JFP **19** (3 & 4): 311–333, 2009. © 2009 Cambridge University Press
doi:10.1017/S0956796809007278 Printed in the United Kingdom

Big-step normalisation

THORSTEN ALTENKIRCH
School of Computer Science, University of Nottingham,
Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK
(*e-mail:* txa@cs.nott.ac.uk)

JAMES CHAPMAN
Institute of Cybernetics at Tallinn University of Technology,
Akadeemia tee 21, EE-12618 Tallinn, Estonia
(*e-mail:* james@cs.ioc.ee)

Abstract

Traditionally, decidability of conversion for typed λ-calculi is established by showing that small-step reduction is confluent and strongly normalising. Here we investigate an alternative approach employing a recursively defined normalisation function which we show to be terminating and which reflects and preserves conversion. We apply our approach to the simply typed λ-calculus with explicit substitutions and $\beta\eta$-equality, a system which is not strongly normalising. We also show how the construction can be extended to system T with the usual β-rules for the recursion combinator. Our approach is practical, since it does verify an actual implementation of normalisation which, unlike normalisation by evaluation, is first order. An important feature of our approach is that we are using logical relations to establish equational soundness (identity of normal forms reflects the equational theory), instead of the usual syntactic reasoning using the Church–Rosser property of a term rewriting system.

1 Introduction

Traditionally, decidability of conversion for typed λ-calculi is established by showing that small-step reduction is confluent and strongly normalising; e.g. see Girard *et al.* (1989), where this approach is applied to the simply typed λ-calculus, system F and system T. In fact, decidability is not the only corollary of strong normalisation; we can reason using the structure of normal forms and show for example that certain types are not inhabited.

The small-step approach does not extend easily to stronger conversion relations, e.g. η-conversion; η-reduction preserves strong normalisation, but η-expansion obviously does not. On the other hand η-expansion is preferable because normal terms are in constructor form (i.e. λ-abstractions). This issue can by addressed by careful modification of the reduction relation (Jay & Ghani 1995). A more serious issue arises when introducing substitution as an explicit operation (Abadi *et al.* 1990) – this is better because it treats substitution in the same way as other operators such as application. It was hoped that the small-step semantics for substitution would mix well with β-reduction – this hope was dashed by Melliès's (1995) observation that $\sigma\beta$-reduction is not strongly normalising.

All these issues can be addressed by ingenious modifications of the small-step semantics. However, it is doubtful that anybody would actually want to implement a normalisation function by laboriously applying one-step reductions to a term. This criticism applies only to small-step term rewriting; clearly it is computationally sensible to model the computation of a normal form by performing small steps of an abstract machine.

We observe that normalisation can be expressed by the following specification: We introduce a notion of normal forms $n : \mathsf{Nf}\,\Gamma\,\sigma$ indexed over context Γ and type σ which can be embedded back into terms $\ulcorner n \urcorner : \mathsf{Tm}\,\Gamma\,\sigma$. We assume that we can realise a function **nf** which for any $t : \mathsf{Tm}\,\Gamma\,\sigma$ calculates a normal form $\mathbf{nf}\,t : \mathsf{Nf}\,\Gamma\,\sigma$ such that the following properties hold:

(a) Soundness. Normalisation takes convertible terms to identical normal forms:

$$\frac{t \simeq_{\beta\eta\sigma} t'}{\mathbf{nf}\,t = \mathbf{nf}\,t'}$$

(b) Completeness. Terms are convertible to their normal forms:

$$t \simeq_{\beta\eta\sigma} \ulcorner \mathbf{nf}\,t \urcorner$$

Our terminology here is motivated by the view that the normal forms form a syntactic model of the calculus.

As a consequence we obtain that convertibility corresponds to having the same normal form:

$$t \simeq_{\beta\eta\sigma} u \iff \mathbf{nf}\,t = \mathbf{nf}\,u$$

Since the equality of normal forms is obviously decidable, we have that conversion is decidable. Additionally, we would like that the notion of normal form contains no redundant elements – and hence we can establish additional properties by induction over the structure of normal forms. We can capture this property by additionally demanding the next property.

(c) Stability. Normalisation is stable on normal forms:

$$\frac{n : \mathsf{Nf}\,\Gamma\,\sigma}{\mathbf{nf}\,\ulcorner n \urcorner = n}$$

Strong normalisation gives rise to one way of implementing this specification. An alternative is *normalisation by evaluation* (Berger & Schwichtenberg 1991; Coquand & Dybjer 1997). Normalisation by evaluation exploits a complete model construction in which the evaluation function can be inverted. The composition of the evaluation function and its inverse gives rise to a normalisation function. This function can be executed, since all the steps take place in a constructive metatheory. normalisation by evaluation overcomes many of the shortcomings of the small-step approach. Indeed, decidability for strong equality of λ-calculus with coproducts was shown using normalisation by evaluation (Altenkirch *et al.* 2001; Balat 2002). More recently, Lindley (2007) showed that a small-step semantics for this calculus is weakly normalising, simplifying earlier work by Ghani (1995). Moreover, normalisation by evaluation is practical; it has been used in the actual implementation of Schwichtenberg's Minlog system and the Epigram system (Chapman *et al.* 2007).

Here we investigate yet another alternative: big-step normalisation. This is in some way the most naive approach to normalisation: we use an environment machine, implemented as a functional program, to reduce programs to values and apply this method recursively to quote values as normal forms. We apply this approach here to $\lambda^{\beta\eta\sigma}$, simply typed λ-calculus with explicit substitutions, a calculus which is difficult to capture using small-step reduction. Unlike normalisation by evaluation our approach is first order, we do not need higher-order functions in any essential way to implement normalisation, while normalisation by evaluation assumes that we already have a means to evaluate higher-order programs, i.e. λ-terms.

Big-step normalisation shares the logical structure of small-step normalisation. The normalisation function is specified as an inductive relation using only first-order means. This relation is executable; indeed it is derived from a recursive functional program and then shown to be terminating. Unlike normalisation by evaluation there is a strict separation between the first-order structure of the program and the higher-order reasoning needed to establish termination. Note that we do not attempt to give a normalisation argument which can be formalised in first-order arithmetic, such as the one given by David (2001). Indeed, we show in Section 7 that our construction easily extends to system T, whose normalisation proof certainly cannot be formalised in first-order arithmetic.

1.1 Related work

Big-step semantics has been used in the metatheory of typed λ-calculi by a number of people. Here are some examples:

- Levy (2001) used Tait's method to show normalisation for the big-step semantics of a simple programming language.
- Watkins *et al.* (2004) used *hereditary substitutions*, which have a structure similar to big-step normalisation, to show normalisation of a logical framework.
- T. Coquand used a variant of big-step reduction to normalise terms in 'type theory' (Coquand 1991). However, he exploited a model of the untyped λ-calculus to implement normalisation, which is not necessary in our approach.
- In our conference paper we have developed big-step normalisation for a combinatory version of system T (Altenkirch & Chapman 2006).

Our work is closely related to C. Coquand's formalisation of another variant of simply typed λ-calculus with substitutions (Coquand 2002) – the main difference from the present work is that she uses normalisation by evaluation.

1.2 Type theory as a metalanguage

We use type Ttheory (Martin-Löf 1984; Nordström *et al.* 1990; Hofmann 1997) as a metalanguage; hence when we define a function we can also run it as a functional program. However, since we do not exploit propositions as types in any essential way, our development can be understood as taking place in naive set theory.

Our notation is very much inspired by the Epigram system (McBride & McKinna 2004; McBride 2005a), which we have used together with Agda (Norell 2007a,b) for a formalisation of the material presented here (Chapman 2007).

We use $\star$ for the type of small types (or sets) and Prop as the type of propositions. We will not use proofs of propositions to make choices; i.e. we assume a proof-irrelevant universe of propositions. We present inductively defined families, types and predicates by giving the constructors in a natural deduction style, inspired by the syntax of the Epigram system. We construct functions and proofs by structural recursion over inductive definitions which, using the tactics implemented in Epigram, are reducible to basic type theory, using only standard combinators.

As in Epigram and other implementations of type theory we hide arguments and types which can be inferred from the context to make the code more readable. If we want to make implicit arguments explicit we put them in subscript position. As an example consider our presentation of an inductive definition of the set of natural numbers

$$\frac{}{\mathsf{Nat}\ :\ \star}\quad \text{where}\quad \frac{}{\mathsf{zero}\ :\ \mathsf{Nat}}\quad \frac{n\ :\ \mathsf{Nat}}{\mathsf{suc}\, n\ :\ \mathsf{Nat}}$$

and the family of finite types

$$\frac{n\ :\ \mathsf{Nat}}{\mathsf{Fin}\, n\ :\ \star}\quad \text{where}\quad \frac{}{\mathsf{fzero}\ :\ \mathsf{Fin}\,(\mathsf{suc}\, n)}\quad \frac{i\ :\ \mathsf{Fin}\, n}{\mathsf{fsuc}\, i\ :\ \mathsf{Fin}\,(\mathsf{suc}\, n)}$$

Note that we omit the declaration of $n\ :\ \mathsf{Nat}$ as an implicit argument to the constructors fzero and fsuc because it can be automatically inferred by the system. More details and examples can be found in the Epigram tutorial (McBride 2005b).

1.3 Overview of the paper

We introduce a simply typed λ-calculus with explicit substitution and $\beta\eta$-equality in Section 2. We then implement a recursive normalisation function in partial type theory in Section 3. Using a technique introduced by Bove and Capretta (2001) we give a relational presentation, i.e. a *big-step reduction* relation of the partial functions in total type theory to be able to characterize the graph of our normalisation function in Section 4. Using a variant of strong computability[1] (Tait 1967), incorporating Kripke logical predicates, we then show that our partial normalisation function terminates and returns a result convertible to the input in Section 5. It remains to show soundness; we do this using Kripke logical relations in Section 6. We show that this approach is easily extensible to system T with β-rules for the recursion combinator in Section 7. We finish with general observations about our approach and sketch future work (Section 8).

[1] Also called strong reducibility. The strength here refers to the necessary strengthening of the induction hypothesis and to strong normalisation.

2 Simply typed λ-calculus with explicit substitutions

We present here the simply typed λ-calculus with explicit substitutions, much in the spirit of the λ^σ-calculus (Abadi *et al.* 1990). This approach avoids the special status of substitution which traditionally, unlike other term formers, is defined by recursion over the syntax. In our presentation, substitution is a term former like any other, with a set of equationally specified properties. We also diverge from the conventional strategy of defining pre-terms first and then introducing a typing judgment. Instead we directly present the family of well-typed terms as an inductively defined family. We are, after all, only interested in the well-typed terms.

2.1 Syntax

The inductive definition of the set of types $\mathsf{Ty} : \star$ with one base type and contexts $\mathsf{Con} : \star$ as backward-written lists of types are straightforward:

$$\frac{}{\bullet : \mathsf{Ty}} \qquad \frac{\sigma : \mathsf{Ty} \quad \tau : \mathsf{Ty}}{(\sigma \to \tau) : \mathsf{Ty}} \qquad\qquad \frac{}{\varepsilon : \mathsf{Con}} \qquad \frac{\Gamma : \mathsf{Con} \quad \sigma : \mathsf{Ty}}{(\Gamma ; \sigma) : \mathsf{Con}}$$

We define inductive families of well-typed terms and substitutions mutually:

$$\frac{\Gamma : \mathsf{Con} \quad \sigma : \mathsf{Ty}}{\mathsf{Tm}\,\Gamma\,\sigma : \star} \qquad \frac{\Gamma, \Delta : \mathsf{Con}}{\mathsf{Subst}\,\Gamma\,\Delta : \star}$$

The syntax of terms uses categorical combinators which subsume variables. There is a term $\varnothing$ which refers to the last variable in the context, and $t\,[\vec{t}\,]$ is the application of an explicit substitution to a term. Variables other than the last can be constructed by combining $\varnothing$ with weakening substitutions $\uparrow_\sigma$, which corresponds to $+1$ in a de Bruijn representation:

$$\frac{}{\varnothing : \mathsf{Tm}\,(\Gamma ; \sigma)\,\sigma} \qquad \frac{t : \mathsf{Tm}\,\Delta\,\sigma \quad \vec{t} : \mathsf{Subst}\,\Gamma\,\Delta}{t\,[\vec{t}\,] : \mathsf{Tm}\,\Gamma\,\sigma}$$

$$\frac{t : \mathsf{Tm}\,(\Gamma ; \sigma)\,\tau}{\lambda_\sigma t : \mathsf{Tm}\,\Gamma\,(\sigma \to \tau)} \qquad \frac{t : \mathsf{Tm}\,\Gamma\,(\sigma \to \tau) \quad u : \mathsf{Tm}\,\Gamma\,\sigma}{t\,u : \mathsf{Tm}\,\Gamma\,\tau}$$

Our syntax for substitutions uses the standard categorical combinators: id_Γ the identity substitution, $\vec{t} \circ \vec{u}$ composition of substitutions, $\vec{t} ; t$ extension of a substitution and $\uparrow_\sigma$ weakening or projection:

$$\frac{}{\mathsf{id}_\Gamma : \mathsf{Subst}\,\Gamma\,\Gamma} \qquad \frac{\vec{t} : \mathsf{Subst}\,\Gamma\,\Delta \quad \vec{u} : \mathsf{Subst}\,B\,\Gamma}{\vec{t} \circ \vec{u} : \mathsf{Subst}\,B\,\Delta}$$

$$\frac{\vec{t} : \mathsf{Subst}\,\Gamma\,\Delta \quad t : \mathsf{Tm}\,\Gamma\,\sigma}{(\vec{t} ; t) : \mathsf{Subst}\,\Gamma\,(\Delta ; \sigma)} \qquad \frac{}{\uparrow_\sigma : \mathsf{Subst}\,(\Gamma ; \sigma)\,\Gamma}$$

As a special case we can derive substitution of the last variable by a term: given $t : \mathsf{Tm}\,(\Gamma ; \sigma)\,\tau$ and $u : \mathsf{Tm}\,\Gamma\,\sigma$, we obtain t with $\varnothing$ substituted by u as $t\,[u] = t\,[\mathsf{id}_\Gamma ; u] : \mathsf{Tm}\,\Gamma\,\tau$.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

As an example we represent the λ-term implementing the S combinator (given $\sigma, \tau, \rho \;:\; \mathsf{Ty}$)

$$\vdash \lambda f.\lambda g.\lambda x.f\; x\,(g\; x) \;:\; (\sigma{\to}\tau{\to}\rho){\to}(\sigma{\to}\tau){\to}\sigma{\to}\tau$$

as

$$\lambda(\lambda(\lambda((\varnothing[\uparrow_{\sigma\to\tau}][\uparrow_\sigma])\,\varnothing\,((\varnothing[\uparrow_\sigma])\,\varnothing)))) \;:\; \mathsf{Tm}\,\varepsilon\,((\sigma{\to}\tau{\to}\rho){\to}(\sigma{\to}\tau){\to}\sigma{\to}\tau).$$

2.2 Equational theory

We define weak conversion $\simeq_{\mathsf{w}\sigma}$ and strong (or $\beta\eta$) conversion for terms and substitutions $\simeq_{\beta\eta\sigma}$. Each of these is defined mutually for terms and substitutions:

$$\frac{t, u \;:\; \mathsf{Tm}\,\Gamma\,\sigma}{\begin{array}{c} t \simeq_{\beta\eta\sigma} u \;:\; \mathsf{Prop} \\ t \simeq_{\mathsf{w}\sigma} u \;:\; \mathsf{Prop}\end{array}} \qquad \frac{\vec{t}, \vec{u} \;:\; \mathsf{Subst}\,\Gamma\,\Sigma}{\begin{array}{c} \vec{t} \simeq_{\beta\eta\sigma} \vec{u} \;:\; \mathsf{Prop} \\ \vec{t} \simeq_{\mathsf{w}\sigma} \vec{u} \;:\; \mathsf{Prop}\end{array}}.$$

Weak conversion corresponds to combinatorial equality and excludes the η-rule and the ξ-rule (the congruence rule for λ). Since the axioms and rules defining weak equality are simply a subset of the rules defining $\beta\eta$-equality we adopt the convention that we write $\simeq$ if the rule applies to both but use $\simeq_{\beta\eta\sigma}$ if it only applies to strong equality. Intuitively, the weak equality captures the fragment in which we never go under a λ.

2.2.1 Conversion for terms

First we show the rules for how terms interact with substitutions:

$$\begin{array}{llll}
\varnothing[\vec{t}\,;t] & \simeq & t & \mathsf{proj} \\
t\,[\mathsf{id}_\Gamma] & \simeq & t & \mathsf{id} \\
t\,[\vec{t}\circ\vec{u}] & \simeq & t\,[\vec{t}\,][\vec{u}] & \mathsf{comp} \\
(\lambda_\sigma t)[\vec{t}\,] & \simeq_{\beta\eta\sigma} & \lambda_\sigma(t\,[\vec{t}\circ\uparrow_\sigma;\varnothing]) & \mathsf{lam} \\
(t\;u)[\vec{t}\,] & \simeq & t\,[\vec{t}\,]\,(u\,[\vec{t}\,]) & \mathsf{capp}
\end{array}$$

Note that in the weak theory we are not allowed to push a substitution under λ because otherwise we could derive ξ from the other congruences. The β equation is replaced by $\beta\sigma$:

$$(\lambda_\sigma t)[\vec{u}]\,u \quad\simeq\quad t\,[\vec{u}\,;u] \quad\; \beta\sigma$$

As we will show below this $\beta\sigma$ is equivalent to β in the strong theory, however, it is stronger in the weak theory. We also add the η-rule for the $\beta\eta$-equality:

$$t \quad\simeq_{\beta\eta\sigma}\quad \lambda_\sigma(t\,[\uparrow_\sigma]\,\varnothing) \quad\; \eta$$

In addition we have refl, sym and trans and all congruence rules for terms except for ξ which only holds for the strong equality:

$$\frac{\begin{array}{c} t, u \in \mathsf{Tm}\,(\Gamma\,;\sigma)\,\tau \\ t \simeq_{\beta\eta\sigma} u \end{array}}{\lambda_\sigma t \simeq_{\beta\eta\sigma} \lambda_\sigma u} \quad \xi$$

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

2.2.2 Conversion for substitutions

The conversion for substitutions is given by the usual laws defining a category

$$\begin{array}{lcll} (\vec{t} \circ \vec{u}) \circ \vec{v} & \simeq & \vec{t} \circ (\vec{u} \circ \vec{v}) & \mathsf{assoc} \\ \mathsf{id}_\Gamma \circ \vec{u} & \simeq & \vec{u} & \mathsf{idl} \\ \vec{u} \circ \mathsf{id}_\Gamma & \simeq & \vec{u} & \mathsf{idr} \end{array}$$

and the following laws which formalise the existence of finite products:

$$\begin{array}{lcll} \uparrow_\sigma \circ (\vec{u}\,; u) & \simeq & \vec{u} & \mathsf{wk} \\ (\vec{t}\,; t) \circ \vec{u} & \simeq & (\vec{t} \circ \vec{u}); t[\vec{u}] & \mathsf{cons} \\ \mathsf{id}_{\Gamma\,;\sigma} & \simeq & (\mathsf{id}_\Gamma \circ \uparrow_\sigma); \varnothing & \mathsf{sid} \end{array}$$

The choice of laws is motivated by the need to show soundness and completeness of our normalisation algorithm. In addition we have refl, sym and trans and all congruence rules for substitutions.

2.2.3 The β- and $\beta\sigma$-equations

We note that the usual β-rule

$$(\lambda_\sigma t)\, u \quad \simeq_{\beta\eta\sigma} \quad t[u] \quad \beta$$

is too weak for the weak equality because we cannot reduce a λ-term with a delayed substitution. However, as announced earlier, we have the following:

Proposition 1
The rules β and $\beta\sigma$ are inter-derivable in the strong theory.

Proof
First of all it is easy to see that $\beta\sigma$ implies β: $(\lambda_\sigma t)\, u \simeq (\lambda_\sigma t)[\mathsf{id}_\Gamma]\, u$ using id and $(\lambda_\sigma t)[\mathsf{id}_\Gamma]\, u \simeq t[\mathsf{id}_\Gamma\,; u]$ using $\beta\sigma$. Second we show that the other direction is provable:

$$\begin{array}{cll} & (\lambda_\sigma t)[\vec{u}]\, u & \\ \simeq & (\lambda_\sigma t[\vec{u} \circ \uparrow_\sigma; \varnothing])\, u & \{\mathsf{lam}\} \\ \simeq & t[\vec{u} \circ \uparrow_\sigma; \varnothing][\mathsf{id}_\Gamma\,; u] & \{\beta\} \\ \simeq & t[(\vec{u} \circ \uparrow_\sigma; \varnothing) \circ (\mathsf{id}_\Gamma\,; u)] & \{\mathsf{comp}\} \\ \simeq & t[(\vec{u} \circ \uparrow_\sigma) \circ (\mathsf{id}_\Gamma\,; u); \varnothing[\mathsf{id}_\Gamma\,; u]] & \{\mathsf{cons}\} \\ \simeq & t[(\vec{u} \circ \uparrow_\sigma) \circ (\mathsf{id}_\Gamma\,; u); u] & \{\mathsf{proj}\} \\ \simeq & t[\vec{u} \circ (\uparrow_\sigma \circ (\mathsf{id}_\Gamma\,; u)); u] & \{\mathsf{assoc}\} \\ \simeq & t[\vec{u} \circ \mathsf{id}_\Gamma\,; u] & \{\mathsf{wk}\} \\ \simeq & t[\vec{u}\,; u] & \{\mathsf{idr}\} \end{array}$$

□

3 Recursive normalisation

We start with a recursive implementation of normalisation and will later verify that it is terminating, sound and complete. However, since our implementation uses

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

dependent types the function is automatically type correct – we will never have to verify a property like subject reduction.

We first sketch the top-level structure of the algorithm before going into the details of the implementation. We take the liberty of referring to values, environments and normal forms before defining them. Normalisation proceeds in two steps: we define a simple evaluator, basically an environment machine, which produces values, or weak normal forms:

$$\frac{t \;:\; \mathsf{Tm}\,\Delta\,\sigma \qquad \vec{v} \;:\; \mathsf{Env}\,\Gamma\,\Delta}{\mathbf{eval}\,t\,\vec{v} \;:\; \mathsf{Val}\,\Gamma\,\sigma}$$

The evaluator is parameterised by an environment, which assigns to every free variable a value of the appropriate type and returns a value. To complete normalisation we define a quoting function which returns a normal form by recursively evaluating the term

$$\frac{v \;:\; \mathsf{Val}\,\Gamma\,\sigma}{\mathbf{quote}\,v \;:\; \mathsf{Nf}\,\Gamma\,\sigma}$$

Hence we obtain **nf** by combining **eval** and **quote**:

$$\frac{t \;:\; \mathsf{Tm}\,\Gamma\,\sigma}{\mathbf{nf}\,t \;:\; \mathsf{Nf}\,\Gamma\,\sigma} \quad \text{where} \quad \mathbf{nf}\,t \Rightarrow \mathbf{quote}\,(\mathbf{eval}\;t\;\mathbf{id}_\Gamma)$$

Here $\mathbf{id}_\Gamma \;:\; \mathsf{Env}\,\Gamma\,\Gamma$ is the identity environment, which we define later by recursion over Γ.

Having completed our sketch we start to fill in the details. We begin with the definition of de Bruijn variables – the variable $\varnothing$ is the variable at the end (right-hand side) of the context; $\mathsf{vsuc}\,\varnothing$ is the next one; and so on

$$\frac{\Gamma \;:\; \mathsf{Con} \qquad \sigma \;:\; \mathsf{Ty}}{\mathsf{Var}\,\Gamma\,\sigma \;:\; \star} \quad \text{where} \quad \frac{}{\varnothing \;:\; \mathsf{Var}\,(\Gamma\,;\sigma)\,\sigma} \quad \frac{x \;:\; \mathsf{Var}\,\Gamma\,\sigma}{\mathsf{vsuc}_\tau\,x \;:\; \mathsf{Var}\,(\Gamma\,;\tau)\,\sigma}$$

We define a type of neutral values, representing computations which are stuck due to the presence of variables in a key position. Since we need neutral values and neutral normal forms we parameterise the definition by an abstract type of values:

$$\frac{T \;:\; \mathsf{Con} \to \mathsf{Ty} \to \star \qquad \Gamma \;:\; \mathsf{Con} \qquad \sigma \;:\; \mathsf{Ty}}{\mathsf{Ne}_T\,\Gamma\,\sigma \;:\; \star} \quad \text{where}$$

$$\frac{x \;:\; \mathsf{Var}\,\Gamma\,\sigma}{x \;:\; \mathsf{Ne}_T\,\Gamma\,\sigma} \quad \frac{f \;:\; \mathsf{Ne}_T\,\Gamma\,(\sigma\to\tau) \qquad a \;:\; T\,\Gamma\,\sigma}{f\,a \;:\; \mathsf{Ne}_T\,\Gamma\,\tau}$$

Now a value is either a λ-closure or a neutral value. We also define the type of environments, since it has to be defined mutually with the type of values:

$$\frac{\Gamma : \mathsf{Con} \quad \sigma : \mathsf{Ty}}{\mathsf{Val}\,\Gamma\,\sigma : \star} \quad \frac{\Gamma, \Delta : \mathsf{Con}}{\mathsf{Env}\,\Gamma\,\Delta : \star} \quad \text{where}$$

$$\frac{t : \mathsf{Tm}\,(\Delta;\sigma)\,\tau \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta}{\lambda_\sigma t\,[\vec{v}] : \mathsf{Val}\,\Gamma\,(\sigma\to\tau)} \quad \frac{n : \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma}{n : \mathsf{Val}\,\Gamma\,\sigma}$$

$$\frac{}{\varepsilon : \mathsf{Env}\,\Gamma\,\varepsilon} \quad \frac{v : \mathsf{Val}\,\Gamma\,\sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta}{(\vec{v};v) : \mathsf{Env}\,\Gamma\,(\Delta;\sigma)}$$

We introduce a family of (overloaded) embedding operations $\ulcorner\_\urcorner$ from $\mathsf{Var}\,\Gamma\,\sigma$, $\mathsf{Val}\,\Gamma\,\sigma$ and $\mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma$ into $\mathsf{Tm}\,\Gamma\,\sigma$ and from $\mathsf{Env}\,\Gamma\,\Delta$ into $\mathsf{Subst}\,\Gamma\,\Delta$. We are ready to define evaluation which has to be defined mutually with evaluation of substitutions and applications of values:

$$\frac{t : \mathsf{Tm}\,\Delta\,\sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta}{\mathbf{eval}\,t\,\vec{v} : \mathsf{Val}\,\Gamma\,\sigma} \quad \frac{\vec{t} : \mathsf{Subst}\,\Delta\,\Sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta}{\vec{\mathbf{eval}}\,\vec{t}\,\vec{v} : \mathsf{Env}\,\Gamma\,\Sigma}$$

$$\frac{f : \mathsf{Val}\,\Gamma\,(\sigma\to\tau) \quad a : \mathsf{Val}\,\Gamma\,\sigma}{f\,@\,a : \mathsf{Val}\,\Gamma\,\tau}$$

We define **eval**, an environment-based evaluator:

$$\begin{array}{llcll}
\mathbf{eval} & \varnothing & (\vec{v};v) & \Rightarrow & v \\
\mathbf{eval} & t\,[\vec{t}] & \vec{v} & \Rightarrow & \mathbf{eval}\,t\,(\vec{\mathbf{eval}}\,\vec{t}\,\vec{v}) \\
\mathbf{eval} & \lambda t & \vec{v} & \Rightarrow & \lambda t\,[\vec{v}] \\
\mathbf{eval} & t\,u & \vec{v} & \Rightarrow & (\mathbf{eval}\,t\,\vec{v})\,@\,(\mathbf{eval}\,u\,\vec{v})
\end{array}$$

We also have to evaluate substitutions:

$$\begin{array}{llcll}
\vec{\mathbf{eval}} & \mathsf{id} & \vec{v} & \Rightarrow & \vec{v} \\
\vec{\mathbf{eval}} & \vec{t}\circ\vec{u} & \vec{v} & \Rightarrow & \vec{\mathbf{eval}}\,\vec{t}\,(\vec{\mathbf{eval}}\,\vec{u}\,\vec{v}) \\
\vec{\mathbf{eval}} & (\vec{t};t) & \vec{v} & \Rightarrow & (\vec{\mathbf{eval}}\,\vec{t}\,\vec{v});(\mathbf{eval}\,t\,\vec{v}) \\
\vec{\mathbf{eval}} & \uparrow_\sigma & (\vec{v};v) & \Rightarrow & \vec{v}
\end{array}$$

Application of values recursively calls **eval** on the term with the context extended by the argument, while in the case of a neutral value n, the argument is added to the spine:

$$\begin{array}{lccll}
\lambda t\,[\vec{v}] & @ & a & \Rightarrow & \mathbf{eval}\,t\,(\vec{v};a) \\
n & @ & a & \Rightarrow & n\,a
\end{array}$$

To define full normalisation we first define the so-called η-long β-normal forms, reusing the definition of neutral values:

$$\frac{\Gamma : \mathsf{Con} \quad \sigma : \mathsf{Ty}}{\mathsf{Nf}\,\Gamma\,\sigma : \star} \quad \text{where} \quad \frac{n : \mathsf{Nf}\,(\Gamma;\sigma)\,\tau}{\lambda_\sigma n : \mathsf{Nf}\,\Gamma\,(\sigma\to\tau)} \quad \frac{n : \mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\bullet}{n : \mathsf{Nf}\,\Gamma\,\bullet}$$

Alternatively, β-normal forms can be defined by allowing any type in the last rule, instead of restricting it to base type. We extend the family of embedding operations $\ulcorner\_\urcorner$ with instances for $\mathsf{Nf}\,\Gamma\,\sigma$ and $\mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\sigma$ into $\mathsf{Tm}\,\Gamma\,\sigma$ and also into $\mathsf{Val}\,\Gamma\,\sigma$.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

Weakening is defined by mutual recursion for neutral values, values and environments:

$$\frac{v \;:\; \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma}{v^{+_\tau} \;:\; \mathsf{Ne}_{\mathsf{Val}}\,(\Gamma\,;\tau)\,\sigma} \quad \text{where} \quad \begin{array}{lcl} x^{+_\tau} & \Rightarrow & \mathsf{vsuc}_\tau\, x \\ (n\,v)^{+_\tau} & \Rightarrow & (n^{+_\tau})\,(v^{+_\tau}) \end{array}$$

$$\frac{v \;:\; \mathsf{Val}\,\Gamma\,\sigma}{v^{+_\tau} \;:\; \mathsf{Val}\,(\Gamma\,;\tau)\,\sigma} \quad \text{where} \quad \begin{array}{lcl} n^{+_\tau} & \Rightarrow & n^{+_\tau} \\ (\lambda_\rho t\,[\vec{v}])^{+_\tau} & \Rightarrow & \lambda_\rho t\,[\vec{v}^{+_\tau}] \end{array}$$

$$\frac{v \;:\; \mathsf{Env}\,\Gamma\,\Delta}{v^{+_\tau} \;:\; \mathsf{Env}\,(\Gamma\,;\tau)\,\Delta} \quad \text{where} \quad \begin{array}{lcl} \varepsilon^{+_\tau} & \Rightarrow & \varepsilon \\ (\vec{v}\,;v)^{+_\tau} & \Rightarrow & (\vec{v}^{+_\tau}\,;v^{+_\tau}) \end{array}$$

We can iterate weakenings using contexts: here $+\!\!+$ is concatenation of contexts. We give the instance for values as an example:

$$\frac{\Delta \;:\; \mathsf{Con} \quad v \;:\; \mathsf{Val}\,\Gamma\,\sigma}{v^{+_\Delta} \;:\; \mathsf{Val}\,(\Gamma +\!\!+ \Delta)\,\sigma} \quad \text{where} \quad \begin{array}{lcl} v^{+_\varepsilon} & \Rightarrow & v \\ v^{+_{(\Delta;\sigma)}} & \Rightarrow & (v^{+_\Delta})^{+_\sigma} \end{array}$$

The same principle applies to all weakening operations.

Having defined weakening for values we can finally derive the identity environment, which is used by **nf**, by recursion over Γ:

$$\frac{\Gamma \;:\; \mathsf{Con}}{\mathbf{id}_\Gamma \;:\; \mathsf{Env}\,\Gamma\,\Gamma} \quad \text{where}$$

$$\begin{array}{lcl} \mathbf{id}_\varepsilon & \Rightarrow & \varepsilon \\ \mathbf{id}_{(\Gamma\,;\sigma)} & \Rightarrow & (\mathbf{id}_\Gamma)^{+_\sigma}\,;\varnothing \end{array}$$

We are ready to define **quote** for values simultaneously with $\overline{\mathbf{quote}}$ for neutral values:

$$\frac{v \;:\; \mathsf{Val}\,\Gamma\,\sigma}{\mathbf{quote}_\sigma\, v \;:\; \mathsf{Nf}\,\Gamma\,\sigma} \quad \frac{n \;:\; \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma}{\overline{\mathbf{quote}}\, n \;:\; \mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\sigma} \quad \text{where}$$

$$\begin{array}{lllcl} \mathbf{quote}_\bullet & n & \Rightarrow & \overline{\mathbf{quote}}\, n \\ \mathbf{quote}_{(\sigma\to\tau)} & f & \Rightarrow & \lambda_\sigma \mathbf{quote}_\tau\,(f^{+_\sigma} \mathbin{@} \varnothing) \end{array}$$

$$\begin{array}{lllcl} \overline{\mathbf{quote}} & x & \Rightarrow & x \\ \overline{\mathbf{quote}} & n\,v & \Rightarrow & (\overline{\mathbf{quote}}\, n)\,(\mathbf{quote}\, v) \end{array}$$

Note that we define **quote** by recursion over the type. This can be avoided if we are only interested in β-normal forms. In this case we would define **quote** as follows:

$$\begin{array}{lllcl} \mathbf{quote}^\beta & \lambda_\sigma t\,[\vec{v}] & \Rightarrow & \lambda_\sigma \mathbf{quote}^\beta_\tau\,(\mathbf{eval}\, t\,(\vec{v}^{+_\sigma}\,;\varnothing)) \\ \mathbf{quote}^\beta & n & \Rightarrow & \overline{\mathbf{quote}}^\beta\, n \end{array}$$

While **quote** and $\overline{\mathbf{quote}}$ remind us of reify and reflect (sometimes also called quote and unquote) as they appear in normalisation by evaluation, the precise relationship is less clear: while reify maps semantic values to normal forms and reflect maps neutral terms to semantic values, here both **quote** and $\overline{\mathbf{quote}}$ go basically in the same direction, respectively mapping computational values and neutral values to normal forms and neutral normal forms.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

4 Big-step semantics

The functions defined in the previous section are not structurally recursive, and hence it is not obvious how to implement them in total type theory. To bridge this gap we will exploit a technique pioneered by Bove and Capretta (2001): we inductively define the graph of our function and then show that the graph is total; i.e. for every input there exists an output. We can use this proof to actually run our function without having to employ a choice principle – i.e. we keep the separation of propositions and types.

4.1 The Bove–Capretta technique

As an example consider the following function which is defined using nested recursion:

$$\frac{n \ : \ \mathsf{Nat}}{\mathbf{f}\, n \ : \ \mathsf{Nat}} \quad \text{where} \quad \begin{array}{lll} \mathbf{f} & \mathsf{zero} & \Rightarrow \ \mathsf{zero} \\ \mathbf{f} & (\mathsf{suc}\, n) & \Rightarrow \ \mathbf{f}\,(\mathbf{f}\, n) \end{array}$$

While it is obvious to us that $\mathbf{f}$ is total, it is not obviously structurally recursive. However, we can inductively define the graph of the function as a relation – its big-step semantics:

$$\frac{n, n' \ : \ \mathsf{Nat}}{\mathsf{f}\, n \Downarrow n' \ : \ \mathsf{Prop}} \quad \text{where} \quad \frac{}{\mathsf{fz} \ : \ \mathsf{f}\,\mathsf{zero} \Downarrow \mathsf{zero}} \quad \frac{p \ : \ \mathsf{f}\, n \Downarrow n' \quad p' \ : \ \mathsf{f}\, n' \Downarrow n''}{\mathsf{fs}\, p\, p' \ : \ \mathsf{f}\,(\mathsf{suc}\, n) \Downarrow n''}$$

We adopt the convention that the relation corresponding to the recursive definition of $\mathbf{f} \ : \ \mathsf{Nat} \to \mathsf{Nat}$ is written as $\mathsf{f} - \Downarrow - \ : \ \mathsf{Nat} \to \mathsf{Nat} \to \mathsf{Prop}$. We can now define a structurally recursive version of $\mathbf{f}$ called $\mathbf{f}^{\mathbf{str}}$:

$$\frac{p \ : \ \mathsf{f}\, n \Downarrow m}{\mathbf{f}^{\mathbf{str}}\, n\, p \ : \ \Sigma n' {:} \mathsf{Nat}.\, n' = m} \quad \text{where}$$

$$\begin{array}{llllll} \mathbf{f}^{\mathbf{str}} & \mathsf{zero} & \mathsf{fz} & \Rightarrow & (\mathsf{zero}, \mathsf{refl}) & \\ \mathbf{f}^{\mathbf{str}} & (\mathsf{suc}\, n) & (\mathsf{fs}\, p\, p') & \underline{\text{with}} & \mathbf{f}^{\mathbf{str}}\, n\, p & \\ & & & | & (n', \mathsf{refl}) & \Rightarrow \ \mathbf{f}^{\mathbf{str}}\, n'\, p' \end{array}$$

And once we have established that $\mathsf{f} - \Downarrow -$ is total, we have the following:

Theorem 2

$$\frac{n \ : \ \mathsf{Nat}}{\mathsf{f}\, n \Downarrow \mathsf{zero}}$$

Proof
By induction on n. □

We can now redefine $\mathbf{f}$ as a structurally recursive function[2]:

$$\frac{n \ : \ \mathsf{Nat}}{\mathbf{f}\, n \ : \ \mathsf{Nat}} \quad \text{where} \quad \mathbf{f} \quad n \quad \Rightarrow \quad \mathbf{fst}\,(\mathbf{f}^{\mathbf{str}}\, n\,(\mathbf{theorem2}\, n))$$

[2] When using theorems as proof terms in programs we write **theoremi**, where **i** is the number of the theorem.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

4.2 *Big-step semantics of* nf

We will now apply this technique to the recursive definition of normalisation from the previous section. The big-step semantics is given by the following inductively defined relations:

$$\frac{t : \mathsf{Tm}\,\Delta\,\sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta \quad v : \mathsf{Val}\,\Gamma\,\sigma}{\mathsf{eval}\,t\,\vec{v} \Downarrow v : \mathsf{Prop}}$$

$$\frac{\vec{t} : \mathsf{Subst}\,\Delta\,\Sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta \quad \vec{w} : \mathsf{Env}\,\Gamma\,\Sigma}{\vec{\mathsf{eval}}\,\vec{t}\,\vec{v} \Downarrow \vec{w} : \mathsf{Prop}}$$

$$\frac{f : \mathsf{Val}\,\Gamma\,(\sigma\to\tau) \quad a : \mathsf{Val}\,\Gamma\,\sigma \quad v : \mathsf{Val}\,\Gamma\,\tau}{f \,@\, a \Downarrow v : \mathsf{Prop}}$$

$$\frac{v : \mathsf{Val}\,\Gamma\,\sigma \quad n : \mathsf{Nf}\,\Gamma\,\sigma}{\mathsf{quote}\,v \Downarrow n : \mathsf{Prop}} \qquad \frac{v : \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma \quad n : \mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\sigma}{\overline{\mathsf{quote}}\,v \Downarrow n : \mathsf{Prop}}$$

$$\frac{t : \mathsf{Tm}\,\Gamma\,\sigma \quad n : \mathsf{Nf}\,\Gamma\,\sigma}{\mathsf{nf}\,t \Downarrow n : \mathsf{Prop}}$$

The inductive definition of those relations is straightforward from the recursive definition of the functions in the previous section. To illustrate this we give the constructors for $\mathsf{eval}\,t\,\vec{v} \Downarrow v$:

$$\frac{}{\mathsf{rlam} : \mathsf{eval}\,(\lambda_\sigma t)\,\vec{v} \Downarrow \lambda_\sigma t\,[\vec{v}]} \qquad \frac{}{\mathsf{rvar} : \mathsf{eval}\,\varnothing\,(\vec{v}\,;v) \Downarrow v}$$

$$\frac{p : \vec{\mathsf{eval}}\,\vec{t}\,\vec{v} \Downarrow \vec{v}' \quad q : \mathsf{eval}\,t\,\vec{v}' \Downarrow v}{\mathsf{rsubs}\,p\,q : \mathsf{eval}\,(t\,[\vec{t}])\,\vec{v} \Downarrow v}$$

$$\frac{p : \mathsf{eval}\,t\,\vec{v} \Downarrow f \quad q : \mathsf{eval}\,u\,\vec{v} \Downarrow v \quad r : f \,@\, v \Downarrow w}{\mathsf{rapp}\,p\,q\,r : \mathsf{eval}\,(t\,u)\,\vec{v} \Downarrow w}$$

We can now augment our evaluation algorithm, making it structurally recursive on the big-step relation. To make the induction go through we have to show simultaneously that the functions calculate the specified results. We define structurally recursive functions corresponding to the recursive ones in the previous section by structural recursion over the proofs of termination:

$$\frac{t : \mathsf{Tm}\,\Delta\,\sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta \quad p : \mathsf{eval}\,t\,\vec{v} \Downarrow v}{\mathbf{eval}^{\mathbf{str}}\,t\,\vec{v}\,p : \Sigma v' : \mathsf{Val}\,\Gamma\,\sigma\,.\,v' = v}$$

$$\frac{\vec{t} : \mathsf{Subst}\,\Delta\,\Sigma \quad \vec{v} : \mathsf{Env}\,\Gamma\,\Delta \quad p : \vec{\mathsf{eval}}\,\vec{t}\,\vec{v} \Downarrow \vec{w}}{\vec{\mathbf{eval}}^{\mathbf{str}}\,\vec{t}\,\vec{v}\,p : \Sigma w' : \mathsf{Env}\,\Gamma\,\Sigma\,.\,w' = w}$$

$$\frac{f : \mathsf{Val}\,\Gamma\,(\sigma\to\tau) \quad a : \mathsf{Val}\,\Gamma\,\sigma \quad p : f \,@\, a \Downarrow v}{\mathbf{app}^{\mathbf{str}}\,f\,a\,p : \Sigma v' : \mathsf{Val}\,\Gamma\,\tau.\,v' = v}$$

$$\frac{v : \mathsf{Val}\,\Gamma\,\sigma \quad p : \mathsf{quote}\,v \Downarrow n}{\mathbf{quote}^{\mathbf{str}}\,v\,p : \Sigma n' : \mathsf{Nf}\,\Gamma\,\sigma.\,n' = n} \qquad \frac{m : \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma \quad p : \overline{\mathsf{quote}}\,m \Downarrow n}{\overline{\mathbf{quote}}^{\mathbf{str}}\,m\,p : \Sigma n' : \mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\sigma.\,n' = n}$$

$$\frac{p : \mathsf{nf}\,t \Downarrow n}{\mathbf{nf}^{\mathbf{str}}\,t\,p : \mathsf{Nf}\,\Gamma\,\sigma}$$

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

The definition of the structurally recursive operator proceeds along the lines of our example $\mathbf{f}^{\mathsf{str}}$; e.g. in the case of $\mathbf{eval}^{\mathsf{str}}$ this becomes

$$\begin{array}{lllllllll}
\mathbf{eval}^{\mathsf{str}} & \varnothing & (\vec{v}\,;v) & \mathsf{rvar} & \Rightarrow & (v,\,\mathsf{refl}) & & \\
\mathbf{eval}^{\mathsf{str}} & t\,[\vec{t}\,] & \vec{v} & (\mathsf{rsubs}\,p\,q) & \underline{\mathrm{with}} & \vec{\mathbf{eval}}^{\mathsf{str}}\,\vec{t}\,\vec{v}\,p & & \\
 & & & & \mid & (\vec{v}',\,\mathsf{refl}) & \Rightarrow & \mathbf{eval}^{\mathsf{str}}\,t\,\vec{v}'\,q \\
\mathbf{eval}^{\mathsf{str}} & \lambda t & \vec{v} & \mathsf{rlam} & \Rightarrow & (\lambda t\,[\vec{v}\,],\,\mathsf{refl}) & & \\
\mathbf{eval}^{\mathsf{str}} & t\,u & \vec{v} & (\mathsf{rapp}\,p\,q\,r) & \underline{\mathrm{with}} & \mathbf{eval}^{\mathsf{str}}\,t\,\vec{v}\,p & \mid & \mathbf{eval}^{\mathsf{str}}\,u\,\vec{v}\,q \\
 & & & & \mid & (f,\,\mathsf{refl}) & \mid & (a,\,\mathsf{refl})
\end{array}$$
$$\quad\Rightarrow\ \mathbf{app}^{\mathsf{str}}\,f\;a\;r$$

We are using the with-construct here to allow us to pattern match on an intermediate value – see the relevant literature (McBride & McKinna 2004; Norell 2007b) for further details. The derivation of structurally recursive versions of @, $\vec{\mathbf{eval}}$, **quote** and $\overline{\mathbf{quote}}$ proceeds analogously.

5 Termination and completeness

We use the notion of strong computability to show that our normalisation function terminates and that the result is $\beta\eta$-equivalent to the input. Since we are evaluating under λ, we introduce a Kripke-style extension of computability at higher type,

$$\frac{v\ :\ \mathsf{Val}\,\Gamma\;\sigma}{\mathsf{SCV}_{\Gamma,\sigma}\,v\ :\ \mathsf{Prop}}$$

which is defined by recursion over σ:

$$\frac{\overline{\mathsf{quote}}\,n\ \Downarrow\ m\qquad \ulcorner n\urcorner \simeq_{\beta\eta\sigma} \ulcorner m\urcorner}{\mathsf{SCV}_{\Gamma,\bullet}\,n}$$

$$\frac{\forall\Delta.\forall v:\mathsf{Val}\,(\Gamma +\!\!+ \Delta)\,\sigma\,.\,\mathsf{SCV}\,v \to \exists w.f^{+_\Delta}\ @\ v\ \Downarrow\ w \wedge \ulcorner f^{+_\Delta}\urcorner\ulcorner v\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner w\urcorner \wedge \mathsf{SCV}w}{\mathsf{SCV}_{\Gamma,(\sigma\to\tau)}f}$$

It is straightforward to extend strong computability to environments:

$$\frac{\vec{v}\ :\ \mathsf{Env}\,\Gamma\;\Delta}{\mathsf{SCE}_{\Gamma,\Delta}\,\vec{v}\ :\ \mathsf{Prop}}\quad\text{where}\quad \frac{}{\mathsf{SCE}\,\varepsilon}\qquad \frac{\mathsf{SCE}\,\vec{v}\qquad \mathsf{SCV}\,v}{\mathsf{SCE}\,(\vec{v}\,;v)}$$

We will need that strong computability is closed under weakening:

Lemma 3

$$\frac{\mathsf{SCV}_{\Gamma,\sigma}\,v}{\mathsf{SCV}_{(\Gamma+\!\!+\Delta),\sigma}\,v^{+_\Delta}}\qquad \frac{\mathsf{SCE}_{\Gamma,\Sigma}\,\vec{v}}{\mathsf{SCE}_{(\Gamma+\!\!+\Delta),\Sigma}\,\vec{v}^{+_\Delta}}$$

Proof
By induction over σ and Σ. □

Our main technical lemma is that **quote** terminates for all strongly computable values and that the result is $\beta\eta\sigma$-convertible to the input. Our proof proceeds by induction over the type; to deal with the negative occurrence of types we show at the same time that termination of quote for neutral terms implies strong computability.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

The second component of our proof is also required to show that the identity environment is strongly computable. This structure of establishing two propositions by mutual induction over types is common to conventional strong normalisation proofs and can also be found in the normalisation by evaluation construction.

Lemma 4

$$\frac{\mathsf{SCV}_{\Gamma,\sigma}\, v}{\exists m : \mathsf{Nf}\,\Gamma\,\sigma.\, \mathsf{quote}_{\Gamma,\sigma}\, v \Downarrow m \wedge \ulcorner v \urcorner \simeq_{\beta\eta\sigma} \ulcorner m \urcorner}(q) \qquad \frac{\overline{\mathsf{quote}}_{\Gamma,\sigma}\, n \Downarrow m \quad \ulcorner n \urcorner \simeq_{\beta\eta\sigma} \ulcorner m \urcorner}{\mathsf{SCV}_{\Gamma,\sigma}\, n}(u)$$

Proof
By mutual induction over σ. In the base case both implications follow trivially from the definition of SCV and the observation that all values of base type are neutral. Consider $(\sigma \to \tau)$:

- **(q)** Given $\mathsf{SCV}_{\Gamma,\sigma} f$. Using ind.hyp. (u) for σ we can show that $\mathsf{SCV}_{(\Gamma;\sigma),\sigma}\, \varnothing$, and hence $f^{+_\sigma}\ @\ \varnothing \Downarrow v$ (1), $\ulcorner f^{+_\sigma} \urcorner \varnothing \simeq_{\mathsf{w}\sigma} \ulcorner v \urcorner$ (2) and $\mathsf{SCV}_{\Gamma;\sigma,\tau}\, v$. Now using ind.hyp. (q) for τ we know that $\mathsf{quote}\, v \Downarrow n$ (3) and $\ulcorner v \urcorner \simeq_{\beta\eta\sigma} \ulcorner n \urcorner$ (4). By the definition of the big-step semantics and (1) and (2) we can infer that $\mathsf{quote}_{\Gamma,\sigma} f \Downarrow \lambda_\sigma n$ and using η, ξ and with (2) and (4) that $\ulcorner f \urcorner \simeq_{\beta\eta\sigma} \ulcorner \lambda_\sigma n \urcorner$.
- **(u)** Given $\overline{\mathsf{quote}}_{\Gamma,(\sigma\to\tau)}\, n \Downarrow m$ and $\ulcorner n \urcorner \simeq_{\beta\eta\sigma} \ulcorner m \urcorner$ (1). To show $\mathsf{SCV}_{\Gamma,(\sigma\to\tau)}\, n$, assume as given $\mathsf{SCV}_{(\Gamma +\!\!+ \Delta),\sigma}\, v$. Certainly $n^{+_\Delta}\ @\ v \Downarrow n^{+_\Delta}\, v$, since n is neutral. By ind.hyp. (q) for σ we know that $\mathsf{quote}_{\Gamma,\sigma}\, v \Downarrow u$ and $\ulcorner v \urcorner \simeq_{\beta\eta\sigma} \ulcorner u \urcorner$ (2). Hence, $\overline{\mathsf{quote}}_{\Gamma,\sigma}\, (n^{+_\Delta}\, v) \Downarrow m\, u$ (3), and from (1) and (2) we can infer $\ulcorner n^{+_\Delta} \urcorner \ulcorner v \urcorner \simeq_{\beta\eta\sigma} \ulcorner n^{+_\Delta}\, v \urcorner$ (4). $\mathsf{SCV}_{(\Gamma +\!\!+ \Delta),\tau}\, (n^{+_\Delta}\, v)$ follows from (3) and (4) by ind. hyp. (u) for τ. □

A simple consequence of the second component of the lemma is that variables are strongly computable, and hence the identity environment is strongly computable.

Corollary 5

$$\frac{x\ :\ \mathsf{Var}\,\Gamma\,\sigma}{\mathsf{SCV}\, x}(1) \qquad \frac{\Gamma\ :\ \mathsf{Con}}{\mathsf{SCE}\, \mathbf{id}_\Gamma}(2)$$

Proof

(i) Since $\overline{\mathsf{quote}}_{\Gamma,\sigma}\, x \Downarrow x$, we just have to apply (u) of Lemma 4.

(ii) By induction over Γ using (i) and Lemma 3. □

We prove the fundamental theorem for our notion of strong computability which has to be shown mutually for terms and substitutions:

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

Theorem 6

$$\frac{t \;:\; \mathsf{Tm}\,\Delta\,\sigma \quad \mathsf{SCE}_{\Gamma,\Delta}\,\vec{v}}{\exists v : \mathsf{Val}\,\Gamma\,\sigma.\,\mathsf{eval}\,t\,\vec{v} \Downarrow v \wedge t[\ulcorner\vec{v}\urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner v \urcorner \wedge \mathsf{SCV}\,v}$$

$$\frac{\vec{t} \;:\; \mathsf{Subst}\,\Delta\,\Sigma \quad \mathsf{SCE}_{\Gamma,\Delta}\,\vec{v}}{\exists\,\vec{w} \;:\; \mathsf{Env}\,\Gamma\,\Sigma.\,\vec{\mathsf{eval}}\,\vec{t}\,\vec{v} \Downarrow \vec{w} \wedge \vec{t} \circ \ulcorner\vec{v}\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner\vec{w}\urcorner \wedge \mathsf{SCE}\,\vec{w}}$$

Proof
By induction over $t \;:\; \mathsf{Tm}\,\Delta\,\sigma$ and $\vec{t} \;:\; \mathsf{Subst}\,\Gamma\,\Delta$, using the laws of the weak conversion relation and the definition of the big-step reduction relation. The proof is mostly straightforward adaptation of the fundamental theorem for logical predicates; we just discuss some interesting cases. We assume as given $\mathsf{SCE}_{\Gamma,\Delta}\,\vec{v}$.

- $\lambda_\sigma t$: Since $\lambda_\sigma t$ is a value, we have that $\mathsf{eval}\,(\lambda_\sigma t)\,\vec{v} \Downarrow f$ with $f = \lambda_\sigma t\,[\vec{v}]$, and the equational condition holds trivially. It remains to show that $\mathsf{SCV}_{\Gamma,(\sigma\to\tau)} f$. Assume as given $\mathsf{SCV}_{\Gamma+\!\!+\Delta',\sigma}\,v$; using the induction hypothesis for t and Lemma 3 for $\vec{v}$, we know that there is a w, such that $\mathsf{eval}\,t\,(\vec{v}^{+_{\Delta'}};v) \Downarrow w$, $\ulcorner t\,[\vec{v}^{+_{\Delta'}};v]\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner w \urcorner$ and $\mathsf{SCV}\,w$. By the definition of the big-step relation we have that $f^{+_\Delta}\ @\ v \Downarrow w$, and using $\beta\sigma$ we can show that $\ulcorner f \urcorner \ulcorner v \urcorner \simeq_{\mathsf{w}\sigma} \ulcorner w \urcorner$.
- $(t\,u)$: By ind.hyp for t and u we can infer $\mathsf{eval}\,t\,\vec{v} \Downarrow f$ (1), $t[\ulcorner\vec{v}\urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner f \urcorner$ (2) and $\mathsf{SCV}_{\Gamma,(\sigma\to\tau)} f$ (3); $\mathsf{eval}\,u\,\vec{v} \Downarrow v$ (4), $u[\ulcorner\vec{v}\urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner v \urcorner$ (5) and $\mathsf{SCV}_{\Gamma,\sigma}\,v$ (6). By the definition of SCV using $\Delta = \varepsilon$ and (3) and (6) we get that $f\ @\ v \Downarrow w$ (7), $\ulcorner f \urcorner \ulcorner v \urcorner \simeq_{\mathsf{w}\sigma} \ulcorner w \urcorner$ (8) and $\mathsf{SCV}\,w$. Using the definition of the big-step semantics and (1), (4) and (7) we can show that $\mathsf{eval}\,(t\,u)\,\vec{v} \Downarrow w$ and $(t\,u)[\ulcorner\vec{v}\urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner w \urcorner$ using capp and (2), (5) and (8).
- $(\vec{t};t)$: By ind.hyp. for $\vec{t}$ we get $\vec{\mathsf{eval}}\,\vec{t}\,\vec{v} \Downarrow \vec{w}$ (1), $\vec{t} \circ \ulcorner\vec{v}\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner\vec{w}\urcorner$ (2) and $\mathsf{SCE}\,\vec{w}$ (3). Using the last with the ind.hyp. for t we have that $\vec{\mathsf{eval}}\,t\,\vec{v} \Downarrow v$ (4), $t\,[\ulcorner\vec{v}\urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner v \urcorner$ (5) and $\mathsf{SCV}\,v$ (6). The definition of the big-step reduction and (1) and (4) imply that $\vec{\mathsf{eval}}\,(\vec{t};t)\,\vec{v} \Downarrow (\vec{w};v)$. Using the cons rule and (2) and (5) we can show $(\vec{t};t) \circ \ulcorner\vec{v}\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner \vec{w};v \urcorner$ and $\mathsf{SCE}\,(\vec{w},\,v)$ by (3) and (6). □

Note that the proof never refers to the notion of computability at base type; hence we could have replaced it with any predicate.[3] The fundamental theorem already implies termination and completeness for reduction to values – this corresponds to the result in our workshop paper (Altenkirch & Chapman 2006) which uses combinatory logic corresponding to weak equality of closed terms. Correspondingly we can actually show that the result is weakly equal ($\simeq_{\mathsf{w}\sigma}$) to its input, even though here we only need that it is $\beta\eta\sigma$-equal to its input.

We now can combine the results to infer that **nf** terminates and produces a normal form which is $\beta\eta\sigma$-equivalent to its input.

[3] Including the empty set, indeed there are no closed values of base type.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

Proposition 7

$$\frac{t \;:\; \mathsf{Tm}\,\Delta\,\sigma}{\exists n : \mathsf{Nf}\,\Delta\,\sigma.\,\mathsf{nf}\,t \Downarrow n \;\wedge\; t \simeq_{\beta\eta\sigma} \ulcorner n \urcorner}$$

Proof

By the fundamental Theorem 6 and Corollary 5(ii) we know that $\mathsf{eval}\,t\,\mathbf{id} \Downarrow v$ with $t \simeq_{\mathsf{w}\sigma} t[\ulcorner \mathbf{id} \urcorner] \simeq_{\mathsf{w}\sigma} \ulcorner v \urcorner$ and $\mathsf{SCV}\,v$. Using Lemma 4 we know that $\mathbf{quote}\,v \Downarrow n$ and $\ulcorner v \urcorner \simeq_{\beta\eta\sigma} \ulcorner n \urcorner$; and hence by combining the two steps we obtain the result. □

Since we now know that our functions terminate, we can from now on use the total functions defined in Section 4 together with the termination proofs given in this section:

$$\frac{t \;:\; \mathsf{Tm}\,\Gamma\,\sigma}{\mathbf{nf}\,t \;:\; \mathsf{Nf}\,\Gamma\,\sigma} \quad \text{where} \quad \mathbf{nf}\,t \;\Rightarrow\; \mathbf{nf}^{\mathbf{str}}\,t\,(\mathbf{fst}\,(\mathbf{prop7}\,t))$$

To ease notation we will omit the proof terms altogether but make sure that we only use strongly computable values and environments.

Once we have established that **nf** is terminating it is straightforward to show stability:

Proposition 8 (*stability*)

$$\frac{n \;:\; \mathsf{Nf}\,\Gamma\,\sigma}{\mathbf{nf}\,\ulcorner n \urcorner = n} \qquad \frac{n \;:\; \mathsf{Ne}_{\mathsf{Nf}}\,\Gamma\,\sigma}{\exists\, n' : \mathsf{Ne}_{\mathsf{Val}}\,\Gamma\,\sigma\,.\,\mathbf{eval}\,\ulcorner n \urcorner = n' \;\wedge\; \overline{\mathbf{quote}}\,n' = n}$$

Proof

By simultaneous induction on normal and neutral terms. □

6 Soundness

It remains to be shown that normalisation maps $\beta\eta\sigma$-equivalent terms to equal normal forms. We define a logical relation on values which is preserved by the values obtained from convertible terms and which is mapped to identical normal forms by quote:

$$\frac{v, w \;:\; \mathsf{Val}\,\Gamma\,\sigma}{v \sim_{\Gamma,\sigma} w \;:\; \mathsf{Prop}} \quad \text{where}$$

$$\frac{\overline{\mathbf{quote}}\,m = \overline{\mathbf{quote}}\,n}{m \sim_{\Gamma,\bullet} n}$$

$$\frac{\forall \Delta\,.\,\forall v, w : \mathsf{Val}\,(\Gamma +\!\!+ \Delta)\,\sigma\,.\,v \sim w \;\rightarrow\; f^{+_\Delta} @ v \sim g^{+_\Delta} @ w}{f \sim_{\Gamma,(\sigma\to\tau)} g}$$

The pointwise extension to environments is straightforward:

$$\frac{\vec{v}, \vec{w} \;:\; \mathsf{Env}\,\Gamma\,\Delta}{\vec{v} \sim \vec{w} \;:\; \mathsf{Prop}} \quad \text{where} \quad \frac{}{\varepsilon \sim \varepsilon} \qquad \frac{\vec{v} \sim \vec{w} \quad v \sim w}{(\vec{v}\,; v) \sim (\vec{w}\,; w)}$$

As before for strong computability we will need that $\sim$ is closed under weakening:

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

Lemma 9

$$\frac{v \sim_{\Gamma,\sigma} w}{v^{+\Delta} \sim_{(\Gamma+\!\!+\Delta),\sigma} w^{+\Delta}} \qquad \frac{\vec{v} \sim_{\Gamma,\Sigma} \vec{w}}{v^{+\Delta} \sim_{(\Gamma+\!\!+\Delta),\Sigma} w^{+\Delta}}$$

Proof
By induction over σ and Σ. $\square$

We will also need that we have defined a family of partial equivalence relations (PERs).

Lemma 10
For all $v, v' : \mathsf{Val}\,\Gamma\,\sigma$ such that $v \sim_{\Gamma,\sigma} v'$ is symmetric and transitive, and for all $\vec{v}, \vec{v}' : \mathsf{Env}\,\Gamma\,\Delta$ such that $\vec{v} \sim_{\Gamma,\Delta} \vec{v}'$ is symmetric and transitive.

Proof
By induction over σ for both properties for the value relation and corresponding by induction over Δ for the environment relation. Symmetry for environments requires symmetry for values, and transitivity for environments requires transitivity for values. Note also that we need symmetry of values to establish transitivity of values for the $\sigma \to \tau$ case. $\square$

Before we can establish the fundamental theorem for logical relations we have to show an *identity extension lemma*:

Lemma 11

$$\frac{t : \mathsf{Tm}\,\Gamma\,\sigma \qquad \vec{v} \sim \vec{w}}{\mathbf{eval}\,t\,\vec{v} \sim \mathbf{eval}\,t\,\vec{w}} \qquad \frac{\vec{t} : \mathsf{Subst}\,\Gamma\,\Delta \qquad \vec{v} \sim \vec{w}}{\overrightarrow{\mathbf{eval}}\,\vec{t}\,\vec{v} \sim \overrightarrow{\mathbf{eval}}\,\vec{u}\,\vec{w}}$$

Proof
By simultaneous induction over $t : \mathsf{Tm}\,\Gamma\,\sigma$ and $\vec{t} : \mathsf{Subst}\,\Gamma\,\Delta$. $\square$

To show that quote maps equivalent values to equal normal forms, we have to simultaneously establish a dual property, as before for strong computability.

Lemma 12

$$\frac{v \sim_{\Gamma,\sigma} w}{\mathbf{quote}_{\Gamma,\sigma}\,v = \mathbf{quote}_{\Gamma,\sigma}\,w}(q) \qquad \frac{\overline{\mathbf{quote}}_{\Gamma,\sigma}\,m = \overline{\mathbf{quote}}_{\Gamma,\sigma}\,n}{m \sim_{\Gamma,\sigma} n}(u)$$

Proof
By induction over σ. For base types both properties follow directly from the definition of $\sim$ and the observation that all values of base type are neutral. We show both properties for $(\sigma \to \tau)$:

(q) Given $f \sim_{\Gamma,(\sigma\to\tau)} g$ (1) we have to show $\mathbf{quote}_{\Gamma,(\sigma\to\tau)} f = \mathbf{quote}_{\Gamma,(\sigma\to\tau)}\,g$. This reduces to showing $\lambda_\sigma \mathbf{quote}_{(\Gamma;\sigma),\tau}\,(f^{+\sigma} @ \varnothing) = \lambda_\sigma \mathbf{quote}_{(\Gamma;\sigma),\tau}\,(g^{+\sigma} @ \varnothing)$. Applying Lemma 9 to (1) we obtain $f^{+\sigma} \sim_{(\Gamma;\sigma),(\sigma\to\tau)} g^{+\sigma}$ (2). Using ind.hyp. (u) for σ we can show $\varnothing \sim_{(\Gamma;\sigma),\sigma} \varnothing$ (3), and hence by the definition of $\sim$ and (2) and (3) we get $f^{+\sigma} @ \varnothing \sim_{(\Gamma;\sigma),\tau} f^{+\sigma} @ \varnothing$ (4). By applying ind.hyp (q) for τ to (4) we arrive at $\mathbf{quote}_{(\Gamma;\sigma),\tau}\,(f^{+\sigma} @ \varnothing) \sim_{(\Gamma;\sigma),\tau} \mathbf{quote}_{(\Gamma;\sigma),\tau}\,(f^{+\sigma} @ \varnothing)$.

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

(u) Given $\overline{\mathbf{quote}}_{\Gamma,(\sigma\to\tau)}\, m = \overline{\mathbf{quote}}_{\Gamma,(\sigma\to\tau)}\, n$ (1) we have to show $m \sim_{\Gamma,(\sigma\to\tau)} n$. Unfolding the definition of $\sim$ this means that given $v \sim_{(\Gamma+\Delta),\sigma} w$ (2) we have to show that $m^{+\Delta} @ v \sim_{(\Gamma+\Delta),\tau} n^{+\Delta} @ w$. Using the induction hypothesis (u) for τ reduces to showing that $\overline{\mathbf{quote}}_{(\Gamma+\Delta),\tau}\,(m^{+\Delta}\, v) = \overline{\mathbf{quote}}_{(\Gamma+\Delta),\tau}\,(n^{+\Delta}\, w)$. This follows from (1) and $\mathbf{quote}_{\Gamma+\Delta,\sigma}\, v = \mathbf{quote}_{\Gamma+\Delta,\sigma}\, w$ which we can show by using ind.hyp (q) for σ with (2). □

And also, we can exploit the second property to show that the identity environment is related to itself.

Corollary 13

$$\frac{x\ :\ \mathsf{Var}\,\Gamma\,\sigma}{x \sim x} \qquad \frac{\Gamma\ :\ \mathsf{Con}}{\mathbf{id}_\Gamma \sim \mathbf{id}_\Gamma}$$

We show the fundamental theorem of logical relations:

Theorem 14

$$\frac{t \simeq_{\beta\eta\sigma} u \qquad \vec{v} \sim \vec{w}}{\mathbf{eval}\, t\, \vec{v} \sim \mathbf{eval}\, u\, \vec{w}} \qquad \frac{\vec{t} \simeq_{\beta\eta\sigma} \vec{u} \qquad \vec{v} \sim \vec{w}}{\vec{\mathbf{eval}}\, \vec{t}\, \vec{v} \sim \vec{\mathbf{eval}}\, \vec{u}\, \vec{w}}$$

Proof
By mutual induction over the derivation of $t \simeq_{\beta\eta\sigma} u$ and $\vec{t} \simeq_{\beta\eta\sigma} \vec{u}$, as before we consider some typical cases. We assume that $\vec{v} \sim \vec{w}$(H).

refl, trans and sym Reflexivity follows from Lemma 11 and symmetry and transitivity from Lemma 10.

ξ: To show $\mathbf{eval}\,(\lambda_\sigma t)\vec{v} \sim \mathbf{eval}\,(\lambda_\sigma u)\vec{w}$ its sufficient to show $\lambda_\sigma t\,[\vec{v}] \sim_{\Gamma,\sigma\to\tau} \lambda_\sigma u\,[\vec{w}]$. Given $v \sim_{\Gamma+\Delta,\sigma} w$ we have to show that $\lambda_\sigma t\,[\vec{v}^{+\Delta}] @ v \sim_{\Gamma+\Delta,\tau} \lambda_\sigma u\,[\vec{w}^{+\Delta}] @ w$ which reduces to $\mathbf{eval}\, t\,(\vec{v}^{+\Delta}; v) \sim_{(\Gamma+\Delta),\sigma} \mathbf{eval}\, u\,(\vec{v}^{+\Delta}; w)$; this follows from the induction hypothesis and Lemma 9 applied to (H).

$\beta\sigma$: We have to show $\mathbf{eval}\,(((\lambda_\sigma t)[\vec{u}])\, u)\,\vec{v} \sim \mathbf{eval}\,(t\,[\vec{u}\,; u])\,\vec{w}$. This reduces to having to show $\mathbf{eval}\, t\,(\vec{\mathbf{eval}}\,\vec{u}\,\vec{v}\,; \mathbf{eval}\, u\,\vec{v}) \sim \mathbf{eval}\, t\,(\vec{\mathbf{eval}}\,\vec{u}\,\vec{w}\,; \mathbf{eval}\, u\,\vec{w})$. This follows from applying Lemma 11 to u and (H) to give (1), Lemma 11 to $\vec{u}$ and (H) to give (2) and Lemma 11 to t and (2,1).

assoc: We have to show $\vec{\mathbf{eval}}\,((\vec{s} \circ \vec{t}) \circ \vec{u})\,\vec{v} \sim \vec{\mathbf{eval}}\,(\vec{s} \circ (\vec{t} \circ \vec{u}))\,\vec{w}$. This reduces to showing $\vec{\mathbf{eval}}\,\vec{s}\,(\vec{\mathbf{eval}}\,\vec{t}\,(\vec{\mathbf{eval}}\,\vec{u}\,\vec{v})) \sim \vec{\mathbf{eval}}\,\vec{s}\,(\vec{\mathbf{eval}}\,\vec{t}\,(\vec{\mathbf{eval}}\,\vec{u}\,\vec{w}))$; this follows again from Lemma 11, applied first to $\vec{u}$ and (H) to give (1), then to $\vec{t}$ and (1) to give (2) and finally to $\vec{s}$ and (2). □

By putting everything together we can establish soundness of the normalisation function:

Proposition 15

$$\frac{t \simeq_{\beta\eta\sigma} u}{\mathbf{nf}\, t = \mathbf{nf}\, u}$$

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

Proof
Using Corollary 13 and Theorem 14 we can infer that $\mathbf{eval}\,t\,\mathbf{id} \sim \mathbf{eval}\,u\,\mathbf{id}$; and hence by Lemma 12 we obtain the result. □

7 System T

It is straightforward to extend our system to include a type of natural numbers. We replace the base type $\bullet$ with N and extend the syntax of terms with zero $\mathsf{0}$, successor suc and primitive recursion prec:

$$\frac{}{\mathsf{0} \;:\; \mathsf{Tm}\,\Gamma\,\mathsf{N}} \qquad \frac{t \;:\; \mathsf{Tm}\,\Gamma\,\mathsf{N}}{\mathsf{suc}\,t \;:\; \mathsf{Tm}\,\Gamma\,\mathsf{N}} \qquad \frac{n \;:\; \mathsf{Tm}\,\Gamma\,\mathsf{N} \quad f \;:\; \mathsf{Tm}\,\Gamma\,\mathsf{N}{\to}\sigma{\to}\sigma \quad z \;:\; \mathsf{Tm}\,\Gamma\,\sigma}{\mathsf{prec}\,n\,f\,z \;:\; \mathsf{Tm}\,\Gamma\,\sigma}$$

We add the following $\simeq$ rules to the equational theory (and congruences for suc and prec):

$$\begin{array}{lcll} \mathsf{prec}\,\mathsf{0} f\,z & \simeq & z & \mathsf{cprimrecz} \\ \mathsf{prec}\,(\mathsf{suc}\,n) f\,vz & \simeq & f\,n\,(\mathsf{prec}\,n\,f\,z) & \mathsf{cprimrecs} \end{array}$$

Values Val and normal forms are extended with $\mathsf{0}$ and suc and neutral terms Ne with a constructor to represent primitive recursion applied to a neutral natural number:

$$\frac{}{\mathsf{0} \;:\; \mathsf{Val}\,\Gamma\,\mathsf{N}} \quad \frac{v \;:\; \mathsf{Val}\,\Gamma\,\mathsf{N}}{\mathsf{suc}\,v \;:\; \mathsf{Val}\,\Gamma\,\mathsf{N}} \quad \frac{}{\mathsf{0} \;:\; \mathsf{Nf}\,\Gamma\,\mathsf{N}} \quad \frac{n \;:\; \mathsf{Nf}\,\Gamma\,\mathsf{N}}{\mathsf{suc}\,n \;:\; \mathsf{Nf}\,\Gamma\,\mathsf{N}}$$

$$\frac{n \;:\; \mathsf{Ne}_T\,\Gamma\,\mathsf{N} \quad f \;:\; \mathsf{Val}\,\Gamma\,(\mathsf{N}{\to}\sigma{\to}\sigma) \quad z \;:\; \mathsf{Val}\,\Gamma\,\sigma}{\mathsf{prec}\,n\,f\,z \;:\; \mathsf{Ne}_T\,\Gamma\,\sigma}$$

A separate semantic primitive recursor **pr** is added and **eval** extended to accommodate it:

$$\frac{n \;:\; \mathsf{Val}\,\Gamma\,\mathsf{N} \quad f \;:\; \mathsf{Val}\,\Gamma\,(\mathsf{N}{\to}\sigma{\to}\sigma) \quad z \;:\; \mathsf{Val}\,\Gamma\,\sigma}{\mathbf{pr}\,n\,f\,z \;:\; \mathsf{Val}\,\Gamma\,\sigma}$$

$$\begin{array}{lllcl} \mathbf{pr} & \mathsf{0} & f & z & \Rightarrow\; z \\ \mathbf{pr} & (\mathsf{suc}\,n) & f & z & \Rightarrow\; f\,@\,n\,@\,(\mathbf{pr}\,n\,f\,z) \end{array}$$

$$\begin{array}{llccl} \mathbf{eval} & \mathsf{0} & \vec{v} & \Rightarrow & \mathsf{0} \\ \mathbf{eval} & (\mathsf{suc}\,n) & \vec{v} & \Rightarrow & \mathsf{suc}\,(\mathbf{eval}\,n\,\vec{v}) \\ \mathbf{eval} & (\mathsf{prec}\,n\,f\,z) & \vec{v} & \Rightarrow & \mathbf{pr}\,(\mathbf{eval}\,n\,\vec{v})\,(\mathbf{eval} f\,n)\,(\mathbf{eval}\,z\,n) \end{array}$$

For **quote** we replace the case for $\mathbf{quote}_\bullet$ with cases for $\mathbf{quote}_{\mathsf{N}}$,

$$\begin{array}{llcl} \mathbf{quote}_{\mathsf{N}} & \mathsf{0} & \Rightarrow & \mathsf{0} \\ \mathbf{quote}_{\mathsf{N}} & (\mathsf{suc}\,n) & \Rightarrow & \mathsf{suc}\,(\mathbf{quote}_{\mathsf{N}}\,n) \\ \mathbf{quote}_{\mathsf{N}} & n & \Rightarrow & \overline{\mathbf{quote}}\,n \end{array}$$

and the big-step semantics is updated accordingly. Next we replace the base cases $\bullet$ in the definitions of SCV and $\sim$ with inductively defined notions for N:

$$\frac{}{\mathsf{SCV}_{\Gamma,\mathsf{N}}\,\mathsf{0}} \quad \frac{\mathsf{SCV}_{\Gamma,\mathsf{N}}\,n}{\mathsf{SCV}_{\Gamma,\mathsf{N}}\,(\mathsf{suc}\,n)} \quad \frac{\overline{\mathsf{quote}}\,n \Downarrow m \quad \ulcorner n \urcorner \simeq_{\beta\eta\sigma} \ulcorner m \urcorner}{\mathsf{SCV}_{\Gamma,\mathsf{N}}\,n}$$

$$\frac{}{\mathsf{0} \sim_{\mathsf{N}} \mathsf{0}} \quad \frac{m \sim_{\mathsf{N}} n}{\mathsf{suc}\,m \sim_{\mathsf{N}} \mathsf{suc}\,n} \quad \frac{\overline{\mathbf{quote}}\,m = \overline{\mathbf{quote}}\,n}{m {\sim_{\Gamma,\mathsf{N}}} n}$$

We also require an extra lemma to prove the fundamental theorem:

Lemma 16

$$\frac{\mathsf{SCV}_{\Gamma,(\mathsf{N}\to\sigma\to\sigma)}\,f \quad \mathsf{SCV}_{\Gamma\,\sigma}\,z \quad \mathsf{SCV}_{\Gamma,\mathsf{N}}\,n}{\exists v : \mathsf{Val}\,\Gamma\,\sigma\,.\,\mathsf{pr}\,f\;z\;n \Downarrow v \wedge \mathsf{prec}\,\ulcorner f\urcorner\ulcorner z\urcorner\ulcorner n\urcorner \simeq_{\mathsf{w}\sigma} \ulcorner v\urcorner \wedge \mathsf{SCV}\,v}$$

Proof
By induction over $\mathsf{SCV}_{\Gamma,\mathsf{N}}\,n$. □

8 Conclusions

Let us summarize the main result of this paper:

Theorem 17
We have defined a function in total type theory

$$\frac{t \;:\; \mathsf{Tm}\,\Gamma\,\sigma}{\mathbf{nf}\,t \;:\; \mathsf{Nf}\,\Gamma\,\sigma}$$

with the following properties:

soundness	$\dfrac{t \simeq_{\beta\eta\sigma} t'}{\mathbf{nf}\,t = \mathbf{nf}\,t'}$
completeness	$\dfrac{}{t \simeq_{\beta\eta\sigma} \ulcorner \mathbf{nf}\,t \urcorner}$
stability	$\dfrac{n \;:\; \mathsf{Nf}\,\Gamma\,\sigma}{\mathbf{nf}\,\ulcorner n \urcorner = n}$

Proof
Propositions 7, 8 and 15. □

As we have already indicated, we chose the names because we consider normal forms as a syntactic model construction. Moreover, the second property, completeness, implies that the inverse of soundness holds:

Corollary 18

$$\frac{\mathbf{nf}\,t = \mathbf{nf}\,u}{t \simeq_{\beta\eta\sigma} u}$$

Since our definition of normal form is a first-order inductive definition (see Proposition 3), it is clear that equality of normal forms is decidable. Hence, we obtain the following corollary:

Corollary 19
Given $t, u \;:\; \mathsf{Tm}\,\Gamma\,\sigma$, it is decidable whether $t \simeq_{\beta\eta\sigma} u$ holds.

Moreover, the last property, stability, clearly implies that **nf** is surjective on normal forms. As a consequence, we can prove relevant properties of terms by induction over normal forms.

This is not a new result: it can be obtained by proving strong normalisation of a suitably chosen small-step reduction relation (avoiding Melliès' problem) or by using normalisation by evaluation. What have we gained by our approach?

https://doi.org/10.1017/S0956796809007278 Published online by Cambridge University Press

First of all, the traditional approach using term rewriting does not directly lead to an implementation of normalisation. We can use strong normalisation to justify such an implementation, but this requires yet another proof. Also we wonder why we have to first fight with the non-determinism introduced by the small-step relation only to throw it away in the end anyway. The case of the typed λ-calculus with strong sums (Lindley 2007) is a good example. Lindley's (2007) analysis clearly suggests an algorithm to compute normal forms, but this is lost due to the need to fit it into the framework of term rewriting.

Second, the term-rewriting approach means that we need to prove the Church–Rosser property as a property of our rewriting system. In our experience, it often requires a fair amount of ingenuity to have the Church–Rosser property without losing completeness or strong normalisation. In our setting, equational soundness is shown by using the fundamental property of logical relation. This at least inspires some hope that our construction will be more easily generalizable to other calculi.

What about normalisation by evaluation (NBE)? In our view (Altenkirch *et al.* 1995), it is basically a semantic construction: we provide a complete model construction; we show completeness by constructively inverting evaluation. This approach gives us a beautiful high-level analysis of normalisation; however, its actual computational content is often not immediately clear, and the normalisation functions seems to work by magic. No doubt this counterintuitive nature of NBE has a lot to do with the intensive use of higher-order functions in its implementation. They are also the reason that NBE can be only formalised in a metatheory in which constructive higher-order functions are primitive. This may be one reason why the traditional approach using term rewriting is still more popular: it can be easily formulated within standard set theory. Our approach shares this feature, just replacing small-step reduction by a computationally more realistic big-step semantics.

It has been suggested by one anonymous referee that we should try to derive our algorithm from NBE together with the implementation of an evaluator for the functional functional metalanguage which is used to execute the higher-order program. It seems plausible that this is possible – however, we doubt that much is gained by doing so because we claim that our approach has its own intuitive beauty and does not need to be justified by translation. Let us look back at what we have done.

How do we implement normalisation? We reduce to values, corresponding to weak normal forms and iterate the process (**quote**). This gives rise to a normaliser to β-normal forms. The only modification required for η-equality in this case is to recursively η-expand every functional term. How do we prove termination? We adopt Tait's method of strengthening the induction hypothesis to function types – i.e. by using logical predicates. Since we go under λ's we really need Kripke logical predicates – this is traditionally swept under the carpet by syntactic trickery, using an infinite supply of fresh variables. Completeness, i.e. that the result of normalisation is convertible to its input, can be shown at the same time, since it follows the structure of the algorithm. How do we prove soundness, i.e. that convertible terms are mapped to identical normal forms? We use logical relations, indeed Kripke logical relations, for the same reasons as above. In the present paper we have spelled out the details of this construction in great detail, corresponding to a formalisation using the Agda system (Chapman 2007).

Our recipe, we believe, is applicable to many calculi. Clearly, we have to justify this claim by actually applying our method to well-known difficult cases: typed λ-calculus with sums or other extensions such as bar recursion, dependent types with η rules and the combination, i.e. dependent types with non-empty sums and bar recursion; the last are calculi whose metatheoretic properties are not yet established.

Acknowledgments

We would like to thank Conor McBride for advice related to the formalisation of the material; we would also like to thank Ulf Norell for implementing Agda and help with using it. We are grateful for the anonymous referees for their thorough and competent feedback which helped very much to improve the presentation of the material.

References

Abadi, M., Cardelli, L., Curien, P.-L. & Lèvy, J.-J. (1990) Explicit substitutions. In *Conference Record of 17th Annual ACM Symposium on Principles of Programming Languages, POPL '90 (San Francisco, CA, Jan. 1990)*. ACM Press, pp. 31–46.

Altenkirch, T. & Chapman, J. (2006) Tait in one big step. In *Proceedings of the Workshop on Mathematically Structured Functional Programming, MSFP 2006 (Kuressaare, July 2006)*, McBride, C. & Uustalu T. (eds), Electronic Workshops in Computing. BCS, article 4.

Altenkirch, T., Dybjer, P., Hofmann, M. & Scott, P. (2001) Normalization by evaluation for typed lambda calculus with coproducts. In *Proceedings of the 16th Annual IEEE Symposium on Logic in Computer Science, LICS '01*. Boston, MA.

Altenkirch, T., Hofmann, M. & Streicher, T. (1995) Categorical reconstruction of a reduction free normalization proof. In *Proceedings of the 6th International Confernece on Category Theory and Computer Science, CTCS '95 (Cambridge, August 1995)*, Pitt, D., Rydeheard, D. E. & Johnstone, P. (eds), Lecture Notes in Computer Science, vol. 953. Springer, pp. 182–199.

Balat, V. (2002) *Une étude des sommes fortes: isomorphismes et formes normales*. PhD thesis, Université Denis Diderot.

Berger, U. & Schwichtenberg, H. (1991) An inverse of the evaluation functional for typed λ–calculus. In *Proceedings of the 6th Annual IEEE Symposium on Logic in Computer Science, LICS '91 (Amsterdam, July 1991)*. IEEE CS Press, pp. 203–211.

Bove, A. & Capretta, V. (2001) Nested general recursion and partiality in type theory. In *Proceedings of the 14th International Conference on Theorem Proving in Higher-Order Logics, TPHOLs 2001 (Edinburgh, September 2001)*, Boulton, R. J., & Jackson, P. B. (eds), Lecture Notes in Computer Science, vol. 2152. Springer, pp. 121–135.

Chapman, J. (2007) Formalisation of big-step normalisation for system T [online]. Available at http://cs.ioc.ee/∼james/BSN.html (Accessed 4 May 2009).

Chapman, J., McBride, C. & Altenkirch, T. (2007) Epigram reloaded: a standalone typechecker for ETT. In *Trends in Functional Programming 6*, van Eekelen, M. (ed). Intellect, pp. 79–94.

Coquand, C. (2002) A formalised proof of the soundness and completeness of a simply typed lambda-calculus with explicit substitutions, *Higher-Order Symb. Comput.*, **15** (1): 57–90.

Coquand, T. (1991) An algorithm for testing conversion in type theory. In *Logical Frameworks*, Huet, G. & Plotkin, G. (eds). Cambridge University Press, pp. 255–279.

Coquand, T. & Dybjer, P. (1997) Intuitionistic model constructions and normalization proofs, *Math. Struct. Comput. Sci.*, **7** (1): 75–94.

David, R. (2001) Normalization without reducibility, *Ann. Pure Appl. Logic*, **107** (1–3): 121–130.

Ghani, N. (1995) Beta–eta equality for coproducts. In *Proceedings of the 2nd International Conference on Typed Lambda Calculi and Applications, TLCA '95 (Edinburgh, April 1995)*, Dezani-Ciancaglini, M. & Plotkin, G. (eds), Lecture Notes in Computer Science, vol. 902. Springer, pp. 171–185.

Girard, J.-Y., Lafont, Y. & Taylor, P. (1989) *Proofs and Types*. Cambridge Tracts in Theoretical Computer Science, vol. 7. Cambridge University Press.

Hofmann, M. (1997) Syntax and semantics of dependent types. *Semantics and logics of computation*, vol. 14, Pitts, A. M. & Dybjer, P. (eds). Cambridge University Press, pp. 79–130.

Jay, C. B. & Ghani, N. (1995) The virtues of eta-expansion, *J. Funct. Program.*, **5** (2): 135–154.

Levy, P. B. (2001) *Call-by-Push-Value*. PhD thesis, Queen Mary, University of London.

Lindley, S. (2007) Extensional rewriting with sums. In *Proceedings of the 8th International Conference on Typed Lambda Calculi and Applications, TLCA 2007 (Paris, June 2007)*, Ronchi Della Rocca, S. (ed), Lecture Notes in Computer Science, vol. 4583. Springer, pp. 255–271.

Martin-Löf, P. (1984) *Intuitionistic Type Theory*. Bibliopolis.

McBride, C. (2005a) Epigram [online]. Available at http://www.e-pig.org/ (Accessed 4 May 2009).

McBride, C. (2005b) Epigram: practical programming with dependent types. In *Revised Lectures from 5th International School on Advanced Functional Programming, AFP 2004 (Tartu, August 2004)*, Vene, V. & Uustalu, T. (eds), Lecture Notes in Computer Science, vol. 3622. Springer, pp. 130–170.

McBride, C. & McKinna, J. (2004) The view from the left, *J. Funct. Program.*, **14** (1): 69–111.

Melliès, P.-A. (1995) Typed lambda-calculi with explicit substitutions may not terminate. In *Proceedings of the 2nd International Conference on Typed Lambda Calculi and Applications, TLCA '95 (Edinburgh, April 1995)*, Dezani-Ciancaglini, M. & Plotkin, G. (eds), Lecture Notes in Computer Science, vol. 902. Springer, pp. 328–334.

Nordström, B., Petersson, K. & Smith, J. (1990) *Programming in Martin-Löf's Type Theory: An Introduction*. Oxford University Press.

Norell, U. (2007a) Agda 2 [online]. Available at http://www.cs.chalmers.se/∼ulfn/.

Norell, U. (2007b) *Towards a Practical Programming Language Based on Dependent Type theory*. PhD thesis, Chalmers University of Technology.

Tait, W. W. (1967) Intensional interpretations of functionals of finite type, *J. Symb. Logic*, **32** (2): 198–212.

Watkins, K., Cervesato, I., Pfenning, F. & Walker, D. (2004) A concurrent logical framework: the propositional fragment. In *Revised Selected Papers from 3rd Internationa; Workshop on Types for Proofs and Programs, TYPES 2003 (Torino, April/May 2003)*, Berardi, S., Coppo, M. & Damiani, F. (eds), Lecture Notes in Computer Sci., vol. 3085. Springer, pp. 355–377.