When Bécassine brings automation to Coq

Valentin Blot¹² Louise Dubois de Prisque¹², Chantal Keller², <u>Pierre Vial¹²</u>

 $^{1}\mathrm{Deducteam}$ (Inria Paris-Saclay) $^{2}\mathrm{LMF}$ (Gif-sur-Yvette)

This work is funded by Nomadic Labs

November 4, 2021

Incha

A FEW WORDS ABOUT $\ensuremath{\texttt{Coq}}$

• COQ: proof assistant based on type theory

and the **Curry-Howard isomorphism** Formulas = types, proofs = programs

• Four Colors Theorem

- Feit-Thomson Theorem
- CompCert (certified compiler)
- These successes are possible because of its design
 - Strong type-checking within Coq
 - Rich specification language
 - Highly trusted (small logical kernel)

Goal forall (A : Type) (l : list A) (n : nat), length $l = S n \rightarrow l \neq []$. Proof. intros A l n H H'. rewrite H' in H. discriminate. Qed.

Goal forall (A : Type) (l : list A) (n : nat), length $l = S n \rightarrow l \neq []$. Proof. intros A l n H H'. rewrite H' in H. discriminate. Qed.

Let us admit

 $\label{eq:lemma_search_app:forall (A: Type) (x: A) (11 12: list A), \\ search x (11 ++ 12) = (search x 11) || (search x 12). \\$

Goal forall (A : Type) (l : list A) (n : nat), length $l = S n \rightarrow l \neq []$. Proof. intros A l n H H'. rewrite H' in H. discriminate. Qed.

Let us admit

 $\label{eq:lemma_search_app:forall(A:Type)(x: A) (l1 l2: list A),} \\ search x (l1 ++ l2) = (search x l1) || (search x l2).$

Goal forall (A : Type) (l : list A) (n : nat), length $l = S n \rightarrow l \neq []$. Proof. intros A l n H H'. rewrite H' in H. discriminate. Qed.

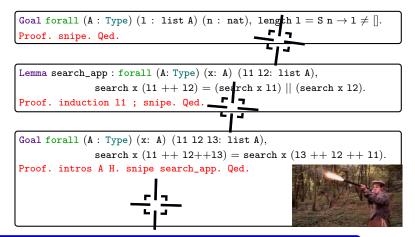
Let us admit

 $\label{eq:lemma_search_app:forall (A: Type) (x: A) (11 12: list A), \\ search x (11 ++ 12) = (search x 11) || (search x 12). \\$

Coq lacks automation

- The user must be very specific
- Difficult for the beginner/non-formal method specialist
- May discourage new users (e.g., maths, industry)





Coq lacks automation

- The user must be very specific
- Difficult for the beginner/non-formal method specialist
- May discourage new users (e.g., maths, industry)

Coq (Proof assistant)	First-order provers
Very expressive logic	Limited expressivity
Checks proofs	Finds proofs
Highly trustable	Less so

- Coq difficult to automatize
- Even the first-order part of the proofs
- FOL highly automated outside Coq
- Line of software development: call external solvers to handle the first-order parts of the proofs (avoid redundant code!)
- Partial transformations from Coq logic to FOL



Coq *vs.* automated provers

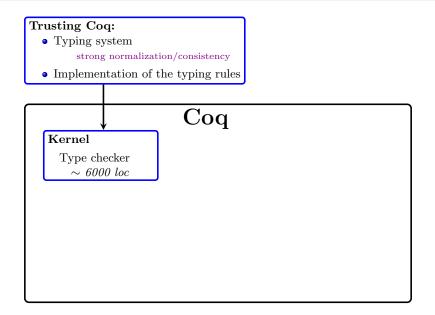


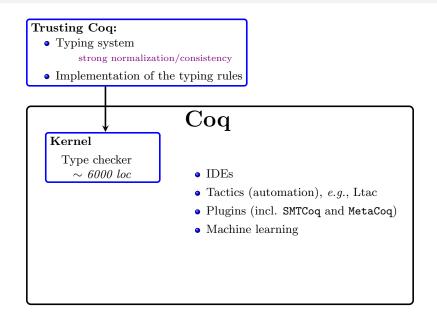
Trusting Coq:

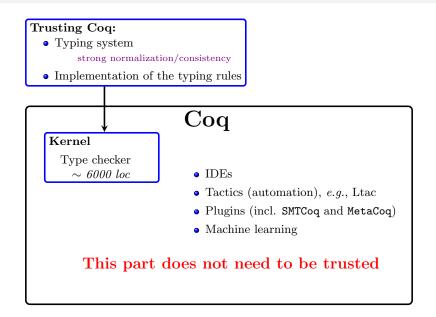
• Typing system

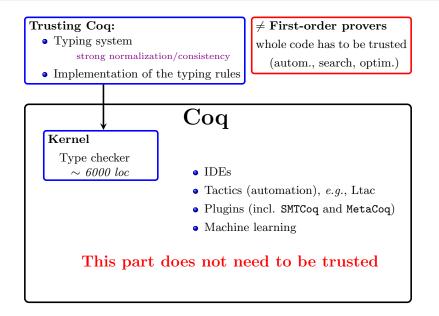
strong normalization/consistency

• Implementation of the typing rules









COQ LOGIC vs. FIRST-ORDER LOGIC

Coq

based on the Calculus of Inductive Constructions (CIC)

First-order logic (FOL)

- functions and relations
- basic datatypes (bool, int, float)
- boolean equality
- quantification over objects

```
incl. linear integer arithmetics, etc
```

In CIC but not in FOL:

• Higher-order computation (functions are first-class objects):

map f [x1 ; ... ; xn] := [f x1 ; ... ; f xn] ~~

map f is a function on lists

• Higher-order quantification

forall (A B C : Type) (f : A \rightarrow B) (g : B \rightarrow C), (map g) o (map f) = map (g o f)

• Dependent types, e.g., Vec A n is definable

the type of lists of length ${\tt n}$ whose elements have type ${\tt A}$

Sniper: Bécassine à la rescousse

COQ LOGIC vs. FIRST-ORDER LOGIC

Coq

based on the Calculus of Inductive Constructions (CIC)

First-order logic (FOL)

- functions and relations
- basic datatypes (bool, int, float)
- boolean equality
- quantification over objects

incl. linear integer arithmetics, etc

COQ LOGIC vs. FIRST-ORDER LOGIC

Coq

based on the Calculus of Inductive Constructions (CIC)

First-order logic (FOL)

- functions and relations
- basic datatypes (bool, int, float)
- boolean equality
- quantification over objects

```
incl. linear integer arithmetics, etc
```

Zoom on Coq inductives

```
• Inductive types
```

Inductive list (A : Type) : Type := [] : list A | _ :: _ : \rightarrow list A \rightarrow list A

Fixpoints and pattern-matching:
Fixpoint length { A : Type } (1: list A) := match l with

[] ⇒ 0 | a :: 10 ⇒ 1 + length 10

• Generic (non-boolean) Leibniz equality on any type Leibniz equality is a dependent type

• When we make two programs interact, we need an interface

```
Theorem destruct_list : forall 1 : list A, {x:A & {tl:list A | 1 = x::tl}}+{1 = nil}.
Proof.
induction 1 as [|a tl].
right; reflexivity.
left; exists a; exists tl; reflexivity.
Qed.
```

A Coq Theorem and its proof

```
1:(input (#1:(= op_3 #2:(op_1 op_4 op_5))))
2:(input (#3:(forall ( RelName10 Tindex_1) (RelName11 Tindex_2) ) #4:(=> #5:(= op_3 #6:(op_1 RelName11 R
3:(tmp_betared (#7:(forall ( (@vr10 Tindex_1) (@vr11 Tindex_2) ) #8:(=> #9:(= op_3 #10:(op_1 @vr10 @vr10))
4:(tmp_qnt_tidy (#11:(forall ( (@vr14 Tindex_1) (@vr16 Tindex_2) ) #12:(=> #13:(= op_3 #14:(op_1 @vr16 @v
5:(forall_inst (#15:(or (not #11) #16:(=> #1 false))))
6:(false ((not false)))
7:(implies_pos ((not #16) (not #1) false))
```

Excerpt of an smt2 certificate

• When we make two programs interact, we need an interface

```
Theorem destruct_list : forall 1 : list A, {x:A & {tl:list A | 1 = x::tl}}+{1 = nil}.
Proof.
induction 1 as [|a tl].
right; reflexivity.
left; exists a; exists tl; reflexivity.
Qed.
```

A Coq Theorem and its proof

• When we make two programs interact, we need an interface

```
Theorem destruct_list : forall 1 : list A, {x:A & {tl:list A | 1 = x::tl}}+{1 = nil}.
Proof.
induction 1 as [|a tl].
right; reflexivity.
left; exists a; exists tl; reflexivity.
Qed.
```

A Coq Theorem and its proof

Need for **reification** (or quoting)

 \rightsquigarrow translating programs of a language \mathcal{L} into another language \mathcal{L}' .

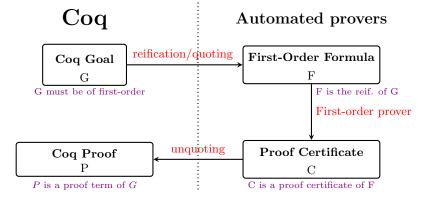
• When we make two programs interact, we need an interface

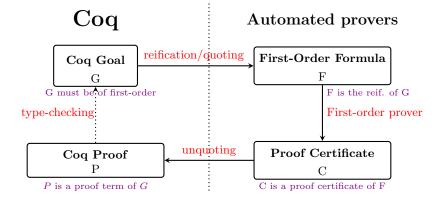
```
Theorem destruct_list : forall 1 : list A, {x:A & {tl:list A | 1 = x::tl}}+{1 = nil}.
Proof.
induction 1 as [|a tl].
right; reflexivity.
left; exists a; exists tl; reflexivity.
Qed.
```

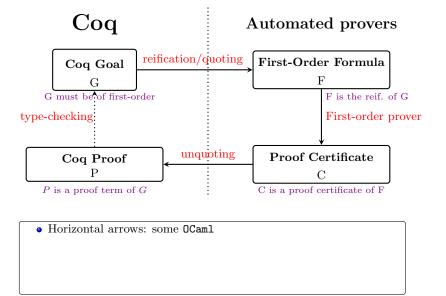
A Coq Theorem and its proof

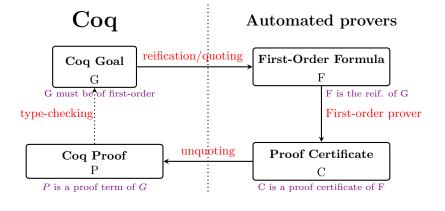
Need for **reification** (or quoting) \rightsquigarrow translating programs of a language \mathcal{L} into another language \mathcal{L}' .

e.g., forall (A : Set), $A \rightarrow A$ (type) \rightarrow Prod (name "A") Set_reif (Prod unnamed A (dB 0) (dB 1)) =reif. with de Bruijn indexes

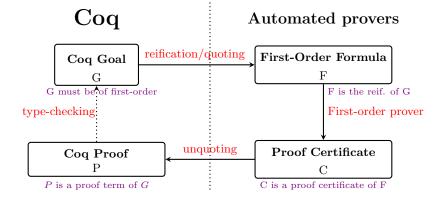




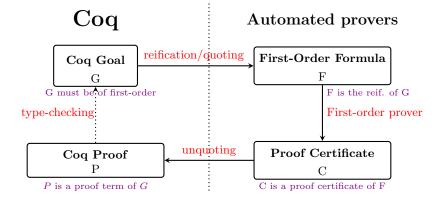


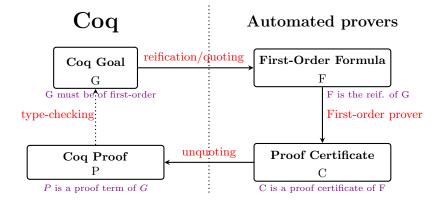


- Horizontal arrows: some OCaml
- Any arrow may fail (reification, solving...)



- Horizontal arrows: some OCaml
- Any arrow may fail (reification, solving...)
- Autarkic approach: each certificate is checked on the run





In our case:

- Plugin = SMTCoq
- Automated provers = SMT solvers, *e.g.*, veriT
- Under the carpet: casting Leibniz equality into boolean eq.

(decidable types only)



• Question. Why aren't we happy with this?



• Question. Why aren't we happy with this?

Problem 1 Avoid harmless polymorphism and higher-order - forall (A : Type) (11 12 : list A), length (11 ++ 12) = length 11 + length 12 - f = g with f,g: nat -> nat

instead of \forall (x : nat), f x = g x



• Question. Why aren't we happy with this?

Problem 1

Avoid harmless polymorphism and higher-order

```
- forall (A : Type) (l1 l2 : list A),
length (l1 ++ l2) = length l1 + length l2
- f = g with f,g: nat -> nat
instead of ∀(x : nat), f x = g x
```

Problem 2

Some info. is lost during goal reification

type constructors uninterpreted e.g., - S $n = S n' \rightarrow n = n'$ is forgotten

- nothing known about List.length



• Question. Why aren't we happy with this?

Problem 1

Avoid harmless polymorphism and higher-order

```
- forall (A : Type) (l1 l2 : list A),
    length (l1 ++ l2) = length l1 + length l2
- f = g with f,g: nat -> nat
```

instead of \forall (x : nat), f x = g x

Problem 2

Some info. is lost during goal reification

type constructors uninterpreted e.g., - S $n = S n' \rightarrow n = n'$ is forgotten

- nothing known about List.length

→ Sniper

- eliminates harmless polymorphism
- helps the first-order provers interpret symbols (constructors and functions)



• Question. Why aren't we happy with this?

Problem 1

Avoid harmless polymorphism and higher-order

```
- forall (A : Type) (l1 l2 : list A),
length (l1 ++ l2) = length l1 + length l2
```

 $-f = g \text{ with } f,g: nat \rightarrow nat$

instead of \forall (x : nat), f x = g x



\rightsquigarrow Sniper

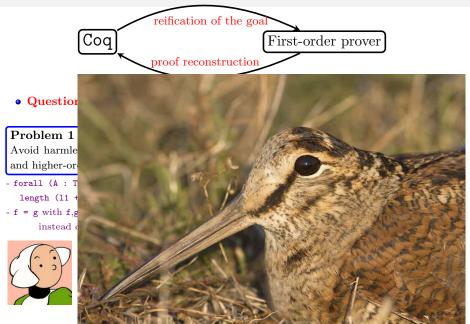
- eliminates harmless polymorphism
- helps the first-order provers interpret symbols (constructors and functions)

Problem 2

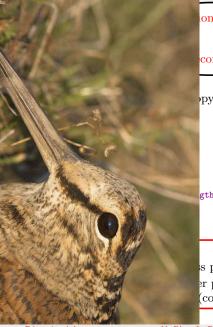
Some info. is lost during goal reification

type constructors uninterpreted e.g., - S $n = S n' \rightarrow n = n'$ is forgotten

- nothing known about List.length



V. Blot, L. Dubois de Prisque, C. Keller, <u>P. Vial</u>





py with this?

Problem 2

Some info. is lost during goal reification

gth 12

type constructors uninterpreted e.g., - S $n = S n' \rightarrow n = n'$ is forgotten

- nothing known about List.length

s polymorphism er provers interpret symbols (constructors and functions)

Sniper: Bécassine à la rescousse



• Question. Why aren't we happy with this?

Problem 1

Avoid harmless polymorphism and higher-order

```
- forall (A : Type) (l1 l2 : list A),
length (l1 ++ l2) = length l1 + length l2
```

 $-f = g \text{ with } f,g: nat \rightarrow nat$

instead of \forall (x : nat), f x = g x



\rightsquigarrow Sniper

- eliminates harmless polymorphism
- helps the first-order provers interpret symbols (constructors and functions)

Problem 2

Some info. is lost during goal reification

type constructors uninterpreted e.g., - S $n = S n' \rightarrow n = n'$ is forgotten

- nothing known about List.length

Coq *vs.* automated provers



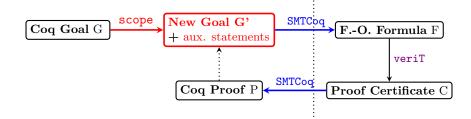
SNIPER (PRINCIPLES)

Sniper is a two-fold tactic.

First Step. The tactic scope

- Eliminates harmless higher-order and polymorphism in the goal if needed
- Produces and proves first-order auxiliary statements in the local context (currently 6 transformations,)

Second step. The transformed goal and the auxiliary statement are sent to the SMT solver veriT, *via* SMTCoq.



Question. What should a first-order prover know about an inductive datatype T?

```
Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)
```

Question. What should a first-order prover know about an inductive datatype T?

```
Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)
```

Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1

Question. What should a first-order prover know about an inductive datatype T?

```
Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)
```

- Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1
- Constructors are injective (I_T)

 $\forall x y l l', x::l = y::l' \rightarrow x=y \land l=l'$

Question. What should a first-order prover know about an inductive datatype T?

Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)

- Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1
- Constructors are injective (I_T)

 $\forall x y l l', x::l = y::l' \rightarrow x=y \land l=l'$

Every term of this type is generated by one of the constructors (G_T)
 ∀(1:list A),∃x l', l = x :: l' ∨ l = []

Question. What should a first-order prover know about an inductive datatype T?

Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)

- Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1
- Constructors are injective (I_T)

 $\forall x y l l', x::l = y::l' \rightarrow x=y \land l=l'$

Every term of this type is generated by one of the constructors (G_T)
 ∀(1:list A),∃x 1', 1 = x :: 1' ∨ 1 = []

Problem.

 $T \mapsto D_T, I_T, G_T$ cannot be defined in Coq or in Ltac

Question. What should a first-order prover know about an inductive datatype T?

Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)

- Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1
- Constructors are injective (I_T)

 $\forall x y l l', x::l = y::l' \rightarrow x=y \land l=l'$

Every term of this type is generated by one of the constructors (G_T)
 ∀(1:list A),∃x 1', 1 = x :: 1' ∨ 1 = []

Problem.

 $T \mapsto D_T, I_T, G_T$ cannot be defined in Coq or in Ltac

Solution. Gain direct access to the syntax of Coq terms

Question. What should a first-order prover know about an inductive datatype T?

Inductive list (A : Type) : Type :=
| nil : list A (* [] *)
| cons : A → list A → list A. (* _ :: _ *)

- Constructors are pairwise disjoints (D_T)
 ∀x 1, [] ≠ x::1
- Constructors are injective (I_T)

 $\forall x y l l', x::l = y::l' \rightarrow x=y \land l=l'$

Every term of this type is generated by one of the constructors (G_T)
 ∀(1:list A),∃x 1', 1 = x :: 1' ∨ 1 = []

Problem.

 $T \mapsto D_T, I_T, G_T$ cannot be defined in Coq or in Ltac

Solution. Gain direct access to the syntax of Coq terms

→ Use MetaCoq reification of Coq in Coq

+ quoting/unquoting mechanisms

How does the transformations work?

How does the transformations work?

- Generation of the reified statements in MetaCoq (e.g., constructors are injective)
- **2** Unreify these statements
- **③** Proof of these statements with Coq regular tactics (Ltac)
- In the statements are now in the local context

How does the transformations work?

How does the transformations work?

- Generation of the reified statements in MetaCoq (e.g., constructors are injective)
- **2** Unreify these statements
- **③** Proof of these statements with Coq regular tactics (Ltac)
- **9** The statements are now in the local context

Currently implemented transformations

- Make explicit the semantics of symbols
- Eliminate higher-order equalities
- Eliminate prenex polymorphism

Action of scope on a goal



A : Type

forall (1 : list A) (n : nat), length A $l = S n \rightarrow l \neq []$

Sniper: Bécassine à la rescousse

Action of scope on a goal

inductive datatypes

A: Type
1: forall B (x y : B) (1 1' : list B), x ::
$$1 = y :: 1' \rightarrow x = y \land 1 = 1'$$

1: forall B (x : B) (1 : list B), [] \neq x :: 1
1: forall (n n': nat), S n = S n' \rightarrow n = n'
1: forall (n : nat), $0 \neq$ S n

forall (1 : list A) (n : nat), length A l = S n \rightarrow l \neq []

Action of scope on a goal

inductive datatypesdefinitions

A: Type 1: forall B (x y : B) (1 1' : list B), x :: $1 = y :: 1' \rightarrow x = y \land 1 = 1'$ 1: forall B (x : B) (1 : list B), [] \neq x :: 1 1: forall (n n': nat), S n = S n' \rightarrow n = n' 1: forall (n : nat), $0 \neq$ S n 2: length = (fun B \Rightarrow fix length 1 := match 1 with ... end)

 $\texttt{forall} (\texttt{l}: \texttt{list A}) (\texttt{n}: \texttt{nat}), \texttt{ length A } \texttt{l} = \texttt{S} \texttt{ n} \rightarrow \texttt{l} \neq []$

Action of ${\tt scope}$ on a goal

- inductive datatypesdefinitions
- expansion

A: Type 1: forall B (x y : B) (l l' : list B), x :: $l = y :: l' \rightarrow x = y \land l = l'$ 1: forall B (x : B) (l : list B), [] \neq x :: l 1: forall (n n': nat), S n = S n' \rightarrow n = n' 1: forall (n : nat), $0 \neq$ S n 2: length = (fun B \Rightarrow fix length l := match l with ... end) 3: forall B l, length B l = (fun B \Rightarrow fix length l := match l with ... end) B l

forall (1 : list A) (n : nat), length A $l = S n \rightarrow l \neq []$

Sniper: Bécassine à la rescousse

Action of ${\tt scope}$ on a goal

- inductive datatypes
- 2 definitions
- expansion
- fixpoints

A : Type

- 1: forall B (x y : B) (l l' : list B), x :: l = y :: l' \rightarrow x = y \wedge l = l'
- 1: forall B (x : B) (1 : list B), [] \neq x :: 1
- 1: forall (n n': nat), S n = S n' \rightarrow n = n'
- 1: forall (n : nat), $0 \neq S$ n
- 2: length = (fun $B \Rightarrow$ fix length l := match l with ... end)
- 3: forall B 1, length $B 1 = (fun B \Rightarrow fix length 1 := match 1 with ... end) B 1$
- 4: forall B 1, length B 1 = match 1 with ... end

forall (1 : list A) (n : nat), length A $l = S n \rightarrow l \neq []$

Action of scope on a goal

- inductive datatypes
- 2 definitions
- expansion
- fixpoints
- **6** elimination of pattern matching

A : Type

1: forall B (x y : B) (l l' : list B), x :: l = y :: l'
$$\rightarrow$$
 x = y \land l = l'

1: forall B (x : B) (1 : list B), [] \neq x :: 1

1: forall (n n': nat), S n = S n'
$$\rightarrow$$
 n = n²

1: forall (n : nat),
$$0 \neq S$$
 m

2: length = (fun B
$$\Rightarrow$$
 fix length l := match l with ... end)

3: forall B 1, length B 1 = (fun B
$$\Rightarrow$$
 fix length 1 := match 1 with ... end) B 1

5: forall B, length B
$$[] = 0$$

5: forall B (1 : list B) (x : B), length B x ::
$$1 = S$$
 (length B 1)

 $\texttt{forall} (\texttt{l}: \texttt{list A}) (\texttt{n}: \texttt{nat}), \texttt{ length A } \texttt{l} = \texttt{S} \texttt{ n} \rightarrow \texttt{l} \neq []$

Action of scope on a goal

- inductive datatypes
- 2 definitions
- expansion
- fixpoints
- elimination of pattern matching
- applied polymorphic hypotheses

A : Type

```
1: forall B (x y : B) (l l' : list B), x :: l = y :: l' \to x = y \land l = l'
  1: forall B (x : B) (1 : list B), [] \neq x :: 1
  1: forall (n n': nat), S n = S n' \rightarrow n = n'
  1: forall (n : nat), 0 \neq S n
  2: length = (fun B \Rightarrow fix length l := match l with ... end)
  3: forall B 1, length B 1 = (fun B \Rightarrow fix length 1 := match 1 with ... end) B 1
  4: forall B 1, length B 1 = match 1 with ... end
  5: forall B, length B [] = 0
  5: forall B (1 : list B) (x : B), length B x :: 1 = S (length B 1)
  6: length A []=0
  6: forall (1 : list A) (x : A), length x :: 1 = S (length A 1)
     forall (1 : list A) (n : nat), length A l = S n \rightarrow l \neq []
Sniper: Bécassine à la rescousse
                                   V. Blot, L. Dubois de Prisque, C. Keller, P. Vial
```

Action of scope on a goal

- inductive datatypes
- 2 definitions
- expansion
- fixpoints
- elimination of pattern matching
- applied polymorphic hypotheses

A : Type

1: forall B (x y : B) (1 1' : list B), x ::
$$1 = y :: 1' \rightarrow x = y \land 1 = 1'$$

1: forall B (x : B) (1 : list B), [] \neq x :: 1
1: forall (n n': nat), S n = S n' \rightarrow n = n'
1: forall (n : nat), $0 \neq$ S n
2: length = (fun B \Rightarrow fix length 1 := match 1 with ... end)
3: forall B 1, length B 1 = (fun B \Rightarrow fix length 1 := match 1 with ... end) B 1
4: forall B 1, length B 1 = match 1 with ... end
5: forall B, length B [] = 0
5: forall B (1 : list B) (x : B), length B x :: 1 = S (length B 1)
6: length A []= 0
6: forall (1 : list A) (x : A), length x :: 1 = S (length A 1)
forall (1 : list A) (n : nat), length A 1 = S n \rightarrow 1 \neq []

V. Blot, L. Dubois de Prisque, C. Keller, P. Vial

CONCLUSION AND FUTURE WORK

- General methodology: small transformations from a subset of Coq logic to FOL
- Proof of concept: six transformations combined in a tactic (snipe = scope + verit) which calls an external SMT solver.

These transformations are independent from SMTCoq!

CONCLUSION AND FUTURE WORK

- General methodology: small transformations from a subset of Coq logic to FOL
- Proof of concept: six transformations combined in a tactic (snipe = scope +verit) which calls an external SMT solver.

These transformations are independent from SMTCoq!

In the Future.

- More complex transformations: (simple) **dependent types**, dependent pattern matching...
- Add user-defined tactics
- Benchmarks

+ improving the performance of our tactic

Try Sniper! https://github.com/smtcoq/sniper

CONCLUSION AND FUTURE WORK

- \bullet General methodology: small transformations from a subset of ${\tt Coq}$ logic to FOL
- Proof of concept: six transformations combined in a tactic (snipe = scope + verit) which calls an external SMT solver.

These transformations are independent from SMTCoq!

In the Future.

- More complex transformations: (simple) **dependent types**, dependent pattern matching...
- Add user-defined tactics
- Benchmarks

+ improving the **performance** of our tactic

Try Sniper!

https://github.com/smtcoq/sniper

