Induction

J.-F. Monin

Univ. Joseph Fourier and
LIAMA-FORMES, Tsinghua Univ., Beijing

Asia-Pacific Coq Summer School, Tsinghua, August 2010

Why induction matters

Tool of choice for proving properties on an infinite (but countable) number of values

Other methods are

- either weaker (prove less properties)
- or rely on induction in a hidden way

Required in many applications in computer science

- reasoning on data structures
- language syntax
- programming language semantics
- proofs of algorithms

Strength of induction

Induction require ingenuity, in general

- a consequence of Gödel incompleteness theorems
- support for induction is a discriminating criterium for automated provers
  Coq supports induction
- proof search $\neq$ proof checking

Several forms of induction

- Basic induction on natural numbers ($\mathbb{N}$)
- Well-founded induction on $(\mathbb{N}, <)$
- Well-founded induction on (S, R), where S is an arbitrary set and R a suitable relation on S
- Transfinite induction
- Structural induction

We will focus on structural induction, because it is

- a very natural extension of basic induction but on lists, trees, terms ... instead of $\mathbb{N}$
- close to computer science concerns
- yet powerful enough to embed all other kinds of induction

Proving something on all natural numbers

Let us define $x \leq y \ \stackrel{\text{def}}{=\!=} \ \exists d, d + x = y$

Prove $\forall x, 2 + x \leq 5 + x$

- Take an arbitrary natural number x
- Remark that $3 + (2 + x) = 5 + x$
- Hence $\exists d, d + (2 + x) = 5 + x$
- By definition of $\leq$ we get: $2 + x \leq 5 + x$

This proof is uniform : it does not depend on the value of x

Looking at x : proof by cases

Prove $\forall x, x \leq 4 \Rightarrow \exists y, x = 2y \vee \; x = 1 + 2y$

The proof is not uniform: different is each case

- Case $x = 0$: take $y = 0$, left, check $0 = 2.0$
- Case $x = 1$: take $y = 0$, right, check $1 = 1 + 2.0$
- Case $x = 2$: take $y = 1$, left, check $2 = 2.1$
- Case $x = 3$: take $y = 1$, right, check $3 = 1 + 2.1$
- Case $x = 4$: take $y = 2$, left, check $4 = 2.2$
- Case $x = 5 + n$: don't care

Common scheme for a proof by cases on $\mathbb{N}$

$$\frac{P(0) \qquad \forall n, 1 \leq n \Rightarrow P(n)}{\forall x, P(x)}$$

More elegant:

$$\frac{P(0) \qquad \forall n, P(1+n)}{\forall x, P(x)}$$

What do you think of the following one?

$x \leq y \stackrel{\text{def}}{=\!=} \exists d, d + x = y$

Prove $\forall x, x \leq 3x$

- Take an arbitrary natural number x
- Remark that $2x + x = 3x$
- Hence $\exists d, d + x = 3x$
- That is $x \leq 3x$

Is this proof uniform? Yes: no case analysis on x

Proof by cases on a finite set

Material:

- a *finite* set $A = \{a_1, \ldots a_n\}$
- a predicate (property) P on A

In order to prove $\forall x, P(x)$,
prove P on each element a_i

$\Rightarrow$ n cases to consider
We can make a completely different proof in each case

Formally

$$\frac{P(a_1) \quad P(a_2) \quad \ldots \quad P(a_n)}{\forall x, P(x)}$$

Proof by cases on all natural numbers

Material:

- $\mathbb{N} = \{0, 1, \ldots, n, \ldots\}$
- a predicate (property) P on $\mathbb{N}$

$$\frac{P(0) \quad P(1) \quad \ldots \quad P(n) \ldots}{\forall x, P(x)}$$

In order to prove $\forall x, P(x)$*,*
prove P *on each natural number* n

∞ cases to consider

Does not work for an infinite number of cases

Unless we have a systematical way to construct a proof of $P(n)$ for each n?

Constructing proofs of $P(n)$, $n \in \mathbb{N}$

1. Prove $P(0)$
2. Prove $P(0) \Rightarrow P(1)$
3. Prove $P(1) \Rightarrow P(2)$
4. etc.

From 1. and 2. we get $P(1)$
From the latter and 3. we get $P(2)$
Etc.

At first sight, no progress:
infinite number of proof obligations

Unless ve prove (uniformly) 2. 3. 4. etc. at once:

$$\forall n,\ P(n) \Rightarrow P(1+n)$$

Example: the product of 2 consecutive numbers is even

Formally: $\forall n, \underbrace{\exists k, n.(1+n) = 2.k}_{P(n)}$

- For $n = 0$: we have $n.(1+n) = 0.1 = 0 = 2.0$, taking $k = 0$ yields $P(0)$
- (Uniform) proof of $\forall n,\ P(n) \Rightarrow P(1+n)$
  - For an arbitrary $n \in nat$, assume $P(n)$ i.e. $n.(1+n) = 2.y$ for some y
  - Then $$\begin{aligned}(1+n).(1+1+n) &= (2+n).(1+n)\\ &= 2.(1+n) + 2.y\\ &= 2.(1+n+y)\end{aligned}$$
  - Taking $k = 1+n+y$, we get $P(1+n)$, QED.

Induction on $\mathbb{N}$

$$\frac{P(0) \qquad \forall n, P(n) \Rightarrow P(1+n)}{\forall n, P(n)}$$

$P(n)$ is called the *induction hypothesis*.

Remark: proof by cases

$$\frac{P(0) \qquad \forall n, P(1+n)}{\forall n, P(n)}$$

is a special case of induction – the induction hypothesis is not used.

Sum of the n first natural numbers

$$\sum_{i=1}^{n} i = \frac{n.(1+n)}{2}$$

Let us define $S_n = \sum_{i=1}^{n} i$ . . . by induction!

- $S_0 = 0$
- $S_{1+n} = 1 + n + S_n$

Prove that $\forall n, 2.S_n = n.(1+n)$

- Case $n = 0$: $2.S_0 = 2.0 = 0 = 0.(0+1)$
- Assume, for n arbitrary: $2.S_n = n.(1+n)$
  Then $2.S_{1+n} = 2.(1+n+S_n) = 2.(1+n) + n.(1+n) = (2+n).(1+n) = (1+n).(2+n)$

Back to: the product of 2 consecutive numbers is even

$\forall n, \exists k, n.(1+n) = 2.k$

Obvious, since n is either even or odd

Induction-free proof?

Formally:

1. Let n be a natural number
2. Case analysis on the parity of n:
   - $n = 2.y$ for some y in $\mathbb{N}$, hence $n.(1+n) = 2.\underbrace{y.(1+n)}_{k}$
   - $n = 1 + 2.y$ for some y in $\mathbb{N}$, hence $n.(1+n) = n.(2+2.y) = 2.\underbrace{n.(1+y)}_{k}$

Question

- How do we know that n is either even or odd?

Any natural number is either even or odd

Formally: $\forall n, \underbrace{\exists y, n = 2.y \vee \; n = 1 + 2.y}_{P(n)}$

- Case $n = 0$: take $y = 0$, left, check $0 = 2.0$
- (Uniform) proof of $\forall n,\; P(n) \Rightarrow P(1 + n)$
  - Given an arbitrary n, assume $P(n)$
  - This yields some y such that $n = 2.y$ or $n = 1 + 2.y$
    - If $n = 2.y$, we have $1 + n = 1 + 2y$
    - If $n = 1 + 2.y$, we have $1 + n = 2.(1 + y)$
  - In each of the previous cases, we have
    $\exists z, 1 + n = 2.z \vee \; 1 + n = 1 + 2.z$,
    i.e. $P(1 + n)$, qed.

What is $\mathbb{N}$?

- $0 \in \mathbb{N}$
- if $n \in \mathbb{N}$, then $(1 + n) \in \mathbb{N}$
- all natural numbers are generated from 0 and the previous rule

Induction sticks to this definition of $\mathbb{N}$.

This presentation assumes an intuitive knowledge of:

- numbers
- addition

But only the successor $(1 + \square)$ is needed

$\rightarrow$ Let us take a more basic intuition

What is $\mathbb{N}$? (cont'd)

- $0 \in \mathbb{N}$
- if $n \in \mathbb{N}$, then $S(n) \in \mathbb{N}$
- all natural numbers are generated from 0 and the previous rule

Induction on $\mathbb{N}$

$$\frac{P(0) \qquad \forall n, P(n) \Rightarrow P(S(n))}{\forall n, P(n)}$$

What is $+$?

Defined by induction, like S_n above

- $0 + m = m$
- $S(n) + m = S(n + m)$

Method for defining such functions f:

- provide the returned value when the argument is 0
- provide the returned value when the argument is $S(n)$
  this value may depend on n and on $f(n)$

Note that f may other fixed arguments

Official name in the jargon of logic : *primitive recursion*
(just for your culture)

What is $+$?

(Almost all) basic properties of $+$ are proved by induction

- $\forall n, 0 + n = n \quad \ldots?$
- $\forall n, n + 0 = n \quad \ldots?$

Commutativity, associativity

Similarly for subtraction, multiplication…

Interest: foundations (Coq library); fundamental exercises

Subtelties with induction

Consider the following version of addition

Coq syntax for function application, see below why

- $add\ 0\ m = m$
- $add\ (S\ n)\ m = add\ n\ (S\ m)$

Beyond primitive recursion, see explanation below

Prove $add\ n\ m = n + m$ forall n and m

First try

Prove $add\ n\ m = n + m$ by induction on n
(Previous model) $\rightarrow$ Fails

Second try

Prove $\forall m, add\ n\ m = n + m$ by induction on n
Works

Explanations on *add*

- $add\ 0\ m = m$
- $add\ (S\ n)\ m = add\ n\ (S\ m)$

Reads

- $add\ 0 \ = \ fun\ m \Rightarrow m$
- $add\ (S\ n) \ = \ fun\ m \Rightarrow add\ n\ (S\ m)$

Official name in the jargon of logic :
higher order primitive recursion

More advanced example (homework)

- $\mathit{fib}\ 0 = 1$
- $\mathit{fib}\ 1 = 1$
- $\mathit{fib}\ (S\ (S\ n)) = \mathit{fib}\ n + \mathit{fib}\ (S\ n)$

Harmless shorthand for a truly primitive recursion, where we define $\mathit{fib}\ n$ and $\mathit{fib}\ (S\ n)$ at the same time.

- $\mathit{lfib}\ 0\ a\ b = a$
- $\mathit{lfib}\ (S\ n)\ a\ b = \mathit{lfib}\ n\ b\ (a + b)$

Prove $\forall n, \mathit{lfib}\ n\ 1\ 1 = \mathit{fib}\ n$.

More advanced example (hints)

Use a more abstract version, and prove something on it

- $gfib\ a\ b\ 0 = a$
- $gfib\ a\ b\ 1 = b$
- $gfib\ a\ b\ (S\ (S\ n)) = gfib\ a\ b\ n + gfib\ a\ b\ (S\ n)$

But a more direct proof is also possible. . .

More advanced example (solution)

$\forall n, gfib\,1\,1\,n = fib\ n$
using: $\forall n, gfib\,1\,1\,n = fib\ n\ \wedge\ gfib\,1\,1\,(S\ n) = fib\,(S\ n)$

$\forall n,\ \forall ab, lfib\,(S\,(S\ n))\ a\ b = lfib\ n\ a\ b + lfib\,(S\ n)\ a\ b$

$\forall n, lfib\ n\ a\ b = gfib\ a\ b\ n$
using:
$\forall n, lfib\ n\ a\ b = gfib\ a\ b\ n\ \wedge\ lfib\,(S\ n)\ a\ b = gfib\ a\ b\,(S\ n)$

More direct version

$\forall n,\ \forall a, lfib\ n\,(fib\ a)\,(fib\,(S\ a)) = fib\,(a + n)$