Typing Total Recursive Functions in Coq

Dominique Larchey-Wendling
TYPES team, ANR TICAMORE

LORIA – CNRS Nancy, France

https://github.com/DmxLarchey/Coq-is-total

Interactive Theorem Proving ITP 2017, Brazilia, Sept. 2017

Turing completeness for (axiom-free) Coq

- Does Coq contain any μ -recursive function as a term $\mathtt{nat}^k \to \mathtt{nat}$?
- Axiom free Coq defines only *total* functions
 - meta-level (strong/weak) normalization
- The Kleene T predicate method (Bove&Capretta 2005):
 - T is cumbersome = small-step semantics
 - primitive recursive schemes hard to program with
- Lambda calculus method (LW 2017, big dev. of 25k lines):
 - left-most and head normalization (so also small-step sem.)
 - intersection type systems (solvability)

How avoid small-step semantics?

The content of the function type $nat \rightarrow nat$

- What is contained in the type $nat \rightarrow nat$?
 - it depends on axioms (even if only of sort Prop)
 - without axioms, only total recursive functions (normalization)
 - but are every total and recursive functions present?
- What are (total) recursive functions?
 - recursive functions are an inductive class of relations
 - but | totality | depends on meta-theory:
 - * Goodstein sequence (Kirby&Paris)
 - * Finite Ramsey theorem (Paris&Harrington)
- Turing completeness for Coq-provably total recursive functions
 - with (short?) Coq-implementation of this claim

Our method: avoid small-step semantics

- Bove&Capretta's hint (Kleene's normal form theorem):
 - μ -recursive fun. = minimization of primitive recursive fun.
 - Kleene's T pred. relates prog. and computations (prim. rec.)
 - primitive rec. fun. are (trivially) Coq-definable terms
 - unbounded minimization of these terms (mutual recursion?)
- Kleene's T predicate = small-step semantics
 - implement as primitive recursive = awfully complicated
 - a provably correct compiler with prim. rec. schemes
- We avoid small-step semantics
 - unbounded minimization of decidable (& inhabited) predicates
 - cost-aware big-step semantics as Coq decidable predicate

Coq-provably total & computable relations

- \bullet To shorten notations, $\mathcal N$ denotes the type nat
- μ -recursive function $\mathbb{N}^k \longrightarrow \mathbb{N} = \text{func. relation } \mathcal{N}^k \to \mathcal{N} \to \text{Prop}$
 - an *inductive class* of functional/deterministic relations
 - constants, successor, zero, projections, composition, primitive recursion and *unbounded minimization*
 - each μ -recursive function is described by an algorithm
 - algorithm must be given, it *cannot* be extracted
- μ -recursive $R: \mathcal{N}^k \to \mathcal{N} \to \mathsf{Prop}$ is total if $\forall \vec{v}: \mathcal{N}^k, \exists n: \mathcal{N}, R \ \vec{v} \ n$
- R is Coq-computable if $\forall \vec{v} : \mathcal{N}^k, \{n : \mathcal{N} \mid R \ \vec{v} \ n\}$
- Transforming $(\exists n, R \ \vec{v} \ n)$ into $\{n \mid R \ \vec{v} \ n\}$ called reification

Specificity of Coq existential quantifiers

- Three type of existential quantifiers (Σ -types)
 - for $P: X \to \text{Prop}$, non-informative $\exists x: X, P \ x \text{ of type Prop}$
 - for $P: X \to \text{Prop}$, partially info. $\{x: X \mid P \mid x\}$ of type Type
 - for $P: X \to \mathsf{Type}$, fully info. $\{x: X \& P x\}$ of type Type
- Reification is a map $(\exists x : X, P \ x) \rightarrow \{x : X \mid P \ x\}$
 - axiom called Constructive Indefinite Description
 - alternatively, it is a map inhabited $X \to X$
- Reification can be implemented without axioms:
 - when X is an enumerable type (like \mathcal{N})
 - when $P: X \to \{Prop, Type\}$ is decidable
 - implementation by unbounded minimization

Inductive definitions of Coq existential quantifiers

```
Inductive inhabited (P: Type): | Prop | :=
     inhabits: P \rightarrow inhabited P
Inductive ex \{X : \mathsf{Type}\}\ (P : X \to \mathsf{Prop}) : |\mathsf{Prop}| :=
     ex_intro: \forall x: X, P \ x \rightarrow ex \ P \ (also denoted \ \exists x: X, P \ x)
Inductive sig \{X: \mathtt{Type}\}\ (P:X \to \mathtt{Prop}): |\, \mathtt{Type}\,| :=
    exist: \forall x: X, P \ x \to \text{sig} \ P \ \text{(also denoted } \{x: X \mid P \ x\})
Inductive sigT \{X : \mathsf{Type}\}\ (P : X \to \mathsf{Type}) : |\mathsf{Type}| :=
    existT: \forall x: X, P \ x \rightarrow sigT \ P \ (also denoted \{x: X \& P \ x\})
          \exists x: X, P \ x \text{ equivalent to inhabited} \{x: X \mid P \ x\}
```

Unbounded minimization (sample OCaml code)

- Minimization of Boolean function $f: int \rightarrow bool$
 - try f 0, f 1, f 2, ... until f n outputs true
 - if e.g. f 0 does not terminate, then minimization loops as well
- Implemented by this sample code:

```
let rec minimize_rec f n= match f n with  \mid \text{ true } \to n   \mid \text{ false } \to \text{ minimize\_rec } f \ (1+n)  let minimize f=\text{minimize\_rec } f \ 0
```

• This codes does not always terminate, but Coq code must...

Terminating unbounded minimization (OCaml)

• How to ensure termination: decorate with a decreasing argument let rec minimize_rec f n H_n = match f n with $| \text{ true } \rightarrow n$ $| \text{ false } \rightarrow \text{ minimize_rec } f \ (1+n) \ \overline{H_{1+n}}$ let minimize f = minimize_rec f 0 $\overline{H_0}$

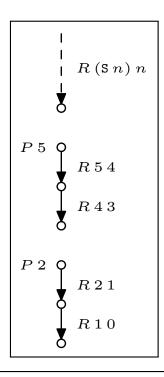
- Problems:
 - Termination input: non-informative proof $H_f: \exists n, f \ n = \texttt{true}$
 - How to obtain H_{1+n} from H_n s.t. H_{1+n} is simpler than H_n ?
 - How to build H_0 from $H_f: \exists n, f \ n =$ true ?
- Solution: H_n is Acc R n for some rel. $R: \mathcal{N} \to \mathcal{N} \to Prop$

Non-informative existence as accessibility

Inductive Acc
$$\{X: \mathtt{Type}\}\ (R: X \to X \to \mathtt{Prop})\ (x: X) :=$$

$$\mid \mathtt{Acc_intro}: (\forall y: X, R\ y\ x \to \mathtt{Acc}\ R\ y) \to \mathtt{Acc}\ R\ x$$

• Accessibility Acc R is the least R-hereditary predicate



- Let e.g. P x iff x = 2 or x = 5
- Let $R \ n \ m$ iff $(n = S \ m) \land \neg P \ m$
- Then Acc R 5 because 5 has no R-antecedent
- Acc R 4 as Acc R 5 (5 only antecedent of 4)
- Acc R 3 as Acc R 4 (4 only antecedent of 3)
- Acc R 2 because 2 has no antecedent
- Then Acc R 1 and Acc R 0
- But $\neg Acc R i \text{ for } i \geq 6$

Well-founded unbounded minimization (1)

Variables $(P: \mathcal{N} \to \mathtt{Prop}) \ ig(H_P: orall n: \mathcal{N}, \ \{P\ n\} + \{\neg P\ n\} ig)$

Let $R(n m : \mathcal{N}) := (n = 1 + m) \land \neg P m$

Let P_Acc_R : $\forall n: \mathcal{N}, P \ n \to Acc \ R \ n$

Let Acc_R_dec : $\forall n: \mathcal{N}, Acc R (1+n) \rightarrow Acc R n$

Let Acc_R_zero : $\forall n: \mathcal{N}, Acc R n \rightarrow Acc R 0$

Let $ex_P_Acc_R_zero$: $(\exists n: \mathcal{N}, P \ n) \to Acc \ R \ 0$

Let Acc_R_eq : $\forall n: \mathcal{N}, Acc R n \iff \exists i: \mathcal{N}, n \leqslant i \land P i$

Well-founded unbounded minimization (2)

```
Let R \ n \ m := (n = 1 + m) \land \neg P \ m
Let Acc_{inv}(n:\mathcal{N}) (H_n:Acc|R|n) (F:\neg P|n):Acc|R|(1+n):=
     let F' := \text{conj eq\_refl } F
     in match H_n with Acc_intro _-H\mapsto H_-F' end
Fixpoint Acc_P (n:\mathcal{N}) (H_n: Acc R n): \{x:\mathcal{N}\mid P x\}:=
     match H_P n with
          | left T \mapsto \mathsf{exist}_{-} n T
          | right F \mapsto Acc_P (1+n) (Acc_inv _ H_n F)
     end.
```

Reification of decidable predicates

• For $P: \mathcal{N} \to \text{Prop and } H_P: \forall n, \{P n\} + \{\neg P n\}$

Theorem nat_reify: $(\exists n : \mathcal{N}, P \ n) \rightarrow \{n : \mathcal{N} \mid P \ n\}$

- Proof:
 - intros $H: \exists n, P \ n$, goal is now $\{n: \mathcal{N} \mid P \ n\}$
 - apply Acc_P with (n := 0), goal is now Acc R : 0 : Prop
 - apply ex_P_Acc_R_zero, goal is now $\exists n : \mathcal{N}, P \ n$
 - assumption, goal solved by hypothesis H
- We also get the fully specified:

Theorem minimize: $(\exists n, P \ n) \rightarrow \{n \mid P \ n \land \forall m, P \ m \rightarrow n \leqslant m\}$

Reification of dec. and informative predicates

• Decidability for informative predicates P: Type

$$\mathtt{decidable_t}\ P := P + (P \to \mathtt{False})$$

• For $P: \mathcal{N} \to \mathsf{Type}$ and $H_P: \forall n, (P \ n) + (P \ n \to \mathsf{False})$

Theorem $nat_reify_t: (\exists n: \mathcal{N}, inhabited(P n)) \rightarrow \{n: \mathcal{N} \& P n\}$

- Hypothesis $\exists n : \mathcal{N}$, inhabited $(P \ n)$ has no informative content
- It computes:
 - -n (minimal) such that P n is inhabited
 - but it also computes an inhabitant of (that) P n
- The proof is very similar to that of nat_reify

An inductive type for recursive algorithms

- X^n is the type of vectors on X: Type of dimension $n:\mathcal{N}$
- \mathcal{A}_k is a notation for recalg $(k:\mathcal{N})$
- recalg : $\mathcal{N} \to \mathtt{Set}$ dependently defined by inductive rules:

$$\frac{f:\mathcal{A}_k \quad \vec{g}:\mathcal{A}_i^k}{\operatorname{comp} f \, \vec{g}:\mathcal{A}_i} \qquad \frac{f:\mathcal{A}_k \quad g:\mathcal{A}_{2+k}}{\operatorname{rec} f \, g:\mathcal{A}_{1+k}} \qquad \frac{f:\mathcal{A}_{1+k}}{\min f:\mathcal{A}_k}$$

• Working with dependent types might involve some difficulties...

Beware fixpoint definitions are not compositional

```
Variable (P : forall k, recalg k -> Type)
    (Pcst: forall n, P (cst n)) (Pzero ....
Fixpoint recalg_rect k f { struct f } : P k f :=
 match f with
   | cst n =  Pcst n
   | zero => Pzero
   | succ => Psucc
   | proj p => Pproj p
   | comp f gj => Pcomp [|f|] (fun p => [|vec_pos gj p|])
   | rec f g => Prec [|f|] [|g|]
   | min f => Pmin [|f|]
 end where "[|f|]" := (recalg_rect_f).
```

Dependencies might involve type castings

- eq_rect maps a term of type P i into P j using a proof e: i = j
- Alternatively, use heterogeneous equality JMeq (John Major's eq.)
- Injection lemmas involve type castings:

Fact ra_comp_inj
$$k \ k' \ i \ (f:\mathcal{A}_k) \ (f':\mathcal{A}_{k'}) \ (\vec{g}:\mathcal{A}_i^k) \ (\vec{g}':\mathcal{A}_i^{k'}) :$$

$$\operatorname{comp} f \ \vec{g} = \operatorname{comp} f' \ \vec{g}' \to \exists e: k = k', \land \begin{cases} \operatorname{eq_rect} \ _-f \ _e = f' \\ \operatorname{eq_rect} \ _-\vec{g} \ _e = \vec{g}' \end{cases}$$

• These difficulties might be frightening for casual Coq users

Relational semantics for recursive algorithms

- We denote $[\![f]\!]$ for ra_rel k $(f:\mathcal{A}_k):\mathcal{N}^k\to\mathcal{N}\to\operatorname{Prop}$
- [f] \vec{v} x: the computation of f on input \vec{v} halts and outputs x

- A simple exercise (given a good recursion principle for A_k ;-)
- But $x \mapsto [\![f]\!] \vec{v} x$ is not a decidable relation.

Big-step semantics for recursive algorithms

- We denote $[f; \vec{v}] \leadsto x$ for ra_bs $k \ f \ \vec{v} \ x$: Prop (or Type...)
- Same meaning as [f] \vec{v} x but defined as an inductive predicate

- Easy (intuitive?) definition
 - $\text{ ra_bs} : \forall k, \mathcal{A}_k \to \mathcal{N}^k \to \mathcal{N} \to \text{Prop}, \llbracket f \rrbracket \ \vec{v} \ x \Longleftrightarrow [f; \vec{v}] \leadsto x$
 - ra_bs: $\forall k, \mathcal{A}_k \to \mathcal{N}^k \to \mathcal{N} \to \mathsf{Type}$ is a type of computations
- Let us transform ra_bs into a decidable predicate

Cost aware big-step semantics

- ullet We denote $[f; \vec{v}]$ $-[lpha]\!\!\rangle x$ for ra_ca k f \vec{v} lpha x: Prop
- α represents the cost (or size) of the computation

- for ra_ca, we have $\llbracket f \rrbracket \ \vec{v} \ x \Longleftrightarrow \exists \alpha : \mathcal{N}, \ [f; \vec{v}] \ \neg [\alpha] \rangle x$
- $x \mapsto [f; \vec{v}] [\alpha] \times x$ is a decidable predicate:
 - from α , recover comp. $[f; \vec{v}] \rightsquigarrow x$: Type by prim. rec. means

Properties of cost aware semantics

• Inversion lemmas:

Lemma ra_ca_rec_S_inv $(k:\mathcal{N})$ $(f:\mathcal{A}_k)$ $(g:\mathcal{A}_{2+k})$ \vec{v} n γ x:

$$[\operatorname{rec} f \ g; 1 + n \# \vec{v}] - [\gamma \rangle \rangle \ x \to \exists y \ \alpha \ \beta, \wedge \begin{cases} \gamma = 1 + \alpha + \beta \\ [\operatorname{rec} f \ g; n \# \vec{v}] - [\alpha \rangle \rangle \ y \\ [g; n \# y \# \vec{v}] - [\beta \rangle \rangle \ x \end{cases}$$

• Functionality:

Theorem ra_ca_fun
$$(k:\mathcal{N})$$
 $(f:\mathcal{A}_k)$ $(\vec{v}:\mathcal{N}^k)$ $(\alpha \beta x y : \mathcal{N})$:
$$[f;\vec{v}] - |\alpha\rangle\rangle x \rightarrow [f;\vec{v}] - |\beta\rangle\rangle y \rightarrow \alpha = \beta \wedge x = y$$

• Decidability:

Theorem ra_ca_decidable_t $(k:\mathcal{N})$ $(f:\mathcal{A}_k)$ $(\vec{v}:\mathcal{N}^k)$ $(\alpha:\mathcal{N})$: $\{x \mid [f;\vec{v}] - |\alpha\rangle\rangle x\} + \{x \mid [f;\vec{v}] - |\alpha\rangle\rangle x\} \rightarrow \mathtt{False}$

Typing total recursive functions

- For $f: \mathcal{A}_k$ and $\vec{v}: \mathcal{N}^k$ fixed, f terminates on \vec{v} iff:
 - $-\exists x, \llbracket f \rrbracket \ \vec{v} \ x$
 - $-\exists x\exists \alpha, [f; \vec{v}] |\alpha\rangle\rangle x$
 - $-\exists \alpha \exists x, [f; \vec{v}] |\alpha\rangle\rangle x$
 - $-\exists \alpha, \mathtt{inhabited} \{x \mid [f; \vec{v}] \vdash [\alpha] \rangle x\}$
- For any α , the type $\{x \mid [f; \vec{v}] [\alpha] \mid x\}$ is decidable:
 - nat_reify_t computes $\{\alpha: \mathcal{N} \& \{x: \mathcal{N} \mid [f; \vec{v}] [\alpha] \mid x\}\}$
 - from which we extract x s.t. $\llbracket f \rrbracket \ \vec{v} \ x$

Theorem Coq_is_total $(k:\mathcal{N})$ $(f:\mathcal{A}_k)$:

 $(\forall \vec{v}: \mathcal{N}^k, \exists x: \mathcal{N}, \llbracket f \rrbracket \ \vec{v} \ x) \to \{t: \mathcal{N}^k \to \mathcal{N} \mid \forall \vec{v}: \mathcal{N}^k, \llbracket f \rrbracket \ \vec{v} \ (t \ \vec{v})\}$

Other applications: reifying undecidable predicates

- Normal forms (typically λ -calculus)
 - for T: Type, $R: T \to T \to \texttt{Prop}$
 - finitary: $\forall t: T, \{l: \mathtt{list}\ X \mid \forall x, R\ t\ x \Longleftrightarrow \mathtt{In}\ x\ l\}$
 - with normal_form $t n := (\forall x, \neg R \ n \ x) \land R^{\star} \ t \ n$
 - we have: $\forall t, (\exists n, \mathtt{normal_form}\ t\ n) \rightarrow \{n \mid \mathtt{normal_form}\ t\ n\}$
- From cut-admissibility to cut-elimination

 $\forall s \ (p : \mathtt{proof} \ s), (\exists q : \mathtt{proof} \ s, \mathtt{cut_free} \ q) \rightarrow \{q : \mathtt{proof} \ s \mid \mathtt{cut_free} \ q\}$

• Recursively enumerable predicates (of the form $\vec{v} \mapsto [\![f]\!] \vec{v}$ 0)

$$\forall (k:\mathcal{N}) \ (f:\mathcal{A}_k), \ (\exists \vec{v}:\mathcal{N}^k, \llbracket f \rrbracket \ \vec{v} \ 0) \to \{\vec{v}:\mathcal{N}^k \mid \llbracket f \rrbracket \ \vec{v} \ 0\}$$

Conclusion

- Mechanization of the Turing completeness of Coq
 - without using any (extra) axiom
 - by implementing reification of decidable predicates over **nat**
- Coq has a kind of unbounded minimization
 - provided the predicate can be informatively decided
 - and there is a non-informative inhabitation proof
- Kleene's T predicate replaced with cost aware big-step semantics
 - avoid small-step semantics and encodings
 - avoid compiler correctedness
 - show decidability of cost aware big-step semantics
- Reification extended to some undecidable predicates as well