow \neg R(x,y) \Longrightarrow$
$\forall\, x.\ H(a,x) \longrightarrow \neg H(c,x) \Longrightarrow$
$\forall\, x.\ x \neq b \longrightarrow H(a,x) \Longrightarrow$
$\forall\, x.\ \neg R(x,a) \longrightarrow H(b,x) \Longrightarrow$
$\forall\, x.\ H(a,x) \longrightarrow H(b,x) \Longrightarrow$ — typo in text
$\forall\, x.\ \exists\, y.\ \neg H(x,y) \Longrightarrow$
$a \neq b \Longrightarrow$
$K(a,a)$
@**proof**
@**case** $K(b,a)$ @*with* @**have** $\forall\, x.\ H(b,x)$ @**end**
@**qed**

theorem *p56*: $(\forall\, x.\ (\exists\, y.\ F(y) \wedge x = f(y)) \longrightarrow F(x)) \longleftrightarrow (\forall\, x.\ F(x) \longrightarrow F(f(x)))$
by *auto2*

theorem *p57*: $F(f(a,b),f(b,c)) \Longrightarrow F(f(b,c),f(a,c)) \Longrightarrow$

$\forall x\ y\ z.\ F(x,y) \wedge F(y,z) \longrightarrow F(x,z) \Longrightarrow F(f(a,b),f(a,c))$ **by** *auto2*

theorem *p58*: $\forall x\ y.\ f(x) = g(y) \Longrightarrow \forall x\ y.\ f(f(x)) = f(g(y))$
@**proof**
  @**have** $\forall x\ y.\ f(f(x)) = f(g(y))$ @*with*
    @**have** $f(x) = g(y)$
  @**end**
@**qed**

theorem *p59*: $\forall x{::}'a.\ F(x) \longleftrightarrow \neg F(f(x)) \Longrightarrow \exists x.\ F(x) \wedge \neg F(f(x))$
@**proof**
  @**obtain** $x{::}'a$ **where** $x = x$ @**case** $F(x)$
@**qed**

theorem *p60*: $\forall x.\ F(x,f(x)) \longleftrightarrow (\exists y.\ (\forall z.\ F(z,y) \longrightarrow F(z,f(x))) \wedge F(x,y))$ **by** *auto2*

theorem *p61*: $\forall x\ y\ z.\ f(x,f(y,z)) = f(f(x,y),z) \Longrightarrow \forall x\ y\ z\ w.\ f(x,f(y,f(z,w))) = f(f(f(x,y),z),w)$
  by *auto2*

end

3 Primes

theory *Primes-Ex*
  imports *Auto2-Main*
begin

3.1 Basic definition

definition *prime* :: *nat* ⇒ *bool* **where** [*rewrite*]:
  $prime\ p = (1 < p \wedge (\forall m.\ m\ dvd\ p \longrightarrow m = 1 \vee m = p))$

lemma *primeD1* [*forward*]: $prime\ p \Longrightarrow 1 < p$ **by** *auto2*
lemma *primeD2*: $prime\ p \Longrightarrow m\ dvd\ p \Longrightarrow m = 1 \vee m = p$ **by** *auto2*
setup ‹*add-forward-prfstep-cond* @{*thm primeD2*} [*with-cond* $?m \neq 1$, *with-cond* $?m \neq ?p$]›
setup ‹*del-prfstep-thm-eqforward* @{*thm prime-def*}›

theorem *exists-prime* [*resolve*]: $\exists p.\ prime\ p$
@**proof** @**have** *prime 2* @**qed**

lemma *prime-odd-nat*: $prime\ p \Longrightarrow p > 2 \Longrightarrow odd\ p$ **by** *auto2*

lemma *prime-imp-coprime-nat* [*backward2*]: $prime\ p \Longrightarrow \neg\ p\ dvd\ n \Longrightarrow coprime\ p\ n$ **by** *auto2*

lemma *prime-dvd-mult-nat*: $prime\ p \Longrightarrow p\ dvd\ m * n \Longrightarrow p\ dvd\ m \vee p\ dvd\ n$ **by** *auto2*
setup ‹*add-forward-prfstep-cond* @{*thm prime-dvd-mult-nat*}
 (*with-conds* [$?m \neq ?p$, $?n \neq ?p$, $?m \neq ?p * ?m'$, $?n \neq ?p * ?n'$])›

theorem *prime-dvd-intro*: $prime\ p \Longrightarrow p * q = m * n \Longrightarrow p\ dvd\ m \vee p\ dvd\ n$
@**proof** @**have** $p\ dvd\ m * n$ @**qed**
setup ‹*add-forward-prfstep-cond* @{*thm prime-dvd-intro*}
 (*with-conds* [$?m \neq ?p$, $?n \neq ?p$, $?m \neq ?p * ?m'$, $?n \neq ?p * ?n'$])›

lemma *prime-dvd-mult-eq-nat*: $prime\ p \Longrightarrow p\ dvd\ m * n = (p\ dvd\ m \vee p\ dvd\ n)$
by *auto2*

lemma *not-prime-eq-prod-nat* [*backward1*]: $n > 1 \Longrightarrow \neg\ prime\ n \Longrightarrow$
  $\exists\, m\ k.\ n = m * k \wedge 1 < m \wedge m < n \wedge 1 < k \wedge k < n$
@**proof**
 @**obtain** m **where** $m\ dvd\ n \wedge m \neq 1 \wedge m \neq n$
 @**obtain** k **where** $n = m * k$ @**have** $m \leq m * k$ @**have** $k \leq m * k$
@**qed**

lemma *prime-dvd-power-nat*: $prime\ p \Longrightarrow p\ dvd\ x\widehat{\ }n \Longrightarrow p\ dvd\ x$ **by** *auto2*
setup ‹*add-forward-prfstep-cond* @{*thm prime-dvd-power-nat*} [*with-cond* $?p \neq ?x$]›

lemma *prime-dvd-power-nat-iff*: $prime\ p \Longrightarrow n > 0 \Longrightarrow p\ dvd\ x\widehat{\ }n \longleftrightarrow p\ dvd\ x$
by *auto2*

lemma *prime-nat-code*: $prime\ p = (1 < p \wedge (\forall\, x.\ 1 < x \wedge x < p \longrightarrow \neg\ x\ dvd\ p))$
by *auto2*

lemma *prime-factor-nat* [*backward*]: $n \neq 1 \Longrightarrow \exists\, p.\ p\ dvd\ n \wedge prime\ p$
@**proof**
 @**strong-induct** n
 @**case** $prime\ n$ @**case** $n = 0$
 @**obtain** k **where** $k \neq 1\ k \neq n\ k\ dvd\ n$
 @**apply-induct-hyp** k
@**qed**

lemma *prime-divprod-pow-nat*:
 $prime\ p \Longrightarrow coprime\ a\ b \Longrightarrow p\widehat{\ }n\ dvd\ a * b \Longrightarrow p\widehat{\ }n\ dvd\ a \vee p\widehat{\ }n\ dvd\ b$ **by** *auto2*

lemma *prime-product* [*forward*]: $prime\ (p * q) \Longrightarrow p = 1 \vee q = 1$
@**proof** @**have** $p\ dvd\ q * p$ @**qed**

lemma *prime-exp*: $prime\ (p\ \widehat{\ }\ n) \longleftrightarrow n = 1 \wedge prime\ p$ **by** *auto2*

lemma *prime-power-mult*: $prime\ p \Longrightarrow x * y = p\ \widehat{\ }\ k \Longrightarrow \exists\, i\ j.\ x = p\ \widehat{\ }\ i \wedge y = p\ \widehat{\ }\ j$
@**proof**

@**induct** k **arbitrary** $x\ y$ @*with*
@**subgoal** $k = Suc\ k'$
@**case** $p\ dvd\ x$ @*with*
@**obtain** x' **where** $x = p * x'$ @**have** $x * y = p * (x' * y)$
@**obtain** $i\ j$ **where** $x' = p\ \hat{}\ i\ y = p\ \hat{}\ j$ @**have** $x = p\ \hat{}\ Suc\ i$ @**end**
@**case** $p\ dvd\ y$ @*with*
@**obtain** y' **where** $y = p * y'$ @**have** $x * y = p * (x * y')$
@**obtain** $i\ j$ **where** $x = p\ \hat{}\ i\ y' = p\ \hat{}\ j$ @**have** $y = p\ \hat{}\ Suc\ j$ @**end**
@**endgoal**
@**end**
@**qed**

3.2 Infinitude of primes

theorem *bigger-prime* [*resolve*]: $\exists p.\ prime\ p \wedge n < p$
@**proof**
@**obtain** p **where** $prime\ p\ p\ dvd\ fact\ n + 1$
@**case** $n \geq p$ @*with* @**have** $(p::nat)\ dvd\ fact\ n$ @**end**
@**qed**

theorem *primes-infinite*: $\neg\ finite\ \{p.\ prime\ p\}$
@**proof**
@**obtain** b **where** $prime\ b\ Max\ \{p.\ prime\ p\} < b$
@**qed**

3.3 Existence and uniqueness of prime factorization

theorem *factorization-exists*: $n > 0 \Longrightarrow \exists M.\ (\forall p{\in}\#M.\ prime\ p) \wedge n = (\prod i{\in}\#M.\ i)$
@**proof**
@**strong-induct** n
@**case** $n = 1$ @*with* @**have** $n = (\prod i{\in}\#\ \{\#\}.\ i)$ @**end**
@**case** $prime\ n$ @*with* @**have** $n = (\prod i{\in}\#\ \{\#n\#\}.\ i)$ @**end**
@**obtain** $m\ k$ **where** $n = m * k\ 1 < m\ m < n\ 1 < k\ k < n$
@**apply-induct-hyp** m
@**obtain** M **where** $(\forall p{\in}\#M.\ prime\ p)\ m = (\prod i{\in}\#M.\ i)$
@**apply-induct-hyp** k
@**obtain** K **where** $(\forall p{\in}\#K.\ prime\ p)\ k = (\prod i{\in}\#K.\ i)$
@**have** $n = (\prod i{\in}\#(M{+}K).\ i)$
@**qed**

theorem *prime-dvd-multiset* [*backward1*]: $prime\ p \Longrightarrow p\ dvd\ (\prod i{\in}\#M.\ i) \Longrightarrow \exists n.\ n{\in}\#M \wedge p\ dvd\ n$
@**proof**
@**strong-induct** M
@**case** $M = \{\#\}$
@**obtain** $M'\ m$ **where** $M = M' + \{\#m\#\}$
@**contradiction** @**apply-induct-hyp** M'
@**qed**

theorem *factorization-unique-aux*:
$\forall\, p{\in}\#M.\ prime\ p \Longrightarrow \forall\, p{\in}\#N.\ prime\ p \Longrightarrow (\prod i{\in}\#M.\ i)\ dvd\ (\prod i{\in}\#N.\ i) \Longrightarrow$
$M \subseteq\#\ N$
@**proof**
  @**strong-induct** M **arbitrary** N
  @**case** $M = \{\#\}$
  @**obtain** $M'\ m$ **where** $M = M' + \{\#m\#\}$
  @**have** $m\ dvd\ (\prod i{\in}\#M.\ i)$
  @**obtain** n **where** $n \in\#\ N\ m\ dvd\ n$
  @**obtain** N' **where** $N = N' + \{\#n\#\}$
  @**have** $m = n$
  @**have** $(\prod i{\in}\#M'.\ i)\ dvd\ (\prod i{\in}\#N'.\ i)$
  @**apply-induct-hyp** $M'\ N'$
@**qed**
setup ‹*add-forward-prfstep-cond* @{*thm factorization-unique-aux*} [*with-cond ?M*
$\neq$ *?N*]›

theorem *factorization-unique*:
$\forall\, p{\in}\#M.\ prime\ p \Longrightarrow \forall\, p{\in}\#N.\ prime\ p \Longrightarrow (\prod i{\in}\#M.\ i) = (\prod i{\in}\#N.\ i) \Longrightarrow$
$M = N$
@**proof** @**have** $M \subseteq\#\ N$ @**qed**
setup ‹*del-prfstep-thm* @{*thm factorization-unique-aux*}›

end

References

[1] F. J. Pelletier. Seventy-five problems for testing automatic theorem provers. *Journal of Automated Reasoning*, 2:191–216, 1986.

[2] B. Zhan. Auto2: a saturation-based heuristic prover for higher-order logic. In J. C. Blanchette and S. Merz, editors, *ITP 2016*, pages 441–456, 2016.