An Effectful Way to Eliminate Addiction to Dependence

Pierre-Marie Pédrot\textsuperscript{1}  Nicolas Tabareau\textsuperscript{2}

\textsuperscript{1} University of Ljubljana, \textsuperscript{2} INRIA

Gallium Seminar
4th May 2017
The Most Important Issue of Them All

Let’s start this talk by a **fundamental** flaw of type theory.
The Most Important Issue of Them All

Let’s start this talk by a **fundamental** flaw of type theory.

- Assume you want to show the wonders of Coq to a fellow programmer
- You fire your favourite IDE
- ... and you’re asked the *dreadful* question.
Let’s start this talk by a fundamental flaw of type theory.

- Assume you want to show the wonders of Coq to a fellow programmer
- You fire your favourite IDE
- ... and you’re asked the dreadful question.
A Well-known Limitation

This is pretty much standard. By proof-as-program correspondence,

Intuitionistic Logic $\iff$ Functional Programming
A Well-known Limitation

This is pretty much standard. By proof-as-program correspondence,

Intuitionistic Logic $\iff$ Functional Programming

which means no effects in TT, amongst which:

- no exceptions
- no state
- no non-termination
- no printing
A Well-known Limitation

This is pretty much standard. By proof-as-program correspondence,

\[ \text{Intuitionistic Logic } \leftrightarrow \text{ Functional Programming} \]

which means no effects in TT, amongst which:

- no exceptions
- no state
- no non-termination
- no printing
- ... and thus no Hello World!
On Burritos

In less expressive settings, a few workarounds are known.

Typically, on the programming side, use the **monadic** style.

- A type $T : \Box \rightarrow \Box$
- A combinator `return` : $\alpha \rightarrow T \alpha$
- A combinator `bind` : $T \alpha \rightarrow (\alpha \rightarrow T \beta) \rightarrow T \beta$
- A few equations
On Burritos

In less expressive settings, a few workarounds are known.

Typically, on the programming side, use the monadic style.

- A type $T : \Box \rightarrow \Box$
- A combinator return: $\alpha \rightarrow T\alpha$
- A combinator bind: $T\alpha \rightarrow (\alpha \rightarrow T\beta) \rightarrow T\beta$
- A few equations

Interpret mechanically effectful programs using this (see Moggi).

This is pervasive in e.g. Haskell.
On the logic side, take the issue the other way around.
On the logic side, take the issue the other way around.

Effects are known to implement non-intuitionistic axioms!

- callcc $\sim$ classical logic (Griffin '90)
- exceptions $\sim$ Markov’s rule (Friedman’s trick)
- global monotonous cell $\sim \neg$CH (forcing)
- delimited continuations $\sim$ double negation shift
- ...

Achieve this using logical translations, e.g. double-negation.
We want a type theory with effects!

1. To program more (exceptions, non-termination...)
2. To prove more (classical logic, univalence...)
We want a type theory with effects!

1. To program more (exceptions, non-termination...)
2. To prove more (classical logic, univalence...)
3. To write Hello World.
The Expressivity Wall

Problem is:

Programming and logical techniques do not scale to type theory.
The Expressivity Wall

Problem is:

Programming and logical techniques do not scale to type theory.

- Monads do not acknowledge dependence

$$\text{bind} : T \alpha \rightarrow (\alpha \rightarrow T \beta) \rightarrow T \beta$$

$$\text{dbind} : \Pi x : T \alpha. (\Pi x : \alpha. T (\beta x)) \rightarrow T (\beta ?)$$

- They don’t acknowledge types-as-terms either
- And they don’t preserve the computational rules of TT
The Expressivity Wall

Problem is:

Programming and logical techniques do not scale to type theory.

- Monads do not acknowledge dependence

\[
\text{bind} : T \alpha \rightarrow (\alpha \rightarrow T \beta) \rightarrow T \beta
\]

\[
\text{dbind} : \Pi \hat{x} : T \alpha. (\Pi x : \alpha. T (\beta x)) \rightarrow T (\beta ?)
\]

- They don’t acknowledge types-as-terms either
- And they don’t preserve the computational rules of TT

On the other hand:

- Herbelin showed that CIC + callcc is unsound!
In This Talk

1. Adding a vast range of effects to (almost) full TT
   - reader (already done previously with the forcing translation)
   - writer, exceptions, non-termination, non-determinism...
   - All with the new weaning translation!
In This Talk

1. Adding a vast range of effects to (almost) full TT
   - reader (already done previously with the forcing translation)
   - writer, exceptions, non-termination, non-determinism...
   - All with the new weaning translation!

2. Implementing them thanks to program translations
   - No crazy category theory models!
   - So-called syntactic models.
   - Compile them on-the-fly into vanilla type theory!
In This Talk

1. Adding a vast range of effects to (almost) full TT
   - reader (already done previously with the **forcing translation**)
   - writer, exceptions, non-termination, non-determinism...
   - All with the new **weaning translation**!

2. Implementing them thanks to program translations
   - No crazy category theory models!
   - So-called **syntactic models**.
   - Compile them on-the-fly into vanilla type theory!

3. Introducing a generic notion of effectful dependent type theory
   - A simple, sensible restriction of dependent elimination
   - Seemingly compatible with all known effects
Syntactic Models

Define $[\cdot]$ on the syntax and derive the type interpretation $[[\cdot]]$ from it s.t.

$$\vdash M : A \quad \text{implies} \quad \vdash [M] : [[A]]$$

Obviously, that's subtle.

The correctness of $[[\cdot]]$ lies in the meta (Darn, Gödel!)

The translation must preserve typing (Not easy)

In particular, it must preserve conversion (Argh!)

Yet, a lot of nice consequences.

Does not require non-type-theoretical foundations (monism)

Can be implemented in your favourite proof assistant

Easy to show (relative) consistency, look at $[[False]]$

Easier to understand computationally

Pédrot & al. (U. Ljubljana & INRIA)

An Effectful Way
Define $[\cdot]$ on the syntax and derive the type interpretation $[\cdot]$ from it s.t.

$$\vdash M : A \quad \text{implies} \quad \vdash [M] : [A]$$

Obviously, that’s subtle.

- The correctness of $[\cdot]$ lies in the meta (Darn, Gödel!)
- The translation must preserve typing (Not easy)
- In particular, it must preserve conversion (Argh!)
Syntactic Models

Define $[\cdot]$ on the syntax and derive the type interpretation $\lbrack \cdot \rbrack$ from it s.t.

\[
\vdash M : A \quad \text{implies} \quad \vdash \lbrack M \rbrack : \lbrack A \rbrack
\]

Obviously, that’s subtle.

- The correctness of $[\cdot]$ lies in the meta (Darn, Gödel!)
- The translation must preserve typing (Not easy)
- In particular, it must preserve conversion (Argh!)

Yet, a lot of nice consequences.

- Does not require non-type-theoretical foundations (monism)
- Can be implemented in your favourite proof assistant
- Easy to show (relative) consistency, look at $\lbrack \text{False} \rbrack$
- Easier to understand computationally
There are two essential properties of TT that need to be explicited.

#1. Type theory is call-by-name by construction. This is because of the unrestricted conversion rule. But the usual monadic interpretation is call-by-value! We need to rely on an alternative decomposition (based on CBPV).

#2. Dependent elimination is hardcore intuitionistic. It rules out non-standard inductive terms that exist in CBN + effects. Reminiscent of Brouwer vs. Bishop mathematics. Needs to be weakened in presence of effects (« Bishop-style TT »)
(Mis)understanding Dependent Type Theory

There are two essential properties of TT that need to be explicited.

**#1. Type theory is call-by-name by construction.**

- This is because of the unrestricted conversion rule.
- But the usual monadic interpretation is call-by-value!
- We need to rely on an alternative decomposition (based on CBPV).
(Mis)understanding Dependent Type Theory

There are two essential properties of TT that need to be explicited.

#1. Type theory is call-by-name by construction.

- This is because of the unrestricted conversion rule.
- But the usual monadic interpretation is call-by-value!
- We need to rely on an alternative decomposition (based on CBPV).

#2. Dependent elimination is hardcore intuitionistic.

- It rules out non-standard inductive terms that exist in CBN + effects
- Reminiscent of Brouwer vs. Bishop mathematics
- Needs to be weakened in presence of effects (« Bishop-style TT »)
TT is intrisically call-by-name because of the conversion rule:

$$\frac{\Gamma \vdash M : B \quad A \equiv_{\beta} B}{\Gamma \vdash M : A}$$

where $\equiv_{\beta}$ is generated by:

$$(\lambda x : A. M) \ N \equiv_{\beta} M\{x := N\}$$
My Name is Call, Call-by-Name

TT is intrisically call-by-name because of the conversion rule:

\[
\frac{\Gamma \vdash M : B \quad A \equiv_{\beta} B}{\Gamma \vdash M : A}
\]

where \(\equiv_{\beta}\) is generated by:

\[
(\lambda x : A. M) \ N \equiv_{\beta} M\{x := N\}
\]

To be call-by-value, it would require instead \(\equiv_{\beta v}\) generated by:

\[
(\lambda x : A. M) \ V \equiv_{\beta v} M\{x := V\}
\]

where \(V\) is a value. But that’s not TT...
Tell Me Eleinberg-Moore

Turns out it is easy to give a call-by-name monadic decomposition.

Use the Eleinberg-Moore category, i.e. the category of algebras.
Tell Me Eleinberg-Moore

Turns out it is easy to give a call-by-name monadic decomposition.

Use the Eleinberg-Moore category, i.e. the category of algebras.

For us, a $T$-algebra will be an inhabitant of:

$$\Box := \Sigma A : \Box. \; T \; A \to A$$

A few remarks:

- It is hard to formulate the notion of algebra without higher-order types
- We don’t require any equations in $\Box$ (they’re quite not algebras)
- It turns out it is not necessary...
Required structure

We assume a monad given by universe-polymorphic terms:

\[
T : \square_i \rightarrow \square_i \\
\text{ret} : \Pi(A : \square). A \rightarrow TA \\
\text{bind} : \Pi(A B : \square). TA \rightarrow (A \rightarrow TB) \rightarrow TB
\]

and we require no equations!!
We assume a monad given by universe-polymorphic terms:

\[
\begin{align*}
T & : \square_i \to \square_i \\
\text{ret} & : \Pi(A : \square). A \to T \ A \\
\text{bind} & : \Pi(A \ B : \square). T \ A \to (A \to T \ B) \to T \ B
\end{align*}
\]

and we require no equations!!

Furthermore, in Type Theory, types are terms. We want the monad to be self-algebraic. This is given by:

\[
\begin{align*}
\text{El} & : T \square_i \to \square_i \\
\text{El} \ (\text{ret} \ \square \ M) & \equiv_{\beta} M
\end{align*}
\]

A lot of monads appear to be self-algebraic.
The Weaning Translation of the Negative Fragment

\[
\begin{align*}
[x] & := x \\
[\lambda x : A. M] & := \lambda x : [A]. [M] \\
[M \, N] & := [M] [N] \\
[\square_i] & := \text{ret } \square_{i+1} (T \, \square_i, \mu_{\square}) \\
[\Pi x : A. \, B] & := \text{ret } \square (\Pi x : [A]. [B], \mu_{\Pi}) \\
[A] & := (\text{El } [A]).\pi_1 \\
\mu_{\square} & : T (T \, \square) \rightarrow \square \\
\mu_{\Pi} & : T (\Pi x : [A]. [B]) \rightarrow \Pi x : [A]. [B]
\end{align*}
\]
The Weaning Translation of the Negative Fragment

\[
\begin{align*}
[x] & := x \\
[\lambda x : A. M] & := \lambda x : [A]. [M] \\
[M N] & := [M] [N] \\
[\square_i] & := \text{ret } \square_{i+1} (T \square_i, \mu_\square) \\
[\Pi x : A. B] & := \text{ret } \square (\Pi x : [A]. [B], \mu_\Pi) \\
[A] & := (\text{El } [A]). \pi_1 \\
\mu_\square & : T (T \square) \to \square \\
\mu_\Pi & : T (\Pi x : [A]. [B]) \to \Pi x : [A]. [B]
\end{align*}
\]

- Functional fragment untouched, types mangled into algebras
- \([\square] \equiv_\beta T \square\) and \([\Pi x : A. B] \equiv_\beta \Pi x : [A]. [B]\)
The Weaning Translation of the Negative Fragment

\[
\begin{align*}
[x] & := x \\
[\lambda x : A. M] & := \lambda x : [A]. [M] \\
[M \; N] & := [M] \; [N] \\
[\square_i] & := \text{ret } \square_{i+1} (T \; \square_i, \mu_{\square}) \\
[\Pi x : A. \; B] & := \text{ret } \square (\Pi x : [A]. [B], \mu_{\Pi}) \\
[A] & := (\text{El } [A]).\pi_1 \\
\mu_{\square} & : T (T \; \square) \rightarrow \square \\
\mu_{\Pi} & : T (\Pi x : [A]. [B]) \rightarrow \Pi x : [A]. [B]
\end{align*}
\]

- Functional fragment untouched, types mangled into algebras
- \([\square] \equiv_{\beta} T \; \square\) and \([\Pi x : A. \; B] \equiv_{\beta} \Pi x : [A]. [B]\)

Soundness

If \(\Gamma \vdash M : A\) then \([\Gamma] \vdash [M] : [A]\). (In particular, conversion is preserved.)
Reduction vs. Effects

Nothing fancy in the negative fragment, by the well-known duality.

- Call-by-name: **functions** well-behaved vs. **inductives** ill-behaved
- Call-by-value: **inductives** well-behaved vs. **functions** ill-behaved
Nothing fancy in the negative fragment, by the well-known duality.

- Call-by-name: **functions** well-behaved vs. **inductives** ill-behaved
- Call-by-value: **inductives** well-behaved vs. **functions** ill-behaved

Why is that?

In call-by-name + effects, consider:

\[(\lambda b : \text{bool. } M) \text{ fail} \leadsto \text{non-standard inductive terms}\]

In call-by-value + effects, consider:

\[(\lambda b : \text{unit. fail}) \leadsto \text{invalid } \eta\text{-rule}\]
Weaning Inductive Types

For the sake of explanation, let’s focus on a very simple type:

\[
\text{Inductive bool} := \text{true} \mid \text{false}.
\]

We pose:

\[
\begin{align*}
[\text{bool}] & := \text{ret } \square (T \text{ bool}, \mu_{\text{bool}}) \\
[\text{true}] & := \text{ret bool true} \\
[\text{false}] & := \text{ret bool false} \\
\mu_{\text{bool}} & : T (T \text{ bool}) \to T \text{ bool}
\end{align*}
\]
Weaning Inductive Types

For the sake of explanation, let’s focus on a very simple type:

\[
\text{Inductive bool := true | false.}
\]

We pose:

\[
\begin{align*}
[\text{bool}] & := \text{ret } \square (T \text{ bool}, \mu_{\text{bool}}) \\
[\text{true}] & := \text{ret } \text{bool} \text{ true} \\
[\text{false}] & := \text{ret } \text{bool} \text{ false} \\
\mu_{\text{bool}} & : \ T (T \text{ bool}) \rightarrow T \text{ bool}
\end{align*}
\]

Remark that \([\text{bool}] \equiv_\beta T \text{ bool} \).
We need a bit more structure on \( T \) to implement elimination:

\[
\begin{align*}
\text{hbind} & : \prod(A : \Box)(B : T \Box). T A \to (A \to [B]) \to [B] \\
\text{dbind} & : \prod(A : \Box)(B : A \to T \Box). \Pi(\hat{x} : T A). \\
& \quad \quad (\Pi(x : A). [B \ x]) \to (\text{El} (\text{hbind} A [\Box] \hat{x} B)). \pi_1
\end{align*}
\]

subject to:

\[
\begin{align*}
\text{hbind} A B (\text{ret} A M) F & \equiv_\beta F M \\
\text{dbind} A B (\text{ret} A M) F & \equiv_\beta F M
\end{align*}
\]

Essentially, \text{hbind} and \text{dbind} are variants of \text{bind}.
We need a bit more structure on $\mathcal{T}$ to implement elimination:

\[
\begin{align*}
\text{hbind} & : \Pi(A : \Box)(B : \mathcal{T} \Box). \mathcal{T} A \to (A \to [B]) \to [B] \\
\text{dbind} & : \Pi(A : \Box)(B : A \to \mathcal{T} \Box). \Pi(\hat{x} : TA).
\quad (\Pi(x : A). [B x]) \to (\text{El} (\text{hbind} A [\Box] \hat{x} B)). \pi_1
\end{align*}
\]

subject to:

\[
\begin{align*}
\text{hbind} A B \ (\text{ret} A M) F & \equiv_\beta F M \\
\text{dbind} A B \ (\text{ret} A M) F & \equiv_\beta F M
\end{align*}
\]

Essentially, $\text{hbind}$ and $\text{dbind}$ are variants of bind.

Remark that the second equation is well-typed iff the first holds.
Interpreting Non-Dependent Elimination

It is easy to provide a non-dependent eliminator using $\text{hbind}$:

$$
\begin{align*}
[\text{bool\_case}] & : \ [\Pi P : \Box. P \to P \to \text{bool} \to P] \\
& := \lambda (P : T \Box) \ (p_t \ p_f : [P]) \ (\hat{b} : T \text{bool}) \ . \\
& \hspace{1em} \text{hbind bool } P \ \hat{b} \ (\lambda b. \text{if } b \text{ then } p_t \text{ else } p_f)
\end{align*}
$$

which has the right reduction rules:

$$
\begin{align*}
[\text{bool\_case } P \ p_t \ p_f \ \text{true}] & \ \equiv_{\beta} \ p_t \\
[\text{bool\_case } P \ p_t \ p_f \ \text{false}] & \ \equiv_{\beta} \ p_f
\end{align*}
$$

Remember:

$$
\begin{align*}
\text{hbind} : & \Pi (A : \Box) (B : T \Box). \ T \ A \to (A \to [B]) \to [B] \\
\text{hbind } & \ A \ B \ (\text{ret } A \ M) \ F \ \equiv_{\beta} \ F \ M
\end{align*}
$$
We would like to recover dependent elimination...
Eliminating Addiction to Dependence

We would like to recover dependent elimination...

... but it’s not valid anymore in presence of effects!

As $\begin{array}{c}
\texttt{bool} \\
\equiv_{\beta} \\
T \texttt{bool},
\end{array}$ if $T$ is not the identity then there are closed booleans in the translation which are neither $\texttt{true}$ nor $\texttt{false}$. 
We would like to recover dependent elimination...

... but it’s not valid anymore in presence of effects!

As $[[\text{bool}]] \equiv \beta T \text{bool}$, if $T$ is not the identity then there are closed booleans in the translation which are neither $[\text{true}]$ nor $[\text{false}]$.

- Typical of CBN + effects: recall Herbelin’s paradox
- Already arose in our forcing translation
- We need to restrict dependent elimination the same way!
The trick consists in sprinkling a few storage operators. For bool:

\[
\theta_{\text{bool}} : [[\text{bool} \to (\text{bool} \to \Box) \to \Box]] := \lambda b. \text{bool\_case } (\text{bool} \to \Box) (\lambda k. k \text{ true}) (\lambda k. k \text{ false}) b
\]

- Only defined in the source via non-dependent eliminator
- In particular, agnostic to the actual translation
- CPS-like to enforce CBV in a CBN world
- Trivial in CIC: \( \vdash \Pi b : \text{bool}. \theta_{\text{bool}} b P = P b \)
The trick consists in sprinkling a few storage operators. For bool:

\[
[\theta_{\text{bool}}] : \left[[\text{bool} \to (\text{bool} \to \Box) \to \Box]\right] \\
:= \left[\lambda b. \text{bool\_case} (\text{bool} \to \Box) (\lambda k. k \text{ true}) (\lambda k. k \text{ false}) b\right]
\]

- Only defined in the source via non-dependent eliminator
- In particular, agnostic to the actual translation
- CPS-like to enforce CBV in a CBN world
- Trivial in CIC: \( \vdash \Pi b : \text{bool}. \theta_{\text{bool}} b P = P b \)

Using `dbind`, this allows to implement:

\[
[\text{bool\_rect}] : \left[\Pi P : \text{bool} \to \Box. P \text{ true} \to P \text{ false} \to \Pi b : \text{bool}. \theta_{\text{bool}} b P\right]
\]

with the expected reduction rules.
There are a lot of monads that satisfy the weaning conditions.

- Exception monad \( T A := A + E \)
- Non-determinism \( T A := A \times \text{list} \ A \)
- Non-termination \( T A := \nu X. A + X \)
- Writer \( T A := A \times \text{list} \ \Omega \) (the one we need for `Hello World`)

Note that some lead to a logically inconsistent model.
There are a lot of monads that satisfy the weaning conditions.

- Exception monad \( T A := A + E \)
- Non-determinism \( T A := A \times \text{list } A \)
- Non-termination \( T A := \nu X. A + X \)
- Writer \( T A := A \times \text{list } \Omega \) (the one we need for \texttt{Hello World})

Note that some lead to a logically inconsistent model.

A few monads aren’t self-algebraic, e.g. state, reader and continuation.
In some inconsistent cases, full dependent elimination is valid. Most notably, this is the case for the exception monad.

Let’s use that to do a Friedman $A$-translation on steroids!
Logic, at Last

In some inconsistent cases, full dependent elimination is valid. Most notably, this is the case for the exception monad.

Let’s use that to do a Friedman $A$-translation on steroids!

Lemmatas

With the exception monad $T A := A + E$:

- Full dependent elimination is valid (at the expense of consistency)
- We have $\neg\neg A \cong ([A] \rightarrow E) \rightarrow E$
- If $A$ is a first-order type, then $[A] \rightarrow A + E$. 
Logic, at Last

In some inconsistent cases, full dependent elimination is valid. Most notably, this is the case for the exception monad.

Let’s use that to do a Friedman \( A \)-translation on steroids!

Lemmatas

With the exception monad \( T \ A := A + E \):

- Full dependent elimination is valid (at the expense of consistency)
- We have \( \llbracket \neg \neg A \rrbracket \cong (\llbracket A \rrbracket \to E) \to E \)
- If \( A \) is a first-order type, then \( \llbracket A \rrbracket \to A + E \).

Admissibility of Markov’s rule in CIC

If \( A \) is first-order and \( \vdash_{\text{CIC}} \neg \neg A \) then \( \vdash_{\text{CIC}} A \).
Moi, j'ai dit linéaire, linéaire ? Comme c'est étrange...

Back to restricted elimination. It turns out we have a semantic criterion for valid dependent predicates.

LINEARITY.
Back to restricted elimination. It turns out we have a semantic criterion for valid dependent predicates.

**LINEARITY.**

- A concept invented by G. Munch, rephrased recently by P. Levy.
- Little to do with « linear use of variables »
- Essentially, \( f : A \to B \) linear in CBN if semantically CBV in \( A \).
- Categorically, \( f \) linear iff it is an algebra morphism.
- Storage operators turn freely any morphism into a linear one.
- Can be approximated by a syntactic guard condition.

\[
\Gamma \vdash M : \text{bool} \quad \ldots \quad P \text{ linear in } b \\
\Gamma \vdash \text{if } M \text{ return } \lambda b. P \text{ then } N_1 \text{ else } N_2 : P\{b := M\}
\]
A Bishop-style Type Theory

We can generalize this restriction to form **Baclofen Type Theory**.

- Subset of CIC
- Independent from the actual translation.
- Works with forcing
- Works with weaning
- Prevents Herbelin’s paradox
We can generalize this restriction to form **Baclofen Type Theory**.

- Subset of CIC
- Independent from the actual translation.
- Works with forcing
- Works with weaning
- Prevents Herbelin’s paradox

**BTT is the generic theory to deal with dependent effects**

« Bishop-style, effect-agnostic type theory »

(Take that, Brouwerian HoTT!)
Implementation

A nice paper summarizing this talk.

https://www.pédrot.fr/articles/weaning.pdf

Just as for the forcing translation we have a Coq plugin for weaning.

https://github.com/CoqHott/coq-effects

- Allows to add effects to Coq just today.
- Implement your favourite effectful operators: fail, fix...
- Compile effectful terms on the fly.
- Allows to reason about them in Coq.

(If time permits, small demo here.)
A new effectful translation of TT, the weaning translation
  - Cosmic version of Eilenberg-Moore categories
  - Gives both programming and logical features

An experimentally confirmed notion of effectful type theories, BTT
  - Works for forcing, weaning and CPS
  - Restriction of dependent elimination on linearity guard condition
  - Conjecture: the correct way to add effects to TT

Implementation of a plugin in Coq
  - Try it out today!
Thanks for your attention.