

Meta-programación Tipada de Meta-programas Tipados en Coq

Ignacio Tiraboschi

19 de enero de 2020

Índice general

Capítulo 1

Una introducción

1.1. Introducción

1.2. Coq

En este capítulo introduciremos las características más relevantes del asistente de pruebas Coq. El objetivo de este capítulo es introducir todos los conceptos que utilizaremos más adelante, pero esto significa que no es una introducción completa.

1.2.1. La historia de Coq

El desarrollo de Coq comenzó en 1984 con el apoyo de INRIA como el trabajo de Gérard Huet y Thierry Coquand. En ese momento Coquand estaba implementado un lenguaje llamado *Calculus of Constructions* cuando el 1991 una nueva implementación basada en un Calculus of Inductive Constructions extendido comenzó a ser desarrollado tomando el nombre de Coq.

Ahora mismo Coq es desarrollado por más de 40 desarrolladores activos y es reconocido como uno de los asistentes de prueba principales.

1.2.2. Los lenguajes Coq

Coq no es técnicamente un lenguaje de programación, si no un asistente de pruebas. Pero podemos encontrar múltiples lenguajes dentro de Coq que nos permiten expresarnos.

- *Gallina*: este es el lenguaje de especificación de Coq. Permite desarrollar teorías matemáticas y probar especificaciones de programas.

Utilizaremos extensivamente este lenguaje para definir nuestros programas en `Mtac2`.

- *Ltac*: este es el lenguaje en que se definen las *tácticas* de Coq. Dado que Coq está centrado en las tácticas, Ltac es una de las partes centrales de Coq.
- *Vernacular*.

Aunque Coq no es un lenguaje de programación propiamente dicho, este puede ser utilizado como un lenguaje de programación funcional. Estos programas serán especificados en Gallina. Dada la naturaleza de Coq como provador de teoremas, estos programas son funciones puras, es decir, no producen efectos secundarios y siempre terminan.

1.2.3. La teoría de Coq

Calculus of Inductive Constructions es la base de Coq. Este es un cálculo lambda tipado de alto orden y puede ser interpretado como una extensión de la correspondencia Curry-Howard.

Llamaremos *Terms* (o términos) a los elementos básicos de esta teoría. Terms incluye *Type*, *Prop*, variables, tuplas, funciones, entre otros. Estas son algunas de las herramientas que utilizaremos para escribir nuestros programas.

1.2.4. Set vs. Prop vs. Type

1.2.5. Tipos de datos y Funciones

Ahora aprenderemos a codificar nuestros programas funcionales en Coq. Lo primero que debemos entender es que operamos sobre *términos* y algo es un término si tiene tipo. Coq viene con muchos tipos predefinidos, por ejemplo `unit : Type` es un tipo, y a su vez es un término. También lo es `true : bool`, etcetera.

Veamos cómo se define el tipo `bool`:

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

Como vemos, es un tipo inductivo, especificado por la keyword `Inductive`, con dos constructores `true` y `false`.

Ahora, veamos un ejemplo un poco más complicado, \mathbb{N} .

```

Inductive N : Set :=
| 0 : N
| S : N → N .

```

Notemos que el constructor `S` es una función que recibe un término de tipo `nat`, es decir, `N` es un tipo recursivo. Por ejemplo el término `S (S 0)` es de tipo `N` y representa al número 2.

Para complicar la situación un poco más, veamos el tipo de `list`.

```

Inductive list (A : Type) : Type :=
| nil : list A
| cons : A → list A → list .

```

Este tipo es un tipo polimórfico dado que requiere de un `A : Type` para ser un término. Por ejemplo, un posible elemento es `cons (S 0) nil : list N` que representa a la lista con un único elemento 1.

1.3. Mtac2

1.3.1. Mónadas

1.3.2. Confección de Meta-programas

1.3.3. Meta-tácticas

1.3.4. Alternativas

1.4. Telescopios

En `Mtac2` llamamos *telescopios* a una estructura de datos inductiva que permite expresar una secuencia de tipos posiblemente dependientes de largo arbitrario.

```

Inductive MTele : Type :=
| mBase : MTele
| mTele {X : Type} (F : X → MTele) : MTele.

```

El tipo `MTele` crea una cadena de abstracciones. Este se codifica a través de funciones, lo que permite que sean dependientes, es decir, un telescopio puede tener elementos que dependan de elementos anteriores.

Los telescopios, junto con las funciones que lo acompañan, serán claves a la hora de poder expresar nuestro problema y nuestra solución.

Los telescopios no se caracterizan por su simpleza. Para comenzar a estudiarlos podemos pensarlos en jerarquías. La primera jerarquía serían los

telescopios en sí, elementos de tipo `MTele`. Definiremos la siguiente notación para poder referirnos a estos de manera más simple.

$$\begin{aligned} \text{mBase} &\equiv []_t \\ \text{mTele } (\lambda T : \text{Type} \Rightarrow \text{mTele } R : T \rightarrow \text{Type}) &\equiv [T : \text{Type}; > R : (T \rightarrow \text{Type})]_t \\ \text{mTele } (\lambda T : \text{Type} \Rightarrow \text{mTele } t : T) &\equiv [T : \text{Type}; > t : T]_t \end{aligned}$$

Las otras jerarquías entran en juego cuando utilizamos las distintas funciones telescópicas.

La principal función es `MTele_Sort`. Dado un telescopio t de largo n y un $s : \text{Sort}$, esta función computa $\forall x_1 \dots x_n, s$, es decir, retorna el tipo expresado por ese telescopio y el `Sort` dado. Utilizaremos `MTele_Ty` y `MTele_Prop` para expresar `MTele_Sort Types` y `MTele_Sort Props` respectivamente. Elementos de este tipo serán de la segunda jerarquía.

También debemos poder referirnos a valores de estos tipos telescópicos. Esto significa que dado un elemento de tipo `MTele_Sort s t`, podamos conseguir un elemento de tipo s . En terminos matemáticos, $\lambda x_1 \dots x_n \Rightarrow T x_1 \dots x_n$ con $T : \text{MTele_Sort } s t$. Consideraremos que estos elementos pertenecen a la tercera y última jerarquía.

Deseríamos que este fuera el final de los telescopios, pero nos falta el último ingrediente que nos proveerá del poder que buscamos. Con `MTele_In` podemos ganar acceso a múltiples tipos y valores telescopios al mismo tiempo, así siendo capaces de computar un nuevo tipo telescópico.

Capítulo 2

Lift

2.1. Motivación

Mtac2 nos permite definir funciones monádicas. Estas cuentan con ciertas ventajas. Un ejemplo de un meta-programa es el siguiente.

```
Definition list_max_N :=  
  mfix f (l: list N) : l <> nil → M N :=  
    mtmmatch l as l' return l' <> nil → M N with  
    | [? e] [e] ⇒m λ nonE ⇒ M.ret e  
    | [? e1 e2 l'] (e1 :: e2 :: l') ⇒m λ nonE ⇒  
      let m := Nat.max e1 e2 in  
      f (m :: l') cons_not_nil  
    | [? l' r'] l' ++ r' ⇒ (* cualquier cosa *)  
  end.
```

Esta función calcula el máximo de una lista de números \mathbb{N} : **Set**. Dado que en el último caso de **mtmmatch**, **match** monádico generalizado, analiza una expresión con una función, y no un constructor, es imposible implementar esto sin **Mtac2**. Notar que tampoco podemos utilizar **mmatch** ya que el tipo que retornamos es $l' <> nil \rightarrow M\ \mathbb{N}$.

Ahora supongamos que deseamos parametrizar \mathbb{N} y tener una función que acepte múltiples conjuntos. Sea

```
Definition max (S: Set) : M (S → S → S) :=  
  mmatch S in Set as S' return M (S' → S' → S') with  
  | N ⇒ M.ret Nat.max  
end.
```

la función que retorna la relación máximo en un conjunto S . A primera vista nuestra idea podría funcionar, es decir, conceptualmente no es incorrecta.

```

Definition list_max (S: Set) :=
  max ← max S; (* error! *)
  mfix f (l: list S) : l' <> nil → M S :=
    mtmmatch l as l' return l' <> nil → M S with
    | [? e] [e] ⇒m λ nonE ⇒ M.ret e
    | [? e1 e2 l'] (e1 :: e2 :: l') ⇒m λ nonE ⇒
      m ← max e1 e2;
      f (m :: l') cons_not_nil
    end.

```

Al intentar que Coq interprete la función veremos que esta función no tipa. Esto es debido a la signatura de `bind`. Nuestro `mfix` no puede unificarse a `M B`, ya que tiene tipo `f : forall (l: list S) l' <> nil → M S`.

`bind : forall A B, M A → (A → M B) → M B`.

Solucionar esta situación específica no es un problema. Una alternativa es introducir los parámetros de la función y beta-expandir la definición del fixpoint. Otra es codificar un nuevo `bind` que tenga el tipo necesario. El problema será que ambas soluciones son específicas al problema, entonces en cada situación debemos volver a implementar alguno de estos recursos.

Es por eso que nuestro proyecto es la codificación de un nuevo meta-programa LIFT que automáticamente puede generalizar meta-programas con las dependencias necesarias para que sea utilizado en el contexto. En nuestro ejemplo, con LIFT podemos generalizar `bind` consiguiendo un nuevo meta-programa que se comporta como la función original pero con una signatura distinta, permitiendo su uso.

2.2. LIFT

Denominamos `lift` a la función desarrollada en este trabajo. La misma tiene la tarea de agregar dependencias a meta-programas de manera quasi automática: solo requiere un telescopio.

Lift se basa en analizar los tipos de las funciones y modificarlos añadiendo dependencias triviales en los tipos que se encuentran bajo la mónada, generando así nuevos meta-programas más generales. El significado real de este funcionamiento será trabajado en este capítulo a través de ejemplos que mostrarán el comportamiento de `lift`.

2.2.1. Lifting funciones simples

Una de las funciones monádica más simples es `ret : ∀ A, A → M A`, uno de los operadores monádicos. Supongamos que nos interesa tener

$\text{ret}^\wedge : \forall (A : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \text{Type}), (\forall n n', A n n') \rightarrow (\forall n n', M (A n n'))$

Ahora, nuestro trabajo es poder definir un telescopio que se amolde a la información que necesitamos agregar. En este caso el telescopio es el siguiente: $t := [n : \mathbb{N}; > n' : \mathbb{N}]_t$ dado que queremos que A dependa de dos \mathbb{N} . Efectivamente $\text{ret}^\wedge := \text{lift ret } t$.

Este ejemplo es simple porque no involucra funciones. Pero en ese caso, estudiemos que sucede con `bind`.

$\text{bind} : \forall A B, M A \rightarrow (A \rightarrow M B) \rightarrow M B$

Bind es el operador monádico más importante y es central en estos problemas. Supongamos que queremos utilizar el telescopio $t := [T : \text{Type}; > l : \text{list } T]_t$ donde `list T` es el tipo de las listas con elementos de tipo T . Al momento de liftearlo, no es obvio cual debería ser el resultado, así que veamoslo.

$\text{bind}^\wedge : \forall A B : \forall T : \text{Type}, \text{list } T \rightarrow \text{Type},$
 $(\forall (T : \text{Type}) (l : \text{list } T), M (A T l)) \rightarrow$
 $(\forall (T : \text{Type}) (l : \text{list } T), A T l \rightarrow M (B T l)) \rightarrow$
 $(\forall (T : \text{Type}) (l : \text{list } T), M (B T l))$

Lo más importante es que si observamos la definición de esta función es sumamente simple.

$\lambda (A B : \forall T : \text{Type}, \text{list } T \rightarrow \text{Type})$
 $(\text{ma} : \forall (T : \text{Type}) (l : \text{list } T), M (A T l))$
 $(f : \forall (T : \text{Type}) (l : \text{list } T), A T l \rightarrow M (B T l))$
 $(T : \text{Type}) (l : \text{list } T) \Rightarrow$
 $\text{bind } (A T l) (B T l) (\text{ma } T l) (f T l)$

Esta definición es efectivamente la que buscábamos y funciona perfectamente. El otro aspecto que seguiremos observando es que `Lift` no genera información innecesaria en la función destino.

2.3. Aspectos Técnicos

En esta sección discutiremos los aspectos técnicos de LIFT. Comenzaremos discutiendo su funcionamiento básico y de ahí escalaremos a los detalles más profundos.

2.3.1. TyTree

En términos generales LIFT es un fixpoint sobre un telescopio con un gran análisis por casos sobre los tipos. Entonces surge un problema: ¿Cómo podemos hacer *pattern matching* en los tipos de manera sintáctica? La solución

es utilizar un reflejo de los mismos, de manera de que podamos expresarlos de manera inductiva.

```

Inductive TyTree : Type :=
| val{m : MTele} (T : MTele_Ty m) : TyTree
| M (T : Type) : TyTree
| MFA{m : MTele} (T : MTele_Ty m) : TyTree
| In(s : Sort) {m : MTele} (F : accessor m → s) : TyTree
| imp(T : TyTree) (R : TyTree) : TyTree
| FATEleVal{m : MTele} (T : MTele_Ty m)
  (F : ∀ t : MTele_val T, TyTree) : TyTree
| FATEleType(m : MTele) (F : ∀ (T : MTele_Ty m), TyTree) : TyTree
| FAVal(T : Type) (F : T → TyTree) : TyTree
| FAType(F : Type → TyTree) : TyTree
| base(T : Type) : TyTree
.

```

Con este tipo podemos reescribir todas las signatures de funciones. Varios de los constructores, como por ejemplo `MFA` o `FATEleVal`, tendrán sentido más adelante, dado que reflejan elementos en funciones lifteadas.

Una de las propiedades principales de este reflejo es que a primera vista parece que un tipo puede escribirse de múltiples formas en `TyTree`, pero los tratamos de manera distinta y por eso podemos plantear una biyección entre `Type` y `TyTree`. Utilizamos la función `to_ty : TyTree → Type` para transformar un `TyTree` en su `Type` correspondiente. Notar que esta función no es monádica, y eso es principal, ya que nos permite utilizar la función en las signatures que definimos. En cambio, la función `to_tree : Type → M TyTree` necesariamente es monádica ya que debemos hacer un análisis sintáctico en los tipos de Coq. Haremos uso extensivo de `to_ty`.

Utilizaremos `Constt` para expresar `tyTree_Const`, donde `Const` representa alguno de los nombres de los constructores de la definición de `TyTree`.

Ahora tomaremos una función `f` e iremos modificando su signature para mostrar este reflejo de tipos. Para simplificar escribiremos $P \equiv Q$ para expresar que un tipo es equivalente a un `TyTree` aunque no sea técnicamente correcto en Coq.

$f : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \mathbb{N}$

Su tipo traducido es

$f : \text{imp}(\text{base } \mathbb{N}) (\text{imp} (\text{base } \mathbb{N}) (\text{base } \mathbb{N}))$

Dado que no hay dependencias de tipos, podemos utilizar `imp`. Ahora si parametrizamos \mathbb{N} por cualquier tipo.

$\forall A, A \rightarrow A \rightarrow A$

```

≡
FAType(λ A ⇒ imp(base A) (imp (base A) (base A)))

```

Con **FAType** podemos observar claramente la dependencia de **A**.

Para expresar la dependencia de un valor utilizamos **FAVal**.

```

∀ A (B : A → Type) (a : A), (B a)
≡
FAType(λ A ⇒ FAType(λ B : A → Type ⇒ FAValA (λ a ⇒ base(B a))))

```

El centro de nuestro trabajo son las funciones monádicas, esto significa utilizar **M**.

```

ret : ∀ A, A → M A
≡
ret : FAType(λ A ⇒ imp(base A) (M A))

```

Es importante notar que podemos encontrar usos de **M** en múltiples secciones de la signatura, solo se matcheará **M** en **LIFT** con el retorno de la función.

2.3.2. El tipo de **LIFT**

Analicemos la definición de **LIFT**.

```

Fixpoint lift (m : MTele) (U : ArgsOf m) (p l : bool) (T : TyTree) :
  ∀ (f : to_ty T), M m:{ T : TyTree & to_ty T } := ...

```

Dentro de esta signatura vemos elementos conocidos como el telescopio **m : MTele** que anuncia las dependencias, un **T : TyTree** que representa el tipo de la función a *liftear* y la **f : to_ty T** de tipo **T**. También conocemos

```

M m:{ T : TyTree & to_ty T }
(* Σ-type con un tipo y un elemento del mismo *)

```

En el retorno conseguimos un nuevo **T' : TyTree** que representa la signatura de la función **f** lifteada, y un elemento de tipo **T'**, es decir, la nueva función.

Ahora, ¿qué representan los argumentos que no mencionamos? Es **fácil** comprender quienes son **p** y **l**.

- **p**: la polaridad. Comienza con valor **true**. Este solo se modifica cuando matcheamos **imp** en **LIFT**. No es útil para **lift_in** y representa...
 - **p = true**:
- **l**: comienza en **false**. Representa si nos encontramos a la derecha o izquierda de una implicación de manera inmediata. Es útil para...

En cuanto a $U : \text{ArgsOf } m$, U representa los argumentos que el telescopio añade, y es nuestro truco para poder hacer funcionar LIFT, ya que nos permite transportar argumentos de manera descurricada. Esto quiere decir que transporta a los argumentos en un “contenedor”. Lo que sucede es que cuando encontremos un tipo A cualquiera en nuestra función, este tipo puede o no estar bajo la mónada. En el caso de estarlo debemos modificarlo, es decir, reemplazar $A : \text{Type}$ por $A : \text{MTele_Ty } t$ con $t : \text{MTele}$.

2.3.3. TyTrees monádicos

Ahora nos centraremos en cómo representar funciones lifteadas con **TyTree**. Esto significa entender aún más detalles de los tipos dependientes de telescopios.

En este caso observamos el tipo de **ret** lifteado. Es importante notar que está parametrizado a cualquier telescopio, lo que significa que liftear una función no necesariamente necesita de un telescopio específico.

```

λ m : MTele ⇒
  FATEleType m
  (λ A : MTele_Ty m ⇒
    imp(In Type_s (λ a : accessor m ⇒ acc_sort a A))
    (MFA A))

```

Este es el tipo lifteado de **ret**. En este podemos observar el uso de **FATEleType**, **In** y **MFA**.

Con **FATEleType** podemos introducir tipos telescopios, se comporta de igual manera que **FAType**.

MTele_In nos permite adentrarnos a un tipo telescópico, momentáneamente introduciendo todos los argumentos con un **accessor** y trabajando sobre el tipo de manera directa. Dado que no tenemos interés real en usar estos argumentos del telescopio, no tenemos que hacer demasiado trabajo, simplemente generamos tipos de manera trivial, es decir, ignorando los argumentos. Esto es expresado por **acc_sort a A** que en verdad está produciendo el tipo $\forall x_1 \dots x_n, A x_1 \dots x_n$.

Utilizamos **MFA** para representar tipos monádicos cuantificados. Definimos **MFA** en **Mtac2** de la siguiente manera.

```

Definition MFA {t} (T : MTele_Ty t) := (MTele_val
  (MTele_C Type_sort Prop_sort M T)).

```

Sea $t : \text{MTele}$ de largo n y $T : \text{MTele_Ty } t$, **MFA T** representará

$\forall x_1 \dots x_n, M (T x_1 \dots x_n)$

Finalmente, en el caso anterior, interpretando los tipos de una manera más matemática, tomamos un valor $\forall x_1 \dots x_n, A x_1 \dots x_n$ y retornamos $\forall x_1 \dots x_n, M (T x_1 \dots x_n)$.

En el caso de `ret`, la signature `In Types` ($\lambda a : \text{accessor } m \Rightarrow \text{acc_sort } a \ A$) es simplemente equivalente a `val A` pero Coq no puede inferir esto directamente. La forma en que hemos definido LIFT utiliza esta forma más general en todos los casos.

Si concretizamos el telescopio conseguiremos una signature más parecida a la matemática y la función resultante será muy simple. Para mostrar esto supongamos que tenemos un telescopio $t := [x_1 : T_1; > x_2 : T_2; > x_3 : T_3]_t$. Luego,

```
ret^ :=
λ (A : T1 → T2 → T3 → Type)
  (x : ∀ (t1 : T1) (t2 : T2) (t3 : T3), A t1 t2 t3)
  (x1 : T1) (x2 : T2) (x3 : T3) ⇒ r (A x1 x2 x3) (A x1 x2 x3)
: ∀ T : T1 → T2 → T3 → Type,
  (∀ (t1 : T1) (t2 : T2) (t3 : T3), T t1 t2 t3) →
  ∀ (t1 : T1) (t2 : T2) (t3 : T3), M (T t1 t2 t3)
```

Solo nos falta hablar de `FATeleVal`. Tomemos una función de ejemplo con la siguiente signature.

```
∀ (T : Type) (R : T → Type) (t : T), M (R t)
≡
FAType (λ T ⇒ FAType (λ R : T → Type ⇒ FAVAl T (λ t ⇒ M (R t))))
```

La traducción será muy directa.

```
FAType (λ T ⇒ FATeleType (λ R : T → Type ⇒ FATeleVal T (λ t ⇒ M (R t))))
```

Notar que el primer constructor no es reemplazado ya que `T` no se encuentra bajo la mónada en ningún momento.

2.3.4. El algoritmo

A continuación haremos un recorrido paso a paso de como `ret` es lifteado.

1. Matcheamos `FAType F` donde $F : \text{Type} \rightarrow \text{TyTree}$. Generamos un tipo arbitrario `A` y calculamos `is_m (F A) A`. Esta función retornará `true` ya que efectivamente el tipo `A` se encuentra bajo la mónada en la signature `F A`. Entonces generamos un nuevo `A : MTele_Ty t` aplicamos LIFTde manera recursiva sobre `F (apply_sort A U)`.
2. Ahora debemos liftear algo del siguiente tipo.

$F(\text{apply_sort } A \ U) \equiv$
 $\text{imp}(\text{base}(\text{apply_sort } A \ U)) \ (M(\text{apply_sort } A \ U))$

Nuestra expresión matcheará con el caso `imp` en el cual comenzaremos evaluando `contains_u m U X` donde `X` será `base(apply_sort A U)`. En nuestro caso es claro que se cumple dado que hemos reemplazado `A` por un tipo descurricado. Estos nos lleva a tener que utilizar `lift_in`.

3. Tendremos una llamada `lift_in U (base(apply_sort A U))` matcheando el caso correspondiente. `lift_in` retornará una Σ -Type con un valor $F : \text{accessor } t \rightarrow \text{Type}_s$ y una prueba de que el tipo `base(apply_sort A U)` es igual a `F(uncurry_in_acc U)`. `uncurry_in_acc U` nos retorna el accesor trivial de `U`. Lo importante es que sabemos que `F(uncurry_in_acc U)` es igual al lado izquierdo de la implicación de `ret`.

4. Ahora volvemos a LIFT y generamos un valor $x : X'$ con

$X' := \text{MTele_val}(\text{MTele_In } \text{Type}_s \ F)$

Es decir, un valor `x` del tipo resultante de liftear `X` en el paso anterior. Lo que resta es tomar nuestra función `f` de tipo $X' \rightarrow Y$ y liftear. Esto significa liftear `f x`. No entraremos en los detalles específicos de como expresar `f x` ya que son meros métodos de escribirlo y no portan un valor real.

5. El último paso es `lift t U Y (f x)` sabiendo que $Y = M(\text{apply_sort } A \ U)$, entonces matcheamos con el caso correspondiente. Primero, lo que hacemos es abstraer a `U` de `f` obteniendo

$f = (\lambda \ U \Rightarrow \text{to_ty}(\text{tyTree_M}(\text{apply_sort } A \ U)))$

y luego aplicamos `f` a `curry`, de esa manera, la función pasa a tener tipo `to_ty(MFAA)`.

$f : \forall x_1 \ x_n \Rightarrow \text{MFA}(A \ x_1 \ \dots \ x_n)$

6. Finalmente, LIFT retorna una $T' : \text{TyTree}$ y $\text{ret}^{\wedge} : \text{to_ty}(T')$ con

$T' = \text{FATeleType}(\lambda \ A \Rightarrow \text{imp}(\text{val } A) \ (\text{MFA } A)).$

Capítulo 3

Conclusión

3.1. Conclusión

Capítulo 4

Apéndice

4.1. Apéndice A

4.2. Bibliografía