Verifying programs and proofs
part III. prove program properties

Yves Bertot

October 2020

1 Motivating introduction

To prove that programs do what is intended, we need to cover all possible cases in their execution. Most of the time, the programs take their input in types with an infinite numbers of elements. It is thus impossible to check all possible inputs one by one. Instead, we reason on subsets of the types that correspond to different behaviors of the considered programs. This is usually done by following the structure of the program. Thus we reason logically on the various cases that may arise during the execution of programs.

When functions are recursive, this approach relies on a complex logical tool, called *proof by induction*. In this lecture, we want to understand the various aspects of reasoning about program behavior, including the idea of proofs by recursion. We shall restrict this study to programs that compute on numbers and lists.

There are many ways to describe what is the expected behavior of a program. In this course, we will study only a simple approach, where the expected behavior is described with the help of secondary programs that are used either to produce specific inputs or to perform tests on on the output of a program.

For instance, if we wrote a function `evenb` that computes a boolean value that is true whenever the input is an even number (a multiple of 2), we want to write a function that computes all even numbers.

```
Definition mult2 (n : nat) := 2 * n.
```

Then we just want to prove that `evenb` returns the correct value for every result of `mult2`:

```
Lemma evenb_complete :
  forall n : nat, evenb (mult2 n) = true.
```

This is not enough, because the trivial boolean predicate that always returns true would also satisfy this lemma. We may also want to make sure that the function `evenb` accepts only numbers that should be accepted. One way to express this is with the following statement:

```
Lemma evenb_sound :
  forall n : nat, evenb n -> exists y : nat, n = mult2 y.
```

Again, lemma `evenb_sound` is not enough, because the trivial boolean predicate that always returns `false` would also satisfy this lemma.

In the end, every lemma that we write is like a symbolic test that we run on the input program, but the proof shows that this test is satisfied for all possible inputs instead of a random sample. In this sense, the coverage brought by formal proofs is more complete than the coverage brought by random tests and people like to say that we provide 100% correctness, but we should keep in mind that the lemmas may describe only partially the intended behavior of the program (like the predicate `evenb_complete` and `evenb_sound` taken separately) and that the way we consider the inputs (like the function `mult2`) may also be faulty.

2 Reasoning on pattern-matching constructs

A pattern matching construct describes several possible cases of execution. When proving that a program is correct, we need to cover all possibilities. There are commands to decompose the problem and to observe separately each of the cases. Then, it may occur that some cases are inconsistent, because they exhibit assumptions like `0 = 1` or `true = false`. In some other cases, one may have equalities of the form `a::l = b::l'`, from which one should be able to deduce at least `a = b` and `l = l'`. We shall see different approaches for this.

2.1 The `destruct` tactic and the `eqn:` variant

Let us consider the following goal:

```
x : nat
======================
 match x with 0 => true | S p => negb (even p) end = false
 -> x <> 0
```

The conclusion of this goal contains a pattern matching construct. We don't know the value of `x`, but we know that `x` is a natural number, and by consequences `x` may follow one of 2 cases: `x` is `0` or `x` is `S y` for some `y`. When `x` is `0`, the whole pattern-matching construct will compute to `true`, so the goal's conclusion will become

```
 true = false -> 0 <> 0
```

In the other case, we know that the pattern-matching construct will be reduced to `negb (even y)`, so the goal's conclusion will become

```
negb(even p) -> S y <> 0
```

The `destruct` tactic is designed to implement this form of reasoning. It produces as many goals as the number of cases in the argument's type. Here we use `destruct x` and `x` has type `nat`. Because the type `nat` has two constructors, there are two cases: either `x` is 0 or `x` is `S n` for some other natural number `n`.

```
destruct x.
2 subgoals
  ======================
   true = false -> 0 <> 0

Subgoal 2 is
   negb(even n) -> S n <> 0
```

So the tactic simply produced two instances of the goal where `x` is replaced with cases taken from the type. In the first goal, all occurrences of `x` are replaced by `0`, then the pattern matching construct is computed to take this new information into account.

In this example, the two goals can be proved because they mention equalities that cannot hold. This is taken care of by the tactic described in the next section.

The tactic `destruct` operates by replacing all occurrences of the expression that it analyzes by the possible cases. Sometimes this is not sufficient and one needs to add the information that links the expression with the corresponding case in an equality. This is done by adding an `eqn:` directive.

As an illustration consider the following context and goal:

```
Fixpoint mul3 (x : nat) : nat :=
  match x with 0 => 0 | S x => S (S (S (mul3 x))) end.

Lemma example_destruct x :
  match mul3 x with
  | 0 => true
  | S _ => false
  end = true -> x = 0.
Proof.
```

If we just do `destruct (mul3 x)`, we will obtain two goals of the following form

```
true = true -> x = 0

false = true -> x = 0
```

The second goal is easily provable using `discriminate` (see section 2.3), but the first one is not: we lost the information that `mul3 x` must be 0 for this case to arise.

The Coq system provides a way out of this problem by giving the possibility to perform the analysis while keeping the relation between the analyzed expression and the case as an equality. Here is the full script for a complete proof.

```
destruct (mul3 x) eqn:mul3xval.
2 subgoals

  x : nat
  mul3xval : mul3 x = 0
  ============================
  true = true -> x = 0

subgoal 2 is:
 false = true -> x = 0
  destruct x as [ | y].
    intros; reflexivity.
  discriminate mul3xval.
intros truefalse; discriminate truefalse.
Qed.
```

2.2 the `induction` tactic

Reasoning by cases is not adapted for recursive functions, because we often need hypotheses on the values returned by recursive calls. These hypotheses often have the same form as the statement that one attempts to prove. This follows the a well-known pattern seen in proofs about natural number, known as proof by induction.

induction on natural numbers If a predicate on natural numbers `P` is such that `P 0` holds and for every `n` one can deduce `P (S n)` from `P n`, then this predicate holds for every natural number.

induction on lists If a predicate on lists of natural numbers `P` is such that `P nil` holds and for every list `l` and every natural number `a`, if `P l` holds then we can deduce `P (a::l)`, then this predicate holds for every list of natural numbers. This can be generalized to list of any type of elements.

When performing a proof by induction on natural numbers, we also have two cases to study, the first case for the situation where the value is `0` and the second for the situation where the value has the form `S n` for some `n`. In this respect, the induction tactic is very close to the `destruct` tactic. However, when considering the second case, we have more information: we can use an *induction hypothesis* stating that `n` already satisfies the expected predicate. Let's observe such a proof by induction concerning the addition of a number with 0. We first observe the definition of addition:

```
Locate "_ + _".
Notation            Scope
"n + m" := Nat.add n m  : nat_scope
                        (default interpretation)
"x + y" := sum x y   : type_scope
```

```
Print Nat.add.
Nat.add =
fix add (n m : nat) struct n : nat :=
  match n with
  | 0 => m
  | S p => S (add p m)
  end

     : nat -> nat -> nat
```

We see that addition is given as a recursive function where the first argument decreases at each recursive call. By definition `0 + n` computes to `n` in a single step; on the other hand, computing `n + 0` does not do anything directly, but we can prove that it computes to `n`. Here is the proof in Coq:

```
Lemma example_induction_plus : forall n, n + 0 = n.
induction n.
2 subgoals

  ============================
   0 + 0 = 0

subgoal 2 is:
 S n + 0 = S n
```

As expected, we have two cases where `n` is replaced either by `0` or by `S n`. For the first case, immediate computation yields the result.

```
reflexivity.
1 subgoal

  n : nat
  IHn : n + 0 = n
  ============================
   S n + 0 = S n
```

In this goal, the context contains the hypothesis `IHn`, which states exactly that the property we want to prove already holds for `n`. So the induction principle is being used, and the predicate `P` is instantiated with the function

```
fun x => x + 0 = x
```

2.3 the `discriminate` tactic

The datatypes of boolean values, natural numbers, and lists are all described in the Coq system as *inductive types*, where the data may each time correspond to two patterns. We can see this by calling the command `Print`.

```
Print bool.
Inductive bool : Set :=  true : bool | false : bool

Print nat.
Inductive nat : Set :=  O : nat | S : nat -> nat

Print list.
Inductive list (A : Type) : Type :=
    nil : list A | cons : A -> list A -> list A
```

The description of natural numbers means that numbers are either of the form `O` or of the form `S n`. Moreover, it also means that the number `O` is not of the form `S n`[1]. As a result, any goal whose conclusion has the form `O <> S n` should be easily provable. The Coq system provides a specific tactic for that, called `discriminate`. This tactic also takes care of cases where a goal has an arbitrary conclusion but one of its hypotheses is an hypothesis of the form `O = S n`.

Continuing the example given in the previous section, we had two goals that we repeat again here:

```
2 subgoals
  ======================
   true = false -> 0 <> 0

Subgoal 2 is
   negb(even n) -> S n <> 0
```

For the first goal, we use the `intros` tactic and we get a new hypothesis:

```
intros Htf.
  ...
  Htf : true = false
  ======================
   0 <> 0
```

In this goal, the conclusion would be unprovable in an empty context, but the hypothesis `Htf` assumes an equality between two values that are different by definition. This goal can be solved using the `discriminate` tactic or more precisely the `discriminate Htf`.

The second goal has the form

```
  negb(even n) = true -> S n <> 0
```

Here the ultimate conclusion is the negation of an equality between two cases that are forced to be different. So it falls in the same area of reasoning. Here, the `discriminate` tactics also solves the problem.

[1] This is also simply a consequence from the fact that we can define expressions by pattern-matching on natural numbers! In some documents, this is also referred to as a *non-confusion* property.

2.4 the `injection` tactic

We often have to express that the constructors of types like `nat` and `list` are *injective*. In other words, if they give equal outputs for two sets of inputs, then the inputs must be pairwise equal. for lists this is easily expressed with the following example:

```
Lemma example_injection_list :
  forall (a b : nat) (l1 l2 : list nat),  a::l1 = b::l2 ->
    a = b /\ l1 = l2.
intros a b l1 l2 Hq.
  ...
  Hq : a::l1 = b::l2
  ============================
   a = b /\ l1 = l2
```

In this goal, the hypothesis `Hq` describes an equality between two composed lists. The conclusion expresses that the list components correspond to each other. To go from `Hq` to the conclusion, we call the tactic `injection` with the name of the hypothesis.

```
injection Hq.
  ...
  ============================
   l1 = l2 -> a = b -> a = b /\ l1 = l2
intros ql qa; rewrite ql qa; split; reflexivity.
Qed.
```

In the generated goal, two new implications are created, with the equalities between components appearing as left-hand sides of these implications. The last two lines of the example show how to use these hypotheses.

Similarly, if we know two equal numbers that respect the `S` pattern, we can deduce that the sub-components are equal.

```
Lemma example_injection_nat :
  forall (a b : nat),  Sa = S b -> a = b.
intros a b Hq.
  ...
  Hq : S a = S b
  ============================
   a = b
injection Hq.
  ...
  ============================
   a = b -> a = b
intros q1; exact q1.
Qed.
```

3 Manipulating function computation

In goals and logical statements, the Coq system manipulates functions without executing them. We sometimes need to force at least a few steps of computation.

3.1 The `unfold` tactic

The first approach is to simply require that the system expands the definition. The word used in Coq tactics is `unfold`.

```
Definition add3 (n : nat) := n + 3.

Lemma example_add3 : forall n, add3 n = 3 + n.
intros n.
  ...
  ======================
   add3 n = 3 + n.
```

At this point, we would like to replace `add3 n` with the expression it computes. We use the `unfold` tactic.

```
unfold add3.
  ======================
    n + 3 = 3 + n
```

3.2 The `simpl` tactic

When dealing with a recursive function, the `unfold` tactic often makes goals unreadable, because it expands the value of the recursive function into something that repeats the text of the recursive function several times. To avoid this, there are tactics specifically tuned to handle recursive functions. These tactics are called `simpl` and `cbn`.

The example we have already seen about reasoning on the addition function provides an illustration for this. Let start again with this proof.

```
Lemma example_induction_plus : forall n, n + 0 = n.
induction n.
reflexivity.
1 subgoal

  n : nat
  IHn : n + 0 = n
  ============================
   S n + 0 = S n
```

Here, we can request that Coq performs a little computation with `S n + 0`. We simply need to call the `simpl` tactic:

```
simpl.
  ...
  IHn : n + 0 = n
  ============================
   S (n + 0) = S n
```

The left hand side of the equality in this goal's conclusion is an occurrence of the left hand side of the hypothesis `H`. We can rewrite and conclude the proof.

```
rewrite IHn; reflexivity.
Qed.
```

Often, `cbn` is practical to use, because it makes it possible to control what functions will be unfolded. Interested readers should look at the Coq documentation concerning this tactic.

3.3 Manual computation: the `change` tactic

Sometimes the `simpl` tactic performs too much computation. In this case, it is a good idea to state explicitely the result that we want to see after computation, as long as this result really corresponds to a computation. Here is an example.

```
Lemma example_change_plus :
  forall n m p, (1 + n) * m = p -> (1 + (1 + n)) * m = m + p.
intros n m p H.
1 subgoal

 H : (1 + n) * m = p
  ============================
   (1 + (1 + n)) * m = m + p
change ((1 + (1 + n)) * m) with (m + (1 + n) * m).
  ...
  ============================
   m + (1 + n) * m = m + p
rewrite H; reflexivity.
Qed.
```

3.4 Manual computation with the `replace` tactic

The tactic `change` performs replacements only if the two expressions are the same modulo computation. Sometimes, we want to relax the condition, perform replacement, and keep the obligation to prove the equality between the two expressions for later. For this we use the `replace` tactic. This tactic produces a second goal, because the equality between the two expressions still needs to be proved.

3.5 Generating one-step recusion lemma

Using the `change` tactic gives the user complete control on what is executed, but it is very cumbersome to use, as it requires that one writes a lot of text. One efficient approach is to provide an unfolding lemma for the function that one produces. This can easily be done by copying the body of the function in an equality and encapsulating it the universal quantifications that fit. The proof can usually be done by a simple case analysis on the first argument of the function to be analysed. Here is an example for a function of multiplication by 3:

```
Fixpoint mult3 (n : nat) : nat :=
match n with 0 => 0 | S p => 3 + mult3 p) end.

Lemma mult3_step (n : nat) :
  mult3 n =
  match n with 0 => 0 | S p => 3 + mult3 p) end.
Proof.
intros n; case n; reflexivity.
Qed.
```

4 A complete proof

We first define a function that as input a number n and a list l, and returns a list containing elements of l multiplied by successors of n. The first element of the result is the first element of l multiplied by $n+1$, the second element of the result is the second element of l multiplied by $n+2$, and so on. This can be described easily as a recursive function on lists:

```
Require Import List Arith.
Fixpoint mulsl (n : nat) (l : list nat) :=
  match l with
  | a :: l1 => a * (n + 1) :: mulsl (n + 1) l1
  | nil => nil
  end.
```

We know wish to show that this function is actually injective in its second argument. Here is the statement

```
Lemma mulsl_injective n l1 l2 : mulsl n l1 = mulsl n l2 -> l1 = l2.
Proof.
```

We wish to perform this proof by induction on one of the lists `l1` and `l2`. When observing the function `mulsl` carefully, we see that the list that must be shown to be equal to `l1` as it varies cannot be the same. So `l2` should not be fixed in the context of the goal, but rather a universally quantified variable. Similarly, we see that in recursive calls the variable `n` changes, so again `n` should not be

fixed in the context. To make this change of shape of the goal, we use the `revert` tactic.

```
revert n l2.
induction l1 as [ | a l1 IH].
```

The proof by induction makes use of a predicate on lists of natural numbers that has the following shape.

```
fun l => forall (n : nat)(l : list nat),
          mulsl n l = mulsl l2 -> l = l2
```

We now have two goals, where the first one corresponds to the case where this predicate is applied to the empty list `nil`, and the second one is applied to a non-empty list `a :: l1`. Thus, the first goal has the following text

```
============================
forall (n : nat) (l2 : list nat),
   mulsl n nil = mulsl n l2 -> nil = l2

intros n [ | b l2].
  intros; reflexivity.
simpl.
discriminate.
```

The line `intros n [ | b l2]` is equivalent to `intros n tmp; destruct tmp as [ | b l2]`. Because of this line, there are two goals. The first one requires that we prove `nil = nil`, done by reflexivity. The second goal contains an implication of the form `mulsl n nil = mulsl n (b :: l2) -> ...` After applying the computation rules of `mulsl` this boils down to an equality between different constructors of a datatype. Such a premise is self-contradictory and `discriminate` recognizes it. We wrote `simpl` to make the computation happen so that a human observer of the proof can see the result, but this step is not mandatory and using `discriminate` without the `simpl` step would already work.

The second goal generated by the induction step concerns `a :: l1` (the `cons` form), but since it is an induction proof, we can already use the fact we wish to prove for the sublist `l1`. Thus the goal has the following form:

```
a : nat
l1 : list nat
IH : forall (n : nat) (l2 : list nat),
     mulsl n l1 = mulsl n l2 -> l1 = l2
============================
forall (n : nat) (l2 : list nat),
mulsl n (a :: l1) = mulsl n l2 -> a :: l1 = l2
```

So in this goal, the fact that we already know about `l1` is embodied in the hypothesis `IH` (as we prescribed in the `induction` call).

We can now fix a value `n` for the universal quantification, and reason on the possible cases for the list. In this case we know that `l1` is of the `cons` form, after computation `mulsl` will return a result of the `cons` form. If the second list if of the `nil` form, the computation of `mulsl` returns `nil` and we the equality between `mulsl` results is again an inconsistent one, and the goal is easily solved using `discriminate`.

```
intros n [ | b l2]; simpl.
  discriminate.
```

It remains to study the case where both lists are in `cons` form. The goal then has the following shape:

```
  a : nat
  l1 : list nat
  IH : forall (n : nat) (l2 : list nat),
       mulsl n l1 = mulsl n l2 -> l1 = l2
  n, b : nat
  l2 : list nat
  ============================
  a * (n + 1) :: mulsl (n + 1) l1 =
  b * (n + 1) :: mulsl (n + 1) l2 ->
  a :: l1 = b :: l2
```

Because of the `simpl` step after introducing `b` and `l2`, both computations of `mulsl` have produced results that are both in `cons` form. We know need to use the `injection` tactic to make this equality between two lists transform into two equalities, the first concerning the first elements of these lists, and the second concerning the results of these lists.

```
intros Heq; injection Heq as heads tails.
```

The new hypothesis `tails`, in combination with the induction hypothesis, will make it possible to prove that `l1` and `l2` are equal. We use this fact directly, knowing that the proof will be kept for later.

```
replace l2 with l1.
```

Similarly, the fact `a = b` should be deductible from the hypothesis `heads`. To make this step, we look in the database for a theorem that mentions equalities between results of multiplications. The command for this database inquiry is as follows:

```
Search (_ * ?x = _ * ?x).
```

By using two instances of the named place-holder `?x`, we are able to pinpoint exactly a theorem where the same expression appears at the corresponding place on both side of the equality. The result of the `Search` command is as follows:

```
heads: a * (n + 1) = b * (n + 1)
Nat.mul_cancel_r: forall n m p : nat,
     p <> 0 -> n * p = m * p <-> n = m
```

The hypothesis `tails` is listed, but also a theorem about an equivalence between equalities, under the condition that the common factor `p` is non-zero. We first modify the hypothesis `heads` using this theorem, and we then use the modified hypothesis to modify the goal. What we obtain is a trivial equality, and the `easy` tactic can get rid of this goal.

```
  rewrite Nat.mul_cancel_r in heads.
    rewrite heads; easy.
```

At this point, we still have to prove that the common factor is non-zero, and this can be done by modifying the statement to make it appear that the equality is between two different constructor forms of type `nat`.

```
  rewrite Nat.add_1_r; easy.
```

We still have to prove that the replacement of `l2` by `l1` is justified. For this we use the induction hypothesis. However, in this induction hypothesis, the universally quantified variable does not appear in the hypothesis' conclusion, so we need to guide the proof system by showing how the instantiation will happen. Here is one possibility.

```
apply IH with (n := n + 1).
exact tails.
Qed.
```

5 Exercises

1. Define a function `lo` that takes a natural number `n` as input and returns the list containing the first `n` odd natural numbers. For instance `lo 3 = 5::3::1`.
2. Prove that `length (lo n) = n`.
3. Define a function `sl` that takes a list of natural numbers as input and returns the sum of all the elements in the list.
4. Prove that `sl (lo n) = n * n`.
5. We define a function `add` with the following code:

   ```
   Fixpoint add x y := match x with 0 => y | S p => add p (S y) end.
   ```

   Prove the following lemmas:

   (a) `forall x y, add x (S y) = S (add x y)`

(b) `forall x, add x 0 = x`

(c) `forall x y, add (S x) y = S (add x y)`

(d) `forall x y z, add x (add y z) = add (add x y) z`

(e) `forall x y, add x y = x + y`

6. In the exercises part of the first chapter of these course notes, you are required to define a function that describes when a list of numbers is a licit representation of a natural number (it verifies that all digits are less than 10) and a function that computes the successor of a number when represented as a list of digits. Add the function `to_nat` that maps any list of digits to the natural number it represents and show that the successor function is correct in this context.

6 More information

You can use the book [1] (available in French on internet, otherwise you should find English versions at the library) and the reference manual [4]. There is also a tutorial in French [7]. There are also tutorials on the web [5, 3].

References

[1] Y. Bertot and P. Castéran. *Interactive Theorem Proving and Program Development, Coq'Art:the Calculus of Inductive Constructions*. Springer-Verlag, 2004. Version française `http://www-sop.inria.fr/members/Yves.Bertot/coqartF.pdf`

[2] B. Pierce et al. *Software Foundations* `http://www.cis.upenn.edu/~bcpierce/sf/`

[3] Y. Bertot *Coq in a Hurry* Archive ouverte "cours en ligne", 2008. `http://cel.archives-ouvertes.fr/inria-00001173`

[4] The Coq development team. *The Coq proof Assistant Reference Manual*, Ecole Polytechnique, INRIA, Université de Paris-Sud, 2004. `http://coq.inria.fr/doc/main.html`

[5] G. Huet, G. Kahn, C. Paulin-Mohring, *The Coq proof Assistant, A Tutorial*, Ecole Polytechnique, INRIA, Université de Paris-Sud, 2004. `http://coq.inria.fr/V8.1/tutorial.html`

[6] E. Giménez, P. Castéran, *A Tutorial on Recursive Types in Coq*, INRIA, Université de Bordeaux, 2006. `http://www.labri.fr/Perso/~casteran/RecTutorial.pdf.gz`

[7] A. Miquel, *Petit guide de survie en Coq*, Université de Paris VII. `http://www.pps.jussieu.fr/~miquel/enseignement/mpri/guide.html`