

¹INRIA, ²Ecole Polytechnique, ³ENSIIE/Cedric

Embedding logics in Dedukti

Ali Assaf^{1,2}, Guillaume Burel³

April 12, 2013

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

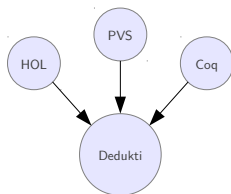
Universal proof checker

Source: HOL, Coq, ...

- ▶ Pure type systems
- ▶ Reconstruction, proof search, ...

Target: Dedukti

- ▶ $\lambda\Pi$ -calculus modulo (see previous talk)
- ▶ Proof checking (no reconstruction, no proof search, ...)



Towards interoperability

Prove $A \Rightarrow B$ in system 1. Prove A in system 2.

- ▶ Can we deduce B ?

$$\frac{\|A \Rightarrow B\|_1 \quad \|A\|_2}{\|B\|_?} ?$$

Shallow embeddings: $\|A \Rightarrow B\|_1 = \|A\|_1 \rightarrow \|B\|_1$

- ▶ Only need to relate $\|A\|_1$ with $\|A\|_2$

Shallow vs. deep embeddings

Reuse features of the target language

- ▶ Variables, binders, ...
- ▶ Rewriting, computation, ...

Advantages:

- ▶ Lighter translations
- ▶ Make interoperability easier

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

Pure type systems

Signature $P = (\mathcal{S}, \mathcal{A}, \mathcal{R})$

- ▶ \mathcal{S} a set of *sorts*
- ▶ $\mathcal{A} \subseteq \mathcal{S} \times \mathcal{S}$ a set of *axioms*
- ▶ $\mathcal{R} \subseteq \mathcal{S} \times \mathcal{S} \times \mathcal{S}$ a set of *rules*

Syntax

sorts	s	\in	\mathcal{S}
terms	t, u, A, B	$::=$	$x \mid s \mid \lambda x:A. t \mid \Pi x:A. B \mid t u$
contexts	Γ	$::=$	$\emptyset \mid \Gamma, x : A$

Typing rules

$$\frac{\Gamma \text{ well-formed} \quad (x : A) \in \Gamma}{\Gamma \vdash x : A} \qquad \frac{\Gamma \text{ well-formed} \quad (s_1 : s_2) \in \mathcal{A}}{\Gamma \vdash s_1 : s_2}$$

$$\frac{\Gamma \vdash A : s_1 \quad \Gamma, x : A \vdash t : B \quad \Gamma \vdash (\Pi x : A. B) : s_3}{\Gamma \vdash (\lambda x : A. t) : (\Pi x : A. B)}$$

$$\frac{\Gamma \vdash A : s_1 \quad \Gamma, x : A \vdash B : s_2 \quad (s_1, s_2, s_3) \in \mathcal{R}}{\Gamma \vdash (\Pi x : A. B) : s_3}$$

$$\frac{\Gamma \vdash t : (\Pi x : A. B) \quad \Gamma \vdash u : A}{\Gamma \vdash (tu) : [u/x]B} \qquad \frac{\Gamma \vdash t : A \quad \Gamma \vdash B : s \quad A \equiv_{\beta} B}{\Gamma \vdash t : B}$$

$$\frac{}{\emptyset \text{ well-formed}} \qquad \frac{\Gamma \text{ well-formed} \quad \Gamma \vdash A : s \quad x \notin \Gamma}{\Gamma, x : A \text{ well-formed}}$$

Example

Example

The calculus of constructions (CC) is the PTS defined by the signature:

$$\begin{aligned} \mathcal{S} &= \text{type, kind} \\ \mathcal{A} &= (\text{type} : \text{kind}) \\ \mathcal{R} &= (\text{type, type, type}), (\text{kind, type, type}), \\ &\quad (\text{type, kind, kind}), (\text{kind, kind, kind}) \end{aligned}$$

The polymorphic identity function $id = (\lambda A:\text{type}. \lambda x:A. x)$ is well-typed in CC:

$$\vdash id : (\Pi A:\text{type}. A \rightarrow A)$$

Translations

- ▶ $\lambda\Pi$ -calculus: types may only depend on terms
- ▶ PTS: terms and types may depend on types (polymorphism, type operators)

Two translations:

- ▶ $|t|$ as a **term**
- ▶ $\|t\|$ as a **type**

Two Translations

As a **type**:

- ▶ $\|s\| = s$, the universe of terms of type s
- ▶ $\|\Pi x:A. B\| = \Pi x:\|A\|. \|B\|$

As a **term**:

- ▶ $|s| = \dot{s}$, a constant
- ▶ $|\Pi x:A. B| = \dot{\pi} |A| (\lambda x:\|A\|. |B|)$

Decoding function

- ▶ A term A can be used as a type
- ▶ Decoding function ε

$$\|A\| = \varepsilon |A|$$

- ▶ Constraints:

$$\begin{array}{lcl} \|s\| & = & s \\ \|\Pi x : A. B\| & = & \Pi x : \|A\|. \|B\| \end{array}$$

Decoding function

- ▶ A term A can be used as a type
- ▶ Decoding function ε

$$\|A\| = \varepsilon |A|$$

- ▶ Constraints:

$$\begin{array}{lcl} \varepsilon |s| & = & s \\ \varepsilon |\Pi x:A. B| & = & \Pi x:\varepsilon |A|. \varepsilon |B| \end{array}$$

Decoding function

- ▶ A term A can be used as a type
- ▶ Decoding function ε

$$\|A\| = \varepsilon |A|$$

- ▶ Constraints:

$$\varepsilon (\dot{\pi} |A| (\lambda x : \|A\| . |B|)) = \Pi x : \varepsilon |A| . \varepsilon |B|$$

Decoding function

- ▶ A term A can be used as a type
- ▶ Decoding function ε

$$\|A\| = \varepsilon |A|$$

- ▶ Constraints:

$$\begin{array}{lcl} \varepsilon \dot{s} & = & s \\ \varepsilon (\dot{\pi} A B) & = & \Pi x : \varepsilon A. \varepsilon (B x) \end{array}$$

Decoding function

- ▶ A term A can be used as a type
- ▶ Decoding function ε

$$\|A\| = \varepsilon |A|$$

- ▶ Constraints:

$$\begin{array}{ccc} \varepsilon \dot{s} & \longrightarrow & s \\ \varepsilon (\dot{\pi} A B) & \longrightarrow & \Pi x : \varepsilon A. \varepsilon (B x) \end{array}$$

The embedding

Constants

s	: Type	$\forall s \in \mathcal{S}$
ε_s	: $s \rightarrow \text{Type}$	$\forall s \in \mathcal{S}$
\dot{s}_1	: s_2	$\forall (s_1 : s_2) \in \mathcal{A}$
$\dot{\pi}_{s_1, s_2, s_3}$: $\Pi A : s_1. (\varepsilon_{s_1} A \rightarrow s_2) \rightarrow s_3$	$\forall (s_1, s_2, s_3) \in \mathcal{R}$

Rewrite rules

$\varepsilon_{s_2}(\dot{s}_1)$	\longrightarrow	s_1	$\forall (s_1 : s_2) \in \mathcal{A}$
$\varepsilon_{s_3}(\dot{\pi}_{s_1, s_2, s_3} A B)$	\longrightarrow	$\Pi x : \varepsilon_{s_1} A. \varepsilon_{s_2}(B x)$	$\forall (s_1, s_2, s_3) \in \mathcal{R}$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$|\lambda A:\text{type}. \lambda x:A. x| \quad : \quad \|\Pi A:\text{type}. A \rightarrow A\|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$|\lambda A:\text{type}. \lambda x:A. x| \quad : \quad \|\Pi A:\text{type}. A \rightarrow A\|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\|\text{type}\|. \lambda x:\|A\|. x \quad : \quad \|\Pi A:\text{type}. A \rightarrow A\|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\varepsilon \text{ |type|}. \lambda x:\varepsilon \text{ |A|}. x \quad : \quad \|\Pi A:\text{type}. A \rightarrow A\|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \|\Pi A:\text{type}. A \rightarrow A\|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \varepsilon|\Pi A:\text{type}. A \rightarrow A|$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \varepsilon(\dot{\pi} \text{ |type| } (\lambda A: \text{ ||type||}. \dots))$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \varepsilon(\dot{\pi} \text{ type } (\lambda A:\text{type}. \dots))$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \Pi A:\varepsilon \text{type}. (\lambda A:\text{type}. \dots)A$$

Example

Example

Polymorphic identity in CC

$$\lambda A:\text{type}. \lambda x:A. x \quad : \quad (\Pi A:\text{type}. A \rightarrow A)$$

Translation:

$$\lambda A:\text{type}. \lambda x:\varepsilon A. x \quad : \quad \Pi A:\text{type}. \varepsilon A \rightarrow \varepsilon A$$

Properties

Theorem

If $\Gamma \vdash t : A$ then $\|\Gamma\| \vdash |t| : \|A\|$. (*Soundness*)

Proof.

In [Cousineau and Dowek 2007]. □

Theorem

If $\|\Gamma\| \vdash t : \|A\|$ then $\Gamma \vdash t' : A$. (*Conservativity*)

Proof.

Coming soon!

Partial result in [Cousineau and Dowek 2007]. □

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

HOL

- ▶ A family of theorem provers
 - (HOL Light, HOL4, ProofPower, ...)
- ▶ Based on *higher order logic*
- ▶ Large formalizations
 - Flyspeck project (Kepler's conjecture)

Higher order logic

- ▶ Formulas are terms of the simply-typed λ -calculus

$$\begin{array}{l} \text{types } A, B \quad ::= \text{prop} \mid A \rightarrow B \\ \text{terms } t, u \quad ::= x \mid \lambda x:A. t \mid t u \mid c \end{array}$$

- ▶ Can be seen as a PTS

$$\begin{array}{l} \mathcal{S} \quad = \text{prop, type, kind} \\ \mathcal{A} \quad = (\text{prop} : \text{type}), (\text{type} : \text{kind}) \\ \mathcal{R} \quad = (\text{prop, prop, prop}), (\text{type, prop, prop}), \\ \quad \quad (\text{type, type, type}), (\text{kind, kind, kind}) \end{array}$$

Challenges

- ▶ Classical formulation
 - Equality as a primitive connective, no PTS structure
 - Axioms (β , extensionality, choice, ...)
- ▶ No proof inspection
 - LCF architecture: theorems as an abstract datatype
 - Proofs are not kept in memory
- ▶ Proof size
 - No computation in proofs (β , definitions, ...)
 - Huge proof trees
- ▶ Free variables
 - Implicit quantification
 - Name clashes

Proof retrieval

HOL kernel

- ▶ OCaml module defining an abstract datatype for theorems + methods for constructing theorems
- ▶ The HOL kernel must be modified!

OpenTheory project [Hurd 2011]

- ▶ A standard for exporting and exchanging HOL proofs
- ▶ A well-defined standard library

The OpenTheory article format

Instructions that are executed to reconstruct the theorems.

Example

Reflexivity on a variable x of type A :

$$\frac{}{\vdash x^A = x^A} \text{ refl}$$

OpenTheory article file

```
"A"
varType
"x"
var
varTerm
refl
```

OCaml execution

```
let A = varType("A") in
let x = varTerm(var("x", A)) in
refl x
```

Proof size

Example

A proof of $t + t = u + u$:

let $p = \dots$ (*A very large proof of $t = u$*) in
`appThm (appThm (refl f) p p)`

$$\frac{\frac{\frac{}{(+)=(+)} \text{refl} \quad \frac{\pi}{t=u}}{(+)t=(+)u} \text{appThm} \quad \frac{\pi}{t=u}}{(+)tt=(+)uu} \text{appThm}$$

Need proof sharing!

Proof sharing

- ▶ Share common subproofs
- ▶ Already provided (partially) by OpenTheory
 - Going further: factorize every step

```
step1 : proof (t = u) := ...
```

```
step2 : proof ((+) = (+)) := refl q.
```

```
step3 : proof ((+) t = (+) u) := appThm step2 step1.
```

```
step4 : proof ((+) t t = (+) u u) := appThm step3 step1.
```

- ▶ Need lambda lifting
 - Abstract over free variables
 - α -renaming to avoid clashes

Proof sharing

- ▶ Share common subproofs
- ▶ Already provided (partially) by OpenTheory
 - Going further: factorize every step

```
step1 : proof (t = u) := ...
```

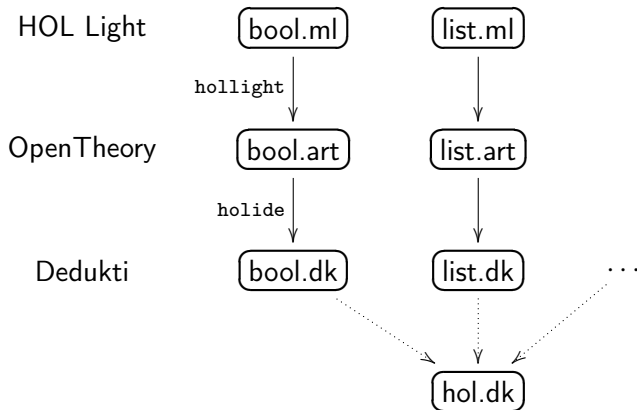
```
step2 : proof ((+) = (+)) := refl q.
```

```
step3 : proof ((+) t = (+) u) := appThm step2 step1.
```

```
step4 : proof ((+) t t = (+) u u) := appThm step3 step1.
```

- ▶ Need lambda lifting
 - Abstract over free variables
 - α -renaming to avoid clashes
- ▶ To-do: extend to terms

Translation process



Results

Package	Size (in kB)		Time (in s)		Percentage
	OpenTheory	Dedukti	Translation	Verification	Verified
unit	26	309	0	0	100%
function	89	1,301	1	3	100%
pair	195	4,943	4	15	100%
bool	305	4,258	2	7	100%
sum	502	20,988	17	99	100%
option	520	23,815	18	77	100%
relation	971	42,572	49	350	100%
list	1377	68,031	39	182	100%
real	1754	68,508	46	1	1%
natural	1952	130,111	38	496	100%
set	2329	90,819	65	431	100%
Total	10020	455,656	280	1,661	85%

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

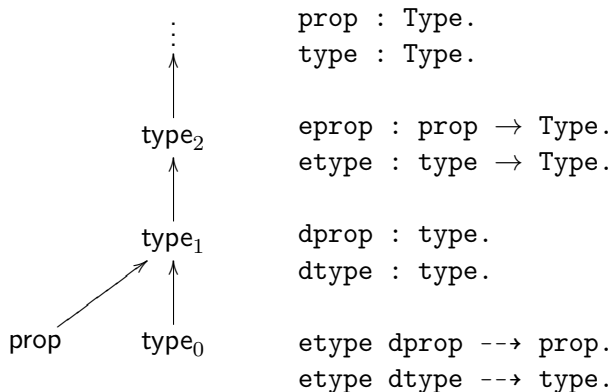
Coq

- ▶ Proof system based on the calculus of inductive constructions
 - Infinite hierarchy of universes type_i
 - Subtyping $\text{type}_i \subseteq \text{type}_{i+1}$
 - Floating universes
 - Inductive types
 - Co-inductive types
 - Modules
 - ...
- ▶ Large formalizations
 - 4 color theorem, Feit–Thompson theorem, ...

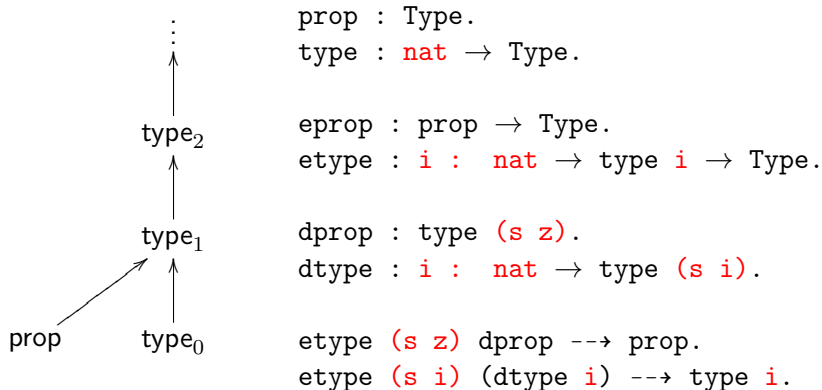
Challenges

- ▶ Rich and complex theory (CIC)
 - Much more than a PTS
 - Can we encode it in the $\lambda\Pi$ -calculus modulo?
- ▶ Proof retrieval
 - No direct export functions
- ▶ Implementation jungle
 - No complete specification
 - Huge code base (100 000s lines of code)
 - Poorly documented

Infinite universe hierarchy



Infinite universe hierarchy



Subtyping

In Coq

$$\frac{\Gamma \vdash t : A \quad \Gamma \vdash B : s \quad A \subseteq B}{\Gamma \vdash t : B} \qquad \frac{A \equiv B}{A \subseteq B}$$

$$\frac{}{\text{prop} \subseteq \text{type}_i} \qquad \frac{i \leq j}{\text{type}_i \subseteq \text{type}_j} \qquad \frac{B \subseteq C}{\prod x : A. B \subseteq \prod x : A. C}$$

In Dedukti

- ▶ Confluence \implies uniqueness of types
- ▶ Uniqueness of types \implies no shallow encoding
- ▶ Use type casting

Type casting

- ▶ No implicit type casting

$$\frac{\Gamma \vdash t : \varepsilon A \quad A \subseteq B}{\Gamma \vdash t : \varepsilon B}$$

- ▶ Cast function

$$\text{cast} : A \subseteq B \rightarrow \varepsilon A \rightarrow \varepsilon B$$

- ▶ Insert when needed

$$\frac{\Gamma \vdash t : \varepsilon A \quad \Gamma \vdash h : A \subseteq B}{\Gamma \vdash \text{cast } h t : \varepsilon B}$$

- ▶ Need to define $A \subseteq B$

Subtyping relation

As an inductive relation

sub_refl : $A \subseteq A$

sub_prop : $\text{prop} \subseteq \text{type}_i$

sub_type : $\text{type}_i \subseteq \text{type}_{i+j}$

sub_prod : $(\prod x:\varepsilon A. (B x) \subseteq (C x)) \rightarrow \dot{\pi} A B \subseteq \dot{\pi} A C$

- ▶ \implies large cast annotations \times

As a rewrite system

$\text{prop} \subseteq \text{type}_i \quad \dashrightarrow \quad \text{prop} \subseteq \text{prop}$

$\text{type}_i \subseteq \text{type}_{i+j} \quad \dashrightarrow \quad \text{type}_i \subseteq \text{type}_i$

$\dot{\pi} A B \subseteq \dot{\pi} A C \quad \dashrightarrow \quad \prod x:\varepsilon A. (B x) \subseteq (C x)$

- ▶ \implies canonical proof sub_refl : $A \subseteq A$ \checkmark

Inductive types

In Coq

```
Inductive nat : Type :=
| z : nat
| s : nat -> nat.
```

In Dedukti

```
nat : type.
z : etype nat.
s : etype nat → etype nat.
```

Pattern matching:

```
nat_case : P : (etype nat → type) →
  fz : etype (P z) →
  fs : (n : eset nat → etype (P (s n))) →
  n : etype nat → etype (P n).
[... ] nat_case P fz fs z --> fz.
[... ] nat_case P fz fs (s x) --> fs n.
```

Fixpoints

Translate fixpoints using rewriting rules

```
Fixpoint f x1 ... xn struct xn := t.
```

```
[...] f x1 ... xn --> t. ✗
```

Termination: CIC fixpoints are unfolded only if their structural argument starts with a constructor.

- ▶ Need to give the same semantics in the $\lambda\Pi$ -calculus modulo

Fixing fixpoints

- ▶ Cannot use pattern matching because of dependent types

```
Fixpoint f (x y : A) (p : x = y) := t.
```

```
[...] f x y (eq_refl x) --> t. ✗
```

Fixing fixpoints

- ▶ Cannot use pattern matching because of dependent types

```
Fixpoint f (x y : A) (p : x = y) := t.
```

```
[...] f x y (eq_refl x) --> t. ✗
```

- ▶ Add guards to constructors

```
pre_nat : type.
```

```
z : etype pre_nat.
```

```
s : etype nat → etype pre_nat.
```

```
nat_constr : pre_nat → nat.
```

$$\begin{aligned} |z| &= \text{nat_constr } z \\ |sn| &= \text{nat_constr } (s \ |n|) \end{aligned}$$

```
[...] f x1 ... (nat_constr xn) --> t. ✓
```

- Increased term size
- no syntactic equivalence

Fixing fixpoints

- ▶ Cannot use pattern matching because of dependent types

```
Fixpoint f (x y : A) (p : x = y) := t.
```

```
[...] f x y (eq_refl x) --> t. ✗
```

- ▶ Add guards to constructors

```
pre_nat : type.
```

```
z : etype pre_nat.
```

```
s : etype nat → etype pre_nat.
```

```
nat_constr : pre_nat → nat.
```

$$\begin{aligned} |z| &= \text{nat_constr } z \\ |s\ n| &= \text{nat_constr } (s\ |n|) \end{aligned}$$

```
[...] f x1 ... (nat_constr xn) --> t. ✓
```

- Increased term size
- no syntactic equivalence

Use recursors?

Modules

In Coq:

- ▶ Nested modules

```
Module A. ... Module B. ... End B. ... End A.
```

- ▶ Functors

```
Module F (A : T). ... End N.
```

- ▶ Module expressions

```
Module B := F(A).
```

In Dedukti:

- ▶ Only separate namespaces for files

Translating modules

Flatten nesting

- ▶ `A.B.C` translated as `A.B_C`

Module is expanded into its individual elements

`Module A := B.`

is translated into

`A_x := B_x.`

`A_y := B_y.`

`...`

Translating modules

Flatten nesting

- ▶ `A.B.C` translated as `A.B_C`

Module is expanded into its individual elements

`Module A (C) := B.`

is translated into

`A_x C_z1 C_z2 ... := B_x.`

`A_y C_z1 C_Z2 ... := B_y.`

`...`

- ▶ Increased term size

Proof retrieval

Coq proof scripts (.v) ✗

- ▶ Notations
- ▶ Implicit arguments
- ▶ Tactics

Coq compiled file (.vo) ✓

- ▶ Low-level proof terms
- ▶ Binary dump of memory content

Using .vo files

Advantages:

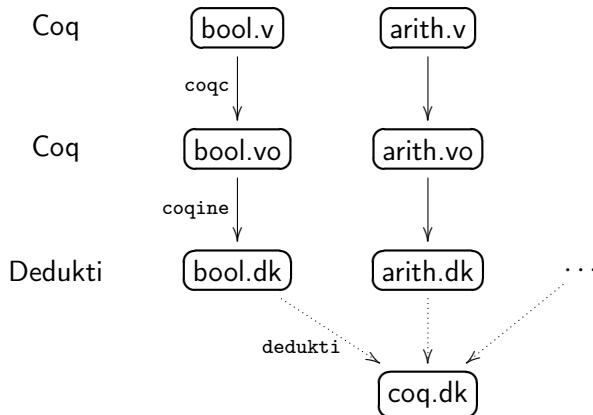
- ▶ Reconstructed proof terms
- ▶ Exactly the ones used by Coq
- ▶ Reuse Coq's code

Disadvantages:

- ▶ Rely on Coq's code
- ▶ Must use the same versions of both `ocaml` and `coqc`

Coqine is built on top of `coqchk`

Translation process



Results

- ▶ Working prototype
- ▶ Coq standard library:
 - Can translate 84%
 - Can type-check 20%
- ▶ File size:
 - `List.v`: 45 kB
 - `List.vo`: 260 kB
 - `List.dk`: 2.2MB

Future work

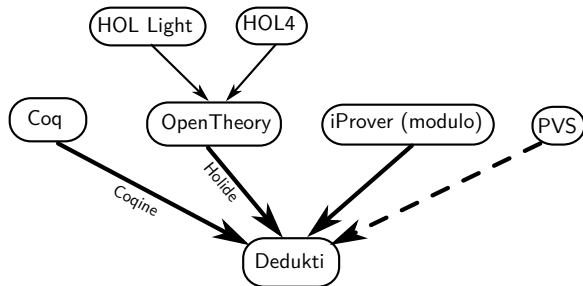
- ▶ Complete the current implementation
- ▶ Solve remaining features
 - Floating universes
 - ▶ Contradicts weakening
 - ▶ Non-modularity, side-effects
 - Co-inductive types
 - New coq features still coming!
- ▶ Long term: better specification for Coq

Outline

- Introduction
- Embedding pure type systems
- Holide: HOL in Dedukti
- Coqine: Coq in Dedukti
- Conclusion

Conclusion

- ▶ Embedding various logics can be fun!
- ▶ **Theoretical** and **practical** challenges
- ▶ Need good specifications and exporting standards

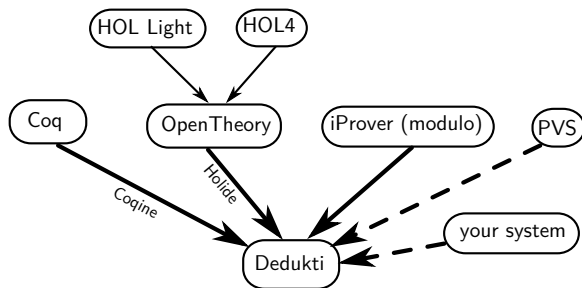


Software available at

<https://www.rocq.inria.fr/deducteam/software.html>

Conclusion

- ▶ Embedding various logics can be fun!
- ▶ **Theoretical** and **practical** challenges
- ▶ Need good specifications and exporting standards



Software available at

<https://www.rocq.inria.fr/deducteam/software.html>