Normalization by evaluation for $\lambda^{\to 2}$

Thorsten Altenkirch[1] ⋆ and Tarmo Uustalu[2] ⋆⋆

[1] School of Computer Science and IT, University of Nottingham
Nottingham NG8 1BB, United Kingdom
txa@cs.nott.ac.uk

[2] Institute of Cybernetics, Tallinn Technical University
Akadeemia tee 21, EE-12618 Tallinn, Estonia
tarmo@cs.ioc.ee

Abstract. We show that the set-theoretic semantics for $\lambda^{\to 2}$ is complete by inverting evaluation using decision trees. This leads to an implementation of normalization by evaluation which is witnessed by the source of this paper being a literate Haskell script. We show the correctness of our implementation using logical relations.

1 Introduction

Which is the simplest typed λ-calculus without uninterpreted base types or type variables? We suggest that the answer should be $\lambda^{\to 2}$: simply typed lambda calculus extended by the type of booleans $\mathtt{Bool}$ with $\mathtt{True}, \mathtt{False} : \mathtt{Bool}$ and $\mathtt{If}\, t\, u_0\, u_1 : \sigma$, given $t : \mathtt{Bool}$ and $u_0, u_1 : \sigma$. The equational theory is given by the usual $\beta\eta$-equations of $\lambda^{\to}$ and the following equations concerning $\mathtt{Bool}$:

$$
\begin{aligned}
\mathtt{If}\,\mathtt{True}\, u_0\, u_1 &=_{\beta\eta} u_0 \\
\mathtt{If}\,\mathtt{False}\, u_0\, u_1 &=_{\beta\eta} u_1 \\
\mathtt{If}\, t\,\mathtt{True}\,\mathtt{False} &=_{\beta\eta} t \\
v\,(\mathtt{If}\, t\, u_0\, u_1) &=_{\beta\eta} \mathtt{If}\, t\,(v\, u_0)\,(v\, u_1)
\end{aligned}
$$

The equations are motivated by the categorical interpretation of $\mathtt{Bool}$ as a boolean object, i.e., an object Bool such that $\mathrm{Hom}(\Gamma \times \mathrm{Bool}, A) \simeq \mathrm{Hom}(\Gamma, A) \times \mathrm{Hom}(\Gamma, A)$ (naturally in Γ and A). The calculus can thus be interpreted in any cartesian closed category with Bool (using the cartesian structure to interpret contexts).

The equational theory introduces some interesting equalities. E.g., consider

$$
\begin{aligned}
\mathsf{once} &= \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\, \lambda x : \mathtt{Bool}.\, f\, x \\
\mathsf{thrice} &= \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\, \lambda x : \mathtt{Bool}.\, f\,(f\,(f\, x))
\end{aligned}
$$

⋆ partially supported by EU Framework 5 thematic networks TYPES and APPSEM II and the Estonian IT Foundation

⋆⋆ partially supported by the Estonian Science Foundation under grant No. 5567

We observe that $\mathsf{once} =_{\beta\eta} \mathsf{thrice}$. To see this, we note that, given $f : \mathtt{Bool} \to \mathtt{Bool}$, we have

$$\begin{aligned}
f\,(f\,(f\,\mathtt{True})) &=_{\beta\eta} \mathtt{If}\,(f\,\mathtt{True})\,(f\,(f\,\mathtt{True}))\,(f\,(f\,\mathtt{False})) \\
&=_{\beta\eta} \mathtt{If}\,(f\,\mathtt{True})\,\mathtt{True}\,(f\,(f\,\mathtt{False})) \\
&=_{\beta\eta} \mathtt{If}\,(f\,\mathtt{True})\,\mathtt{True}\,(\mathtt{If}\,(f\,\mathtt{False})\,(f\,\mathtt{True})\,(f\,\mathtt{False})) \\
&=_{\beta\eta} \mathtt{If}\,(f\,\mathtt{True})\,\mathtt{True}\,(\mathtt{If}\,(f\,\mathtt{False})\,\mathtt{False}\,\mathtt{False}) \\
&=_{\beta\eta} \mathtt{If}\,(f\,\mathtt{True})\,\mathtt{True}\,\mathtt{False} \\
&=_{\beta\eta} f\,\mathtt{True}
\end{aligned}$$

Symmetrically, we can show that $f\,(f\,(f\,\mathtt{False})) =_{\beta\eta} f\,\mathtt{False}$, and hence

$$\begin{aligned}
\mathsf{thrice} &= \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\, \lambda x : \mathtt{Bool}.\, f\,(f\,(f\,x)) \\
&=_{\beta\eta} \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\, \lambda x : \mathtt{Bool}.\, \mathtt{If}\,x\,(f\,(f\,(f\,\mathtt{True})))\,(f\,(f\,(f\,\mathtt{False}))) \\
&=_{\beta\eta} \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\lambda x : \mathtt{Bool}.\, \mathtt{If}\,x\,(f\,\mathtt{True})\,(f\,\mathtt{False}) \\
&=_{\beta\eta} \lambda f : \mathtt{Bool} \to \mathtt{Bool}.\, \lambda x : \mathtt{Bool}. f\,x \\
&= \mathsf{once}
\end{aligned}$$

It is easy to see that once and thrice are equal in the standard semantics where $\mathtt{Bool}$ is interpreted by a two-element set $\mathrm{Bool} = \{\mathrm{true}, \mathrm{false}\}$ and function types are set-theoretic function spaces. We observe that there are only four elements in $\mathrm{Bool} \to \mathrm{Bool} = \{x \mapsto \mathrm{true}, x \mapsto x, x \mapsto \neg x, x \mapsto \mathrm{false}\}$ and that for all the four $f \in \mathrm{Bool} \to \mathrm{Bool}$ we have $f^3 = f$.

May we use set-theoretic reasoning to prove equalities up to $\beta\eta$-convertibility? The answer is yes for $\lambda^{\to 2}$, because for $\lambda^{\to 2}$ we can *invert* set-theoretic *evaluation* of typed closed terms. That is: we can define a function $\mathsf{quote}^\sigma \in [\![\sigma]\!]^{\mathsf{Set}} \to \mathsf{Tm}\ \sigma$ such that $t =_{\beta\eta} \mathsf{quote}^\sigma\ [\![t]\!]^{\mathsf{Set}}$, for any $t \in \mathsf{Tm}\ \sigma$. Consequently, we get that, for any $t, t' \in \mathsf{Tm}\ \sigma$, $t =_{\beta\eta} t' \iff [\![t]\!]^{\mathsf{Set}} = [\![t']\!]^{\mathsf{Set}}$.

The construction can be carried out in a constructive setting, e.g., using Martin-Löf's Type Theory as the metalanguage, and we obtain an implementation of normalization $\mathsf{nf}^\sigma\ t = \mathsf{quote}^\sigma\ [\![t]\!]^{\mathsf{Set}}$. We shall here use the language Haskell as a poor man's Type Theory and obtain a Haskell program to normalize terms. Our implementation is based on decision trees which lead to reasonably simple normal forms.

The existence of quote also implies that $=_{\beta\eta}$ is maximally consistent, i.e., identifying any two non-$\beta\eta$-convertible closed terms would lead to an inconsistent theory. This provides another justification for the specific choice of $=_{\beta\eta}$.

Related work

Inverting evaluation to achieve normalization by evaluation (NBE, aka. reduction-free normalization) was pioneered in [5] for simply typed lambda calculus with type variables and a non-standard semantics; a categorical account in terms of presheaves was given in [2]; this was extended to System F in [3, 4]; see [8] for a recent survey on NBE. The completeness of the set-theoretic model in the presence of coproducts has been shown in [7] and our case arises as a special case when there are no type variables. Normalization procedures for typed λ-calculus with coproducts can be found in [9, 10] using rewriting techniques and [1] using NBE and sheaf theory. Both approaches allow type variables but do not handle the empty type. Here we present a much simpler construction for closed types using the simplest possible semantics of first-order simply typed lambda calculi—the set-theoretic one—and also provide a concrete implementation of quote and nf, whose correctness we show in detail.

2 Implementation of $\lambda^{\to 2}$

The source code of this paper is a literate Haskell script implementing normalization for $\lambda^{\to 2}$ and is available from

`http://www.cs.nott.ac.uk/~txa/Nbe2.lhs`

We start by introducing types $\mathsf{Ty} \in \star$, typing contexts $\mathsf{Con} \in \star$ and untyped terms $\mathsf{UTm} \in \star$ of the object language by the following Haskell datatype definitions:

```
data Ty = Bool | Ty :-> Ty
        deriving (Show, Eq)

type Con = [ (String, Ty) ]

data UTm = Var String
         | TTrue | TFalse | If UTm UTm UTm
         | Lam Ty String UTm | App UTm UTm
        deriving (Show, Eq)
```

We view these recursive definitions as inductive definitions, i.e., we do not consider infinite terms. All the functions we define are total on finite terms.

Implementing typed terms $\mathsf{Tm} \in \mathsf{Con} \to \mathsf{Ty} \to \star$ would take inductive families, which we cannot use in Haskell. But we can implement type inference $\mathsf{infer} \in \mathsf{Con} \to \mathsf{UTm} \to \mathsf{Maybe}\ \mathsf{Ty}$ (where $\mathsf{Maybe}\ X = 1 + X$ as usual):

```
infer :: Con -> UTm -> Maybe Ty
infer gamma (Var x) =
    do sigma <- lookup x gamma
       Just sigma
infer gamma TTrue  = Just Bool
infer gamma TFalse = Just Bool
infer gamma (If t u0 u1) =
    do Bool <- infer gamma t
       sigma0 <- infer gamma u0
       sigma1 <- infer gamma u1
       if sigma0 == sigma1 then Just sigma0 else Nothing
```

```
infer gamma (Lam sigma x t) =
    do tau <- infer ((x, sigma) : gamma) t
       Just (sigma :-> tau)
infer gamma (App t u) =
    do (sigma :-> tau) <- infer gamma t
       sigma' <- infer gamma u
       if sigma == sigma' then Just tau else Nothing
```

This implementation is correct in the sense that $t \in \mathsf{Tm}_\Gamma\ \sigma$ iff $\mathsf{infer}_\Gamma\ t = \mathsf{Just}\ \sigma$.

Evaluation of types $[\![-]\!] : \mathsf{Ty} \to \star$ in a semantics is again an inductive family, which we cannot do in Haskell, and the workaround is to have all $[\![\sigma]\!]$ coalesced into one metalanguage type U (U for 'untyped elements') much the same way as all $\mathsf{Tm}_\Gamma\ \sigma$ appear coalesced in UTm. We use a type class Sem to state what we require of such a coalesced type U:

```
class Sem u where
    true  :: u
    false :: u
    xif :: u -> u -> u -> u
    lam :: Ty -> (u -> u) -> u
    app :: u -> u -> u
```

We introduce untyped environments $\mathsf{UEnv}_U \in \star$:

```
type UEnv u = [ (String, u) ]
```

And now we are ready to implement evaluation of terms $[\![-]\!] \in \mathsf{UEnv}_U \to \mathsf{UTm} \to U$:

```
eval :: Sem u => UEnv u -> UTm -> u
eval rho (Var x) = d
    where (Just d) = lookup x rho
eval rho TTrue  = true
eval rho TFalse = false
eval rho (If t u0 u1) = xif (eval rho t) (eval rho u0) (eval rho u1)
eval rho (Lam sigma x t) = lam sigma (\ d -> eval ((x, d) : rho) t)
eval rho (App t u) = app (eval rho t) (eval rho u)
```

We notice that $[\![-]\!]$ is type-sound: if $\rho \in \mathsf{Env}_\Gamma$ and $t \in \mathsf{Tm}_\Gamma\ \sigma$, then $[\![t]\!]_\rho \in [\![\sigma]\!]$.

The standard set-theoretic semantics can be represented in Haskell as an instance of Sem:

```
data SU = STrue | SFalse | SLam Ty (SU -> SU)

instance Sem SU where
    true  = STrue
    false = SFalse
    xif STrue  d _ = d
    xif SFalse _ d = d
```

```
    lam = SLam
    app (SLam _ f) d = f d
```

We write $[\![-]\!]^{\mathsf{Set}}$ to emphasize evaluation in the set-theoretic semantics (this is automatically inferred by Haskell).

Since sets form a cartesian closed category with a boolean object, the set-theoretic semantics validates all $\beta\eta$-equalities. This is to say that $[\![-]\!]^{\mathsf{Set}}$ is equationally sound:

Proposition 2.1 (Soundness). *If $\rho \in \mathsf{Env}_\Gamma$ and $t =_{\beta\eta} t' \in \mathsf{Tm}_\Gamma\ \sigma$, then $[\![t]\!]^{\mathsf{Set}}_\rho = [\![t']\!]^{\mathsf{Set}}_\rho$.*

Since all the sets we consider are finite, semantic equality can be implemented in Haskell, by making use of the function $\mathsf{enum} \in (\sigma \in \mathsf{Ty}) \to \mathsf{Tree}\ [\![\sigma]\!]$, which we will provide later:

```
instance Eq SU where
    STrue  == STrue  = True
    SFalse == SFalse = True
    (SLam sigma f) == (SLam _ f') =
        and [f d == f' d | d <- flatten (enum sigma)]
    _ == _ = False
```

Using on the same function we can also print elements of SU:

```
instance Show SU where
    show STrue  = "STrue"
    show SFalse = "SFalse"
    show (SLam sigma f) =
        "SLam " ++ (show sigma) ++ " " ++
        (show [ (d, f d) | d <- flatten (enum sigma) ])
```

The equational theory of the calculus itself gives rise to another semantics—the free semantics, or typed closed terms up to $\beta\eta$-convertibility. This semantics can be implemented as another instance of Sem (although of course we have to replace typed closed terms with UTm). To do this properly, we would have to take care of fresh name generation. But it turns out that in our uses of typed closed terms later as names of semantic elements there will be at most one bound variable at any time. We also notice that we get shorter normal forms by implementing a redundancy-avoiding version of `If` which produces a shorter but $\beta\eta$-equivalent term:

```
instance Sem UTm where
    true  = TTrue
    false = TFalse
    xif t TTrue TFalse = t
    xif t u0 u1 = if u0 == u1 then u0 else If t u0 u1
    lam sigma f = Lam sigma "x" (f (Var "x"))
    app = App
```

Also this semantics is clearly equationally sound.

3 Implementation of quote

We now proceed to implementing $\mathsf{quote} \in (\sigma \in \mathsf{Ty}) \to [\![\sigma]\!]^{\mathsf{Set}} \to \mathsf{Tm}\ \sigma$.

To define $\mathsf{quote}^{\sigma\to\tau}$ we use enum^{σ}, which generates a decision tree whose leaves are all the elements of $[\![\sigma]\!]$, and $\mathsf{questions}^{\sigma}$, which generates the list of questions, i.e. elements of $[\![\sigma]\!] \to [\![\mathtt{Bool}]\!]$, based on answers to whom an element of $[\![\sigma]\!]$ can be looked up in the tree enum^{σ}.

Decision trees $\mathsf{Tree} \in \mathsf{Ty} \to \star$ are provided by

```
data Tree u = Val u | Choice (Tree u) (Tree u) deriving (Show, Eq)
```

We will exploit the fact that Tree is a monad

```
instance Monad Tree where
    return = Val
    (Val d) >>= h = h d
    (Choice l r) >>= h = Choice (l >>= h) (r >>= h)
```

(`return` and `>>=` are Haskell for the unit resp. the bind or Kleisli extension operation of a monad) and hence a functor

```
instance Functor Tree where
    fmap h ds = ds >>= return . h
```

(`fmap` is Haskell for the action of a functor on morphisms).

It is convenient to use the function flatten which calculates the list of leaves of a given tree:

```
flatten :: Tree u -> [ u ]
flatten (Val d) = [ d ]
flatten (Choice l r) = (flatten l) ++ (flatten r)
```

We implement enum^{σ} and $\mathsf{questions}^{\sigma}$ by mutual induction on $\sigma \in \mathsf{Ty}$. The precise typings of the functions are $\mathsf{enum} \in (\sigma \in \mathsf{Ty}) \to \mathsf{Tree}\ [\![\sigma]\!]$ and $\mathsf{questions} \in (\sigma \in \mathsf{Ty}) \to [[\![\sigma]\!] \to [\![\mathtt{Bool}]\!]]$. As usual, Haskell cannot express those subtleties due to its lack of dependent types, but we can declare

```
enum :: Sem u => Ty -> Tree u
questions :: Sem u => Ty -> [ u -> u ]
```

The base case for is straightforward: A boolean is true or false, and to know which one it is it suffices to know which one it is.

```
enum Bool = Choice (Val true) (Val false)
```

```
questions Bool = [ \ b -> b ]
```

The implementation of $\mathsf{enum}^{\sigma\to\tau}$ and $\mathsf{questions}^{\sigma\to\tau}$ proceeds from the idea that a function is determined by its graph: to know a function it suffices to know its value on all possible argument values. The main idea in the implementation of $\mathsf{enum}^{\sigma\to\tau}$ is therefore to start with enum^{τ} and to duplicate the tree for each question in $\mathsf{questions}^{\sigma}$ using the bind of Tree:

```
enum (sigma :-> tau) =
    fmap (lam sigma) (mkEnum (questions sigma) (enum tau))

mkEnum :: Sem u => [ u -> u ] -> Tree u -> Tree (u -> u)
mkEnum [] es = fmap (\ e -> \ d -> e) es
mkEnum (q : qs) es = (mkEnum qs es) >>= \ f1 ->
                     (mkEnum qs es) >>= \ f2 ->
                     return (\ d -> xif (q d) (f1 d) (f2 d))
```

$\mathsf{questions}^{\sigma\to\tau}$ produces the appropriate questions by enumerating σ and using questions from τ:

```
questions (sigma :-> tau) =
    [ \ f -> q (app f d) | d <- flatten (enum sigma),
                           q <- questions tau ]
```

As an example, the enumeration and questions for `Bool` → `Bool` return:

```
Choice
   (Choice
       (Val (lam Bool (\ d -> xif d true  true)))
       (Val (lam Bool (\ d -> xif d true  false))))
   (Choice
       (Val (lam Bool (\ d -> xif d false true )))
       (Val (lam Bool (\ d -> xif d false false))))
```

resp.

```
(\ f -> app f true :
   (\ f -> app f false :
       []))
```

We can look up an element in a tree by answering all the questions—this leads to $\mathsf{find}^{\sigma} \in [[\![\mathtt{Bool}]\!]] \to \mathsf{Tree}\,[\![\sigma]\!] \to [\![\sigma]\!]$, in Haskell:

```
find :: Sem u => [ u ] -> Tree u -> u
find [] (Val t) = t
find (a : as) (Choice l r) = xif a (find as l) (find as r)
```

We are now ready to implement $\mathsf{quote}^{\sigma} \in [\![\sigma]\!]^{\mathsf{Set}} \to \mathsf{Tm}\,\sigma$, with Haskell typing

```
quote :: Ty -> SU -> UTm
```

by induction on $\sigma \in \mathsf{Ty}$. As usual, the base case is easy:

```
quote Bool STrue  = TTrue
quote Bool SFalse = TFalse
```

$\mathsf{quote}^{\sigma\to\tau}$ is more interesting: Our strategy is to map $\mathsf{quote}^{\tau} \circ f$ to the set-theoretic enum^{τ} and to then build a tree of `If` expressions by using the syntactic $\mathsf{questions}^{\sigma}$ in conjunction with the syntactic find:

```
quote (sigma :-> tau) (SLam _ f) =
    lam sigma (\ t -> find [ q t | q <- questions sigma ]
                           (fmap (quote tau . f) (enum sigma)))
```

(Notice that in Haskell it is inferred automatically which semantics is meant where.)

As already discussed in the introduction, we implement normalization $\mathsf{nf} \in (\sigma \in \mathsf{Ty}) \to \mathsf{Tm}\ \sigma \to \mathsf{Tm}\ \sigma$ by

```
nf :: Ty -> UTm -> UTm
nf sigma t = quote sigma (eval [] t)
```

Since we can infer types, we can implement $\mathsf{nf}' \in \mathsf{UTm} \to \mathsf{Maybe}\ ((\sigma \in \mathsf{Ty}) \times \mathsf{Tm}\ \sigma)$:

```
nf' :: UTm -> Maybe (Ty, UTm)
nf' t = do sigma <- infer [] t
           Just (sigma, nf sigma t)
```

We test our implementation with the example from the introduction:

```
b2b = Bool :-> Bool
once   = Lam b2b "f" (Lam Bool "x" (App (Var "f") (Var "x")))
twice  = Lam b2b "f" (Lam Bool "x" (App (Var "f")
                                     (App (Var "f") (Var "x"))))
thrice = Lam b2b "f" (Lam Bool "x" (App (Var "f")
                                     (App (Var "f")
                                       (App (Var "f") (Var "x")))))
```

and convince ourselves that $\mathsf{nf}'\ \mathsf{once} == \mathsf{nf}'\ \mathsf{thrice} = \mathsf{true}$ but $\mathsf{nf}'\ \mathsf{once} == \mathsf{nf}'\ \mathsf{twice} = \mathsf{false}$. Since semantic equality is decidable we do not actually have to construct the normal forms to decide convertibility.

Since testing can only reveal the presence of errors we shall use the rest of this paper to prove that quote and hence nf behave correctly.

4 Correctness of quote

The main tool in our proof will be a notion of logical relations, a standard tool for characterization of definable elements in models of typed lambda calculi since the pioneering work of Plotkin [11].

Definition 4.1 (Logical Relations). *We define a family of relations $\mathsf{R}^\sigma \subseteq \mathsf{Tm}\ \sigma \times [\![\sigma]\!]^{\mathsf{Set}}$ by induction on $\sigma \in \mathsf{Ty}$ as follows:*

- *if $t =_{\beta\eta}$ `True`, then $t \mathsf{R}^{\mathsf{Bool}} \mathrm{true}$;*
- *if $t =_{\beta\eta}$ `False`, then $t \mathsf{R}^{\mathsf{Bool}} \mathrm{false}$;*
- *if for all u, d, $u \mathsf{R}^\sigma d$ implies `App` $t\ u \mathsf{R}^\tau f\ d$, then $t \mathsf{R}^{\sigma\to\tau} f$.*

Logical relations obey the following Fundamental Theorem, a kind of substitution lemma for logical relations.

Lemma 4.2 (Fundamental Theorem of Logical Relations). *If* $\theta \mathsf{R}^{\Gamma} \rho$ *and* $t \in \mathsf{Tm}_{\Gamma}\ \sigma$, *then* $t[\theta]\ \mathsf{R}^{\sigma} [\![t]\!]^{\mathsf{Set}}_{\rho}$. *In particular, if* $t \in \mathsf{Tm}\ \sigma$, *then* $t \mathsf{R}^{\sigma} [\![t]\!]^{\mathsf{Set}}$.

The main result required to see that quote is correct is the following lemma:

Lemma 4.3 (Main Lemma). *If* $t \mathsf{R}^{\sigma} d$, *then* $t =_{\beta\eta} \mathsf{quote}^{\sigma}\ d$.

The proof of this lemma is the subject of the next section.

By correctness of quote we mean that it inverts set-theoretic evaluation of typed closed terms.

Theorem 4.4 (Main Theorem). *If* $t \in \mathsf{Tm}\ \sigma$, *then* $t =_{\beta\eta} \mathsf{quote}^{\sigma}\ [\![t]\!]^{\mathsf{Set}}$.

Proof. Immediate from the Fundamental Theorem and the Main Lemma. □

The (constructive) existence and correctness of quote has a number of straightforward important consequences.

Corollary 4.5 (Completeness). *If* $t, t' \in \mathsf{Tm}\ \sigma$, *then* $[\![t]\!]^{\mathsf{Set}} = [\![t']\!]^{\mathsf{Set}}$ *implies* $t =_{\beta\eta} t'$.

Proof. Immediate from the Main Theorem. □

From soundness (Proposition 2.1) and completeness together we get that $=_{\beta\eta}$ is decidable: checking whether $t =_{\beta\eta} t'$ reduces to checking whether $[\![t]\!]^{\mathsf{Set}} = [\![t']\!]^{\mathsf{Set}}$, which is decidable as $[\![-]\!]^{\mathsf{Set}}$ is computable and equality in finite sets is decidable.

Corollary 4.6. *If* $t, t' \in \mathsf{Tm}\ \sigma$, *then* $t =_{\beta\eta} t'$ *iff* $\mathsf{quote}^{\sigma}\ [\![t]\!]^{\mathsf{Set}} = \mathsf{quote}^{\sigma}\ [\![t']\!]^{\mathsf{Set}}$.

Proof. Immediate from soundness (Proposition 2.1) and the Main Theorem. □

This corollary shows that $\mathsf{nf}^{\sigma} = \mathsf{quote}^{\sigma} \circ [\![-]\!]^{\mathsf{Set}} : \mathsf{Tm}\ \sigma \to \mathsf{Tm}\ \sigma$ indeed makes sense as normalization function: apart from just delivering, for any given typed closed term, some $\beta\eta$-equal term, it is actually guaranteed to deliver the same term for t, t', if t, t' are $\beta\eta$-equal (this is Church-Rosser for reduction-free normalization).

Note that although we only stated completeness and normalization for typed closed terms above, these trivially extend to all typed terms as opens terms can always be closed up by lambda-abstractions and this preserves $\beta\eta$-equality.

Corollary 4.7. *If* $t, t' \in \mathsf{Tm}\ \sigma$ *and* $C\ t =_{\beta\eta} C\ t'$ *for any* $C : \mathsf{Tm}\ \sigma \to \mathsf{Tm}\ \mathtt{Bool}$, *then* $t =_{\beta\eta} t'$. *Or, contrapositively, and more concretely, if* $t, t' \in \mathsf{Tm}\ (\sigma_1 \to \ldots \to \sigma_n \to \mathtt{Bool})$ *and* $t \neq_{\beta\eta} t'$, *then there exist* $u_1 \in \mathsf{Tm}\ \sigma_1, \ldots u_n \in \mathsf{Tm}\ \sigma_n$ *such that*

$$\mathsf{nf}^{\mathtt{Bool}}\ (\mathtt{App}\ (\ldots\ (\mathtt{App}\ t\ u_1)\ \ldots)\ u_n) \neq \mathsf{nf}^{\mathtt{Bool}}\ (\mathtt{App}\ (\ldots\ (\mathtt{App}\ t'\ u_1)\ \ldots)\ u_n)$$

Proof. This corollary does not follow from the statement of the Main Theorem, but it follows from its proof. □

Corollary 4.8 (Maximal consistency). *If* $t, t' \in \mathsf{Tm}\ \sigma$ *and* $t \neq_{\beta\eta} t'$, *then from the equation* $t = t'$ *as an additional axiom one would derive* $\mathtt{True} = \mathtt{False}$.

Proof. Immediate from the previous corollary. □

5 Proof of the main lemma

We now present the proof of the main lemma which was postponed in the previous section. To keep the proof readable, we write senum, $\mathsf{squestions}$, sfind to emphasize the uses of the set-theoretic semantics instances of enum, $\mathsf{questions}$, find, while the free semantics instances will be written as tenum, $\mathsf{tquestions}$, tfind.

We first give the core of the proof and prove the lemmas this takes thereafter.

Proof (of the Main Lemma). By induction on σ.

- Case $\mathtt{Bool}$: Assume $t\mathsf{R}^{\mathtt{Bool}}b$. Then either $t =_{\beta\eta} \mathtt{True}$ and $b = \mathrm{true}$, in which case we have

  $$t =_{\beta\eta} \mathtt{True} = \mathsf{quote}^{\mathtt{Bool}}\ \mathrm{true} = \mathsf{quote}^{\mathtt{Bool}}\ b$$

  or $t =_{\beta\eta} \mathtt{False}$ and $b = \mathrm{false}$, in which case we have

  $$t =_{\beta\eta} \mathtt{False} = \mathsf{quote}^{\mathtt{Bool}}\ \mathrm{false} = \mathsf{quote}^{\mathtt{Bool}}\ b$$

- Case $\sigma \to \tau$: Assume $t\mathsf{R}^{\sigma\to\tau}f$, for all u, d, $u\mathsf{R}^{\sigma}d$ implies $\mathtt{App}\ t\ u\mathsf{R}^{\tau}f\ d$. We have

  $$\begin{aligned} t &=_{\beta\eta} \mathtt{Lam}^{\sigma}\ \mathtt{x}\ (\mathtt{App}\ t\ (\mathtt{Var}\ \mathtt{x})) \\ &=_{\beta\eta} \text{(by Lemma 5.2 below)} \\ &\quad\ \mathtt{Lam}^{\sigma}\ \mathtt{x}\ (\mathtt{App}\ t\ (\mathsf{tfind}\ [q\ (\mathtt{Var}\ \mathtt{x}) \mid q \leftarrow \mathsf{tquestions}^{\sigma}]\ \mathsf{tenum}^{\sigma})) \\ &= \mathtt{Lam}^{\sigma}\ \mathtt{x}\ (\mathsf{tfind}\ [q\ (\mathtt{Var}\ \mathtt{x}) \mid q \leftarrow \mathsf{tquestions}^{\sigma}]\ (\mathsf{fmap}\ (\mathtt{App}\ t)\ \mathsf{tenum}^{\sigma})) \\ &=_{\beta\eta} \text{(by Sublemma)} \\ &\quad\ \mathtt{Lam}^{\sigma}\ \mathtt{x}\ (\mathsf{tfind}\ [q\ (\mathtt{Var}\ \mathtt{x}) \mid q \leftarrow \mathsf{tquestions}^{\sigma}]\ (\mathsf{fmap}\ (\mathsf{quote}^{\tau} \circ f)\ \mathsf{senum}^{\sigma})) \\ &= \mathsf{quote}^{\sigma\to\tau}\ f \end{aligned}$$

  The Sublemma is:

  $$\mathsf{fmap}\ (\mathtt{App}\ t)\ \mathsf{tenum}^{\sigma}\ (\mathsf{Tree}\ =_{\beta\eta})\ \mathsf{fmap}\ (\mathsf{quote}^{\tau} \circ f)\ \mathsf{senum}^{\sigma}$$

  For proof, we notice that, by Lemma 5.1 (1) below for σ,

  $$\mathsf{tenum}^{\sigma}\ (\mathsf{Tree}\ \mathsf{R}^{\sigma})\ \mathsf{senum}^{\sigma}$$

  Hence, by assumption

  $$\mathsf{fmap}\ (\mathtt{App}\ t)\ \mathsf{tenum}^{\sigma}\ (\mathsf{Tree}\ \mathsf{R}^{\tau})\ \mathsf{fmap}\ f\ \mathsf{senum}^{\sigma}$$

  Hence, by IH of the Lemma for τ,

  $$\mathsf{fmap}\ (\mathtt{App}\ t)\ \mathsf{tenum}^{\sigma}\ (\mathsf{Tree}\ =_{\beta\eta})\ \mathsf{fmap}\ (\mathsf{quote}^{\tau} \circ f)\ \mathsf{senum}^{\sigma}$$

$\square$

The proof above used two lemmas. One is essentially free, but the other is technical.

Lemma 5.1 ("Free" Lemma).

1. tenum^σ $(\mathsf{Tree}\ \mathsf{R}^\sigma)$ senum^σ.
2. $\mathsf{tquestions}^\sigma$ $[\mathsf{R}^\sigma \rightarrow \mathsf{R}^{\mathtt{Bool}}]$ $\mathsf{squestions}^\sigma$.

Proof. Both follow from the Fundamental Theorem (Lemma 4.2), but in more detail, the proof is simultaneous for (1) and (2) by induction on σ.

- Case $\mathtt{Bool}$: Trivial.
- Case $\sigma \rightarrow \tau$: Proof of (1) uses IH (2) for σ and IH (1) for τ; proof of (2) uses IH (1) for σ and IH (2) for τ.

□

Lemma 5.2 (Technical Lemma). *Define a relation* $< \subseteq \mathsf{UTm} \times [\mathsf{UTm} \rightarrow \mathsf{UTm}] \times \mathsf{Tree}\ \mathsf{UTm}$ *by*

$$t < (qs, ts) \textit{ iff } t =_{\beta\eta} \mathsf{tfind}\ [q\ t \mid q \leftarrow qs]\ ts$$

If $t \in \mathsf{Tm}\ \sigma$*, then* $t < (\mathsf{tquestions}^\sigma, \mathsf{tenum}^\sigma)$*, i.e.,*

$$t =_{\beta\eta} \mathsf{tfind}\ [q\ t \mid q \leftarrow \mathsf{tquestions}^\sigma]\ \mathsf{tenum}^\sigma$$

Proof. By induction on σ.

- Case $\mathtt{Bool}$:

$$\begin{aligned} t &=_{\beta\eta} \mathtt{If'}\ t\ \mathtt{True}\ \mathtt{False} \\ &= \mathtt{If'}\ t\ (\mathsf{tfind}\ []\ (\mathsf{Val}\ \mathtt{True}))\ (\mathsf{tfind}\ []\ (\mathsf{Val}\ \mathtt{False})) \\ &= \mathsf{tfind}\ [t]\ (\mathsf{Choice}\ (\mathsf{Val}\ \mathtt{True})\ (\mathsf{Val}\ \mathtt{False})) \\ &= \mathsf{tfind}\ [q\ t \mid q \leftarrow \mathsf{tquestions}^{\mathtt{Bool}}]\ \mathsf{tenum}^{\mathtt{Bool}} \end{aligned}$$

- Case $\sigma \rightarrow \tau$:

$$\begin{aligned} t &=_{\beta\eta} \mathtt{Lam}^\sigma\mathtt{x}\ (\mathtt{App}\ t\ (\mathtt{Var}\ \mathtt{x})) \\ &=_{\beta\eta} \text{(by Sublemma below, IH for } \sigma) \\ &\qquad \mathtt{Lam}^\sigma\mathtt{x}\ (\mathsf{tfind}\ [q\ (\mathtt{App}\ t\ u) \mid u \leftarrow \mathsf{flatten}\ \mathsf{tenum}^\sigma, q \leftarrow \mathsf{tquestions}^\tau] \\ &\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ (\mathtt{Var}\ \mathtt{x}))\ (\mathsf{tmkenum}\ \mathsf{tquestions}^\sigma\ \mathsf{tenum}^\tau))) \\ &= \mathsf{tfind}\ [q\ (\mathtt{App}\ t\ u) \mid u \leftarrow \mathsf{flatten}\ \mathsf{tenum}^\sigma, q \leftarrow \mathsf{tquestions}^\tau] \\ &\qquad (\mathsf{fmap}\ (\lambda g\ \mathtt{Lam}^\sigma\ \mathtt{x}\ (g\ (\mathtt{Var}\ \mathtt{x})))\ (\mathsf{tmkenum}\ \mathsf{tquestions}^\sigma\ \mathsf{tenum}^\tau))) \\ &= \mathsf{tfind}\ [q\ t \mid q \leftarrow \mathsf{tquestions}^{\sigma\rightarrow\tau}]\ \mathsf{tenum}^{\sigma\rightarrow\tau} \end{aligned}$$

The Sublemma is: If $u \in \mathsf{UTm}$ and $u < (qs, us)$, then

$$\begin{aligned} \mathtt{App}\ t\ u =_{\beta\eta} &\ \mathsf{tfind}\ [q\ (\mathtt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ us, q \leftarrow \mathsf{tquestions}^\tau] \\ &\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)) \end{aligned}$$

The proof is by "lockstep" induction on (qs, us).

- Case $([], \mathsf{Val}\ u^\star)$: Assume $u =_{\beta\eta} \mathsf{tfind}\ []\ (\mathsf{Val}\ u^\star)$, i.e., $u =_{\beta\eta} u^\star$. We get

$$\begin{aligned}
&\texttt{App}\ t\ u\\
=_{\beta\eta}\ &\text{(by assumption)}\\
&\quad \texttt{App}\ t\ u^\star\\
=_{\beta\eta}\ &\text{(by IH of the Lemma for } \tau)\\
&\quad \mathsf{tfind}\ [q\ (\texttt{App}\ t\ u^\star) \mid q \leftarrow \mathsf{tquestions}^\tau]\ \mathsf{tenum}^\tau\\
=\ &\quad \mathsf{tfind}\ [q\ (\texttt{App}\ t\ u^\star) \mid q \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{fmap}\ (\lambda v \lambda u'\ v)\ \mathsf{tenum}^\tau))\\
=\ &\quad \mathsf{tfind}\ [q\ (\texttt{App}\ t\ u') \mid u' \leftarrow [u^\star], q \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ []\ \mathsf{tenum}^\tau))
\end{aligned}$$

- Case $(q : qs, \mathsf{Choice}\ l\ r)$:
Assume $u =_{\beta\eta} \mathsf{tfind}\ (q\ u : [q'u \mid q' \leftarrow qs])\ (\mathsf{Choice}\ l\ r)$, i.e., $u =_{\beta\eta}$ $\texttt{If}'\ (q\ u)\ (\mathsf{tfind}\ [q'\ u \mid q' \leftarrow qs]\ l)\ (\mathsf{tfind}\ [q'\ u \mid q' \leftarrow qs]\ r)$. We get

$$\begin{aligned}
&\texttt{App}\ t\ u\\
=_{\beta\eta}\ &\texttt{If}'\ (q\ u)\ (\texttt{App}\ t\ u)\ (\texttt{App}\ t\ u)\\
=_{\beta\eta}\ &\text{(by IH of the Sublemma for } qs/l,\ qs/r)\\
&\quad \texttt{If}'\ (q\ u)\\
&\qquad (\mathsf{tfind}\ [q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ l, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)))\\
&\qquad (\mathsf{tfind}\ [q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ r, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)))\\
=\ &\quad \mathsf{tfind}\ []\ (\mathsf{Val}\ (\texttt{If}'\ (q\ u)\\
&\qquad (\mathsf{tfind}\ [q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ l, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)))\\
&\qquad (\mathsf{tfind}\ [q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ r, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)))))\\
=_{\beta\eta}\ &\text{(by twice Lemma 5.3)}\\
&\quad \mathsf{tfind}\ ([q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ l, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad +\!\!+\ ([q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ r, q' \leftarrow \mathsf{tquestions}^\tau] +\!\!+ []))\\
&\qquad ((\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)) \ggg\!= \lambda v_0\\
&\qquad\quad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau)) \ggg\!= \lambda v_1\\
&\qquad\quad \mathsf{Val}\ (\texttt{If}'\ (q\ u)\ v_0\ v_1))\\
=\ &\quad \mathsf{tfind}\ ([q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ l, q' \leftarrow \mathsf{tquestions}^\tau]\\
&\qquad\quad +\!\!+\ [q'\ (\texttt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ r, q' \leftarrow \mathsf{tquestions}^\tau])\\
&\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\\
&\qquad\quad ((\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau) \ggg\!= \lambda g_0\\
&\qquad\qquad (\mathsf{tmkenum}\ qs\ \mathsf{tenum}^\tau \ggg\!= \lambda g_1\\
&\qquad\qquad \mathsf{Val}\ (\lambda u'\ \texttt{If}'\ (q\ u')\ (g_0\ u')\ (g_1\ u'))))
\end{aligned}$$

$$
\begin{aligned}
&= \ \mathsf{tfind}\ [q'\ (\mathtt{App}\ t\ u') \mid t \leftarrow (\mathsf{flatten}\ l) \mathbin{+\!\!+} (\mathsf{flatten}\ r), q' \leftarrow \mathsf{tquestions}^{\tau}] \\
&\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ (q : qs)\ \mathsf{tenum}^{\tau})) \\
&= \ \mathsf{tfind}\ [q'\ (\mathtt{App}\ t\ u') \mid u' \leftarrow \mathsf{flatten}\ (\mathsf{Choice}\ l\ r), q' \leftarrow \mathsf{tquestions}^{\tau}] \\
&\qquad (\mathsf{fmap}\ (\lambda g\ g\ u)\ (\mathsf{tmkenum}\ (q : qs)\ \mathsf{tenum}^{\tau}))
\end{aligned}
$$

□

In the last lemma, we used another lemma, a substitution lemma for tfind, which is easy to prove.

Lemma 5.3. *Define a relation* $<> \subseteq [\mathsf{UTm}] \times \mathsf{Tree}\ \mathsf{UTm}$ *by "lockstep" induction on the list and the tree by:*

- *if* $t \in \mathsf{UTm}$*, then* $[] <> \mathsf{Val}\ t$*;*
- *if* $as <> l$*,* $as <> r$*, then* $a : as <> \mathsf{Choice}\ l\ r$*.*

Then,

$$\mathsf{tfind}\ (as \mathbin{+\!\!+} bs)\ (ts \gg\!\!= h) =_{\beta\eta} \mathsf{tfind}\ bs\ (h\ (\mathsf{tfind}\ as\ ts))$$

Proof. By induction on (as, ts).

- Case $([], \mathsf{Val}\ t)$:

$$
\begin{aligned}
&\mathsf{tfind}\ ([] \mathbin{+\!\!+} bs)\ ((\mathsf{Val}\ t) \gg\!\!= h) \\
&= \mathsf{tfind}\ bs\ ((\mathsf{Val}\ t) \gg\!\!= h) \\
&= \mathsf{tfind}\ bs\ (h\ t) \\
&= \mathsf{tfind}\ bs\ (h\ (\mathsf{tfind}\ []\ (\mathsf{Val}\ t)))
\end{aligned}
$$

- Case $(a : as, \mathsf{Choice}\ l\ r)$:

$$
\begin{aligned}
&\mathsf{tfind}\ ((a : as) \mathbin{+\!\!+} bs)\ ((\mathsf{Choice}\ l\ r) \gg\!\!= h) \\
&= \ \mathsf{tfind}\ (a : (as \mathbin{+\!\!+} bs))\ ((\mathsf{Choice}\ l\ r) \gg\!\!= h) \\
&= \ \mathtt{If}'\ a\ (\mathsf{tfind}\ (as \mathbin{+\!\!+} bs)\ (l \gg\!\!= h))\ (\mathsf{tfind}\ (as \mathbin{+\!\!+} bs)\ (r \gg\!\!= h)) \\
&= \ (\text{by IH for } as/l,\ as/r) \\
&\qquad \mathtt{If}'\ a\ (\mathsf{tfind}\ bs\ (h\ (\mathsf{tfind}\ as\ l)))\ (\mathsf{tfind}\ bs\ (h\ (\mathsf{tfind}\ as\ r))) \\
&=_{\beta\eta} \mathsf{tfind}\ bs\ (h\ (\mathtt{If}'\ a\ (\mathsf{tfind}\ as\ l)\ (\mathsf{tfind}\ as\ r))) \\
&= \ \mathsf{tfind}\ bs\ (h\ (\mathsf{tfind}\ (a : as)\ (\mathsf{Choice}\ l\ r)))
\end{aligned}
$$

□

We note that, although not necessary for our purposes, the two last lemmas could have also been formulated generically for any semantics instead of just for the free one. To avoid introduction of extra notation, we refrained here from doing this.

6 Discussion and further work

Instead of decision trees we could have used a direct encoding of the graph of a function, we call this the truth-table semantics. However, this approach leads not only too much longer normal forms but also the semantic equality is less efficient. On the other hand it is possible to go further and use Binary Decision Diagrams (BDDs) [6] instead of decision trees. We plan to explore this in further work.

We have argued that $\lambda^{\to 2}$ is the simplest λ-calculus with closed types, however we are confident that the technique described here works also for closed types in $\lambda^{0+1\times\to}$ (the finitary λ calculus). We leave this extension for a journal version of this work.

One can go even further and implement finitary Type Theory, i.e. $\lambda^{012\Sigma\Pi}$. This could provide an interesting base for a type-theoretic hardware description and verification language.

The approach presented here works only for calculi without typevariables. It remains open to see whether this approach can be merged with the the standard techniques for NBE for systems with type variables, leading to an alternative proof of completeness and maybe even finite completeness for the calculi discussed above.

References

1. T. Altenkirch, P. Dybjer, M. Hofmann, and P. Scott. Normalization by evaluation for typed lambda calculus with coproducts. In *Proc. of 16th Annual IEEE Symposium on Logic in Computer Science*, pages 303–310. IEEE CS Press, 2001.
2. T. Altenkirch, M. Hofmann, and T. Streicher. Categorical reconstruction of a reduction free normalization proof. In D. Pitt, D. E. Rydeheard, and P. Johnstone, editors, *Category Theory and Computer Science*, number 953 in LNCS, pages 182–199, Springer-Verlag, 1995.
3. T. Altenkirch, M. Hofmann, and T. Streicher. Reduction-free normalisation for a polymorphic system. In *Proc. of 11th Annual IEEE Symposium on Logic in Computer Science*, pages 98–106. IEEE CS Press, 1996.
4. T. Altenkirch, M. Hofmann, and T. Streicher. Reduction-free normalisation for system *F*. Unpublished, available on WWW at `http://www.cs.nott.ac.uk/~txa/publ/f97.pdf`, 1997.
5. U. Berger and H. Schwichtenberg. An inverse of the evaluation functional for typed λ-calculus. In *Proc. of 6th Annual IEEE Symposium on Logic in Computer Science*, pages 202–211. IEEE CS Press, 1991.
6. R. E. Bryant. Graph-based algorithms for boolean function manipulation. *IEEE Trans. on Computers*, 35(8):677–691, 1986.
7. D. J. Dougherty and R. Subrahmanyam. Equality between functionals in the presence of coproducts. *Information and Computation*, 157(1–2):52–83, 2000.
8. P. Dybjer and A. Filinski. Normalization and partial evaluation. In G. Barthe et al., editors, *Applied Semantics*, number 2395 in LNCS, pages 137–192. Springer-Verlag, 2002.
9. N. Ghani. *Adjoint Rewriting*. PhD thesis, LFCS, Univ. of Edinburgh, 1995.

10. N. Ghani. $\beta\eta$-equality for coproducts. In M. Dezani-Ciancaglini and G. D. Plotkin, editors, *Typed Lambda Calculus and Applications*, number 902 in LNCS, pages 171–185. Springer-Verlag, 1995.
11. G. D. Plotkin. Lambda-definability and logical relations. Memorandum SAI-RM-4, Lab. for Artif. Intell., Univ. of Edinburgh, Oct. 1973.