J. Functional Programming **7** (3): 303–319, May 1997. Printed in the United Kingdom

Thunks and the λ-calculus

JOHN HATCLIFF†

Computer Science Department, Copenhagen University
Universitetsparken 1, DK-2100 Copenhagen Ø, Denmark
e-mail: `hatcliff@diku.dk`

OLIVIER DANVY‡

Computer Science Department, Aarhus University
Ny Munkegade, Building 540, DK-8000 Aarhus C, Denmark
e-mail: `danvy@brics.dk`

Abstract

Thirty-five years ago, thunks were used to simulate call-by-name under call-by-value in Algol 60. Twenty years ago, Plotkin presented continuation-based simulations of call-by-name under call-by-value and *vice versa* in the λ-calculus. We connect all three of these classical simulations by factorizing the continuation-based call-by-name simulation $\mathscr{C}_{\mathrm{n}}$ with a thunk-based call-by-name simulation $\mathscr{T}$ followed by the continuation-based call-by-value simulation $\mathscr{C}_{\mathrm{v}}$ extended to thunks.

$$\begin{tikzcd}
\Lambda \arrow[r, "\mathscr{T}"] \arrow[dr, "\mathscr{C}_{\mathrm{n}}"'] & \Lambda_{\mathrm{thunks}} \arrow[d, "\mathscr{C}_{\mathrm{v}}"] \\
& \Lambda_{\mathrm{CPS}}
\end{tikzcd}$$

We show that $\mathscr{T}$ actually satisfies all of Plotkin's correctness criteria for $\mathscr{C}_{\mathrm{n}}$ (i.e. his **Indifference**, **Simulation** and **Translation** theorems). Furthermore, most of the correctness theorems for $\mathscr{C}_{\mathrm{n}}$ can now be seen as simple corollaries of the corresponding theorems for $\mathscr{C}_{\mathrm{v}}$ and $\mathscr{T}$.

Capsule Review

Many Continuation-Passing Style (CPS) transformations are complex and can be staged into conceptually different passes. This paper shows that the call-by-name CPS transformation developed by Reynolds and Plotkin can be split into a thunk-introduction phase followed by a call-by-value CPS transformation. Moreover, it proves that the first phase is sufficient for simulation purposes, formalising folklore from the days of Algol 60. The paper stands by itself, but readers may profit from having a copy of Plotkin's 1975 paper nearby.

† Supported by the Danish Research Academy and by the DART project (Design, Analysis and Reasoning about Tools) of the Danish Research Councils.
‡ Supported by BRICS (Basic Research in Computer Science, Centre of the Danish National Research Foundation).

1 Introduction

In his seminal paper 'Call-by-name, call-by-value and the λ-calculus', Plotkin (1975) presents simulations of call-by-name by call-by-value (and *vice versa*). Both of Plotkin's simulations rely on *continuations*. Since Algol 60, however, programming wisdom has it that *thunks* can be used to obtain a simpler simulation of call-by-name by call-by-value. We show that composing a thunk-based call-by-name simulation $\mathcal{T}$ with Plotkin's continuation-based call-by-value simulation $\mathcal{C}_{\mathrm{v}}$ actually yields Plotkin's continuation-based call-by-name simulation $\mathcal{C}_{\mathrm{n}}$ (sections 2 and 3). Revisiting Plotkin's correctness theorems (section 4), we provide a correction to his **Translation** property for $\mathcal{C}_{\mathrm{n}}$, and show that the thunk-based simulation $\mathcal{T}$ satisfies all of Plotkin's properties for $\mathcal{C}_{\mathrm{n}}$. The factorization of $\mathcal{C}_{\mathrm{n}}$ by $\mathcal{C}_{\mathrm{v}}$ and $\mathcal{T}$ makes it possible to derive correctness properties for $\mathcal{C}_{\mathrm{n}}$ from the corresponding results for $\mathcal{C}_{\mathrm{v}}$ and $\mathcal{T}$. This factorization has also found several other applications already (section 5). The extended version of this paper (Hatcliff and Danvy, 1995) gives a more detailed development as well as all proofs.

2 Continuation-based and thunk-based simulations

We consider Λ, the untyped λ-calculus parameterized by a set of basic constants b (Plotkin, 1975, p. 127).

$$\begin{array}{rcl} e & \in & \Lambda \\ e & ::= & b \mid x \mid \lambda x.e \mid e_0\, e_1 \end{array}$$

The sets $Values_{\mathrm{n}}[\Lambda]$ and $Values_{\mathrm{v}}[\Lambda]$ below represent the set of values from the language Λ under call-by-name and call-by-value evaluation, respectively.

$$\begin{array}{rcl} v & \in & Values_{\mathrm{n}}[\Lambda] \\ v & ::= & b \mid \lambda x.e \end{array} \qquad \begin{array}{rcl} v & \in & Values_{\mathrm{v}}[\Lambda] \\ v & ::= & b \mid x \mid \lambda x.e \end{array} \qquad \textit{...where } e \in \Lambda$$

Figure 1 displays Plotkin's call-by-name CPS transformation $\mathcal{C}_{\mathrm{n}}$ (which simulates call-by-name under call-by-value). (Note: the term 'CPS' stands for 'Continuation-Passing Style'. It was coined in Steele's MS thesis (Steele, 1978).) Figure 2 displays Plotkin's call-by-value CPS transformation $\mathcal{C}_{\mathrm{v}}$ (which simulates call-by-value under call-by-name). Figure 3 displays the standard thunk-based simulation of call-by-name using call-by-value evaluation of the language Λ_τ. Λ_τ extends Λ as follows:

$$\begin{array}{rcl} e & \in & \Lambda_\tau \\ e & ::= & \ldots \mid \mathit{delay}\ e \mid \mathit{force}\ e \end{array}$$

The operator *delay* suspends the evaluation of an expression – thereby coercing an expression to a value. Therefore, *delay e* is added to the value sets in Λ_τ:

$$\begin{array}{rcl} v & \in & Values_{\mathrm{n}}[\Lambda_\tau] \\ v & ::= & \ldots \mid \mathit{delay}\ e \end{array} \qquad \begin{array}{rcl} v & \in & Values_{\mathrm{v}}[\Lambda_\tau] \\ v & ::= & \ldots \mid \mathit{delay}\ e \end{array} \qquad \textit{...where } e \in \Lambda_\tau$$

The operator *force* triggers the evaluation of such a suspended expression. This is formalized by the following notion of reduction.

$$
\begin{array}{rcl}
\mathscr{C}_{\mathrm{n}}\langle\!\langle\cdot\rangle\!\rangle & : & \Lambda \rightarrow \Lambda \\
\mathscr{C}_{\mathrm{n}}\langle\!\langle v\rangle\!\rangle & = & \lambda k.k\,\mathscr{C}_{\mathrm{n}}\langle v\rangle \\
\mathscr{C}_{\mathrm{n}}\langle\!\langle x\rangle\!\rangle & = & \lambda k.x\,k \\
\mathscr{C}_{\mathrm{n}}\langle\!\langle e_0\,e_1\rangle\!\rangle & = & \lambda k.\mathscr{C}_{\mathrm{n}}\langle\!\langle e_0\rangle\!\rangle\,(\lambda y_0.y_0\,\mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle\,k) \\
\\
\mathscr{C}_{\mathrm{n}}\langle\cdot\rangle & : & \mathit{Values}_{\mathrm{n}}[\Lambda] \rightarrow \Lambda \\
\mathscr{C}_{\mathrm{n}}\langle b\rangle & = & b \\
\mathscr{C}_{\mathrm{n}}\langle \lambda x.e\rangle & = & \lambda x.\mathscr{C}_{\mathrm{n}}\langle\!\langle e\rangle\!\rangle
\end{array}
$$

Fig. 1. Call-by-name CPS transformation.

$$
\begin{array}{rcl}
\mathscr{C}_{\mathrm{v}}\langle\!\langle\cdot\rangle\!\rangle & : & \Lambda \rightarrow \Lambda \\
\mathscr{C}_{\mathrm{v}}\langle\!\langle v\rangle\!\rangle & = & \lambda k.k\,\mathscr{C}_{\mathrm{v}}\langle v\rangle \\
\mathscr{C}_{\mathrm{v}}\langle\!\langle e_0\,e_1\rangle\!\rangle & = & \lambda k.\mathscr{C}_{\mathrm{v}}\langle\!\langle e_0\rangle\!\rangle\,(\lambda y_0.\mathscr{C}_{\mathrm{v}}\langle\!\langle e_1\rangle\!\rangle\,(\lambda y_1.y_0\,y_1\,k)) \\
\\
\mathscr{C}_{\mathrm{v}}\langle\cdot\rangle & : & \mathit{Values}_{\mathrm{v}}[\Lambda] \rightarrow \Lambda \\
\mathscr{C}_{\mathrm{v}}\langle b\rangle & = & b \\
\mathscr{C}_{\mathrm{v}}\langle x\rangle & = & x \\
\mathscr{C}_{\mathrm{v}}\langle \lambda x.e\rangle & = & \lambda x.\mathscr{C}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle
\end{array}
$$

Fig. 2. Call-by-value CPS transformation.

$$
\begin{array}{rcl}
\mathscr{T} & : & \Lambda \rightarrow \Lambda_{\tau} \\
\mathscr{T}\langle\!\langle b\rangle\!\rangle & = & b \\
\mathscr{T}\langle\!\langle x\rangle\!\rangle & = & \mathit{force}\ x \\
\mathscr{T}\langle\!\langle \lambda x.e\rangle\!\rangle & = & \lambda x.\mathscr{T}\langle\!\langle e\rangle\!\rangle \\
\mathscr{T}\langle\!\langle e_0\,e_1\rangle\!\rangle & = & \mathscr{T}\langle\!\langle e_0\rangle\!\rangle\,(\mathit{delay}\ \mathscr{T}\langle\!\langle e_1\rangle\!\rangle)
\end{array}
$$

Fig. 3. Call-by-name thunk transformation.

$$
\begin{array}{rcl}
\mathscr{C}^{+}_{\mathrm{v}}\langle\!\langle\cdot\rangle\!\rangle & : & \Lambda_{\tau} \rightarrow \Lambda \\
 & \ldots & \\
\mathscr{C}^{+}_{\mathrm{v}}\langle\!\langle \mathit{force}\ e\rangle\!\rangle & = & \lambda k.\mathscr{C}^{+}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle\,(\lambda y.y\,k) \\
\\
\mathscr{C}^{+}_{\mathrm{v}}\langle\cdot\rangle & : & \mathit{Values}_{\mathrm{v}}[\Lambda_{\tau}] \rightarrow \Lambda \\
 & \ldots & \\
\mathscr{C}^{+}_{\mathrm{v}}\langle \mathit{delay}\ e\rangle & = & \mathscr{C}^{+}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle
\end{array}
$$

Fig. 4. Call-by-value CPS transformation (extended to thunks).

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

Definition 1 (τ-reduction)

$$force\ (delay\ e) \longrightarrow_{\tau} e$$

We also consider the conventional notions of reduction β, β_{v}, η, and η_{v} (Barendregt, 1984; Plotkin, 1975; Sabry and Felleisen, 1993).

Definition 2 (β, β_{v}, η, η_{v}-reduction)

$$\begin{array}{rll}
(\lambda x.e_1)\, e_2 & \longrightarrow_{\beta} & e_1[x := e_2] \\
(\lambda x.e)\, v & \longrightarrow_{\beta_{\mathrm{v}}} & e[x := v] \qquad v \in Values_{\mathrm{v}}[\Lambda] \\
\lambda x.e\, x & \longrightarrow_{\eta} & e \qquad x \notin FV(e) \\
\lambda x.v\, x & \longrightarrow_{\eta_{\mathrm{v}}} & v \qquad v \in Values_{\mathrm{v}}[\Lambda]\ \wedge\ x \notin FV(v)
\end{array}$$

For a notion of reduction r, $\longrightarrow_r$ also denotes construct compatible one-step reduction, $\twoheadrightarrow_r$ denotes the reflexive, transitive closure of $\longrightarrow_r$, and $=_r$ denotes the smallest equivalence relation generated by $\longrightarrow_r$ (Barendregt, 1984). We will also write $\lambda r \vdash e_1 = e_2$ when $e_1 =_r e_2$ (similarly for the other relations).

Figure 4 extends $\mathscr{C}_{\mathrm{v}}$ (see Figure 2) to obtain $\mathscr{C}_{\mathrm{v}}^{+}$ which CPS-transforms thunks. $\mathscr{C}_{\mathrm{v}}^{+}$ faithfully implements τ-reduction in terms of β_{v} (and thus β) reduction. We write β_i below to signify that the property holds indifferently for β_{v} and β.

Property 1
For all $e \in \Lambda_{\tau}$, $\lambda\beta_i \vdash \mathscr{C}_{\mathrm{v}}^{+}\langle\!\![force\ (delay\ e)]\!\!\rangle = \mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle$.

Proof

$$\begin{array}{rcl}
\mathscr{C}_{\mathrm{v}}^{+}\langle\!\![force\ (delay\ e)]\!\!\rangle & = & \lambda k.(\lambda k.k\,(\mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle))\,(\lambda y.y\, k) \\
& \longrightarrow_{\beta_i} & \lambda k.(\lambda y.y\, k)\, \mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle \\
& \longrightarrow_{\beta_i} & \lambda k.\mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle\, k \\
& \longrightarrow_{\beta_i} & \mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle
\end{array}$$

The last step holds since a simple case analysis shows that $\mathscr{C}_{\mathrm{v}}^{+}\langle\!\![e]\!\!\rangle$ always has the form $\lambda k.e'$ for some $e' \in \Lambda$. □

3 Connecting the continuation-based and thunk-based simulations

$\mathscr{C}_{\mathrm{n}}$ can be factored into two conceptually distinct steps: (1) the suspension of argument evaluation (captured in $\mathscr{T}$); and (2) the sequentialization of function application to give the usual tail-calls of CPS terms (captured in $\mathscr{C}_{\mathrm{v}}^{+}$).

Theorem 1
For all $e \in \Lambda$, $\lambda\beta_i \vdash (\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\![e]\!\!\rangle = \mathscr{C}_{\mathrm{n}}\langle\!\![e]\!\!\rangle$.

Proof
By induction over the structure of e:

case $e \equiv b$:

$$\begin{array}{rcl}
(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\![b]\!\!\rangle & = & \mathscr{C}_{\mathrm{v}}^{+}\langle\!\![b]\!\!\rangle \\
& = & \lambda k.k\, b \\
& = & \mathscr{C}_{\mathrm{n}}\langle\!\![b]\!\!\rangle
\end{array}$$

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

case $e \equiv x$:

$$
\begin{array}{rcl}
(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle x \rangle\!\rangle & = & \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle \mathit{force}\ x \rangle\!\rangle \\
& = & \lambda k.(\lambda k.k\, x)\,(\lambda y.y\, k) \\
& \longrightarrow_{\beta_i} & \lambda k.(\lambda y.y\, k)\, x \\
& \longrightarrow_{\beta_i} & \lambda k.x\, k \\
& = & \mathscr{C}_{\mathrm{n}}\langle\!\langle x \rangle\!\rangle
\end{array}
$$

case $e \equiv \lambda x.e'$:

$$
\begin{array}{rcl}
(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle \lambda x.e' \rangle\!\rangle & = & \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle \lambda x.\mathscr{T}\langle\!\langle e' \rangle\!\rangle \rangle\!\rangle \\
& = & \lambda k.k\,(\lambda x.(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e' \rangle\!\rangle) \\
& =_{\beta_i} & \lambda k.k\,(\lambda x.\mathscr{C}_{\mathrm{n}}\langle\!\langle e' \rangle\!\rangle) \qquad \textit{...by the ind. hyp.} \\
& = & \mathscr{C}_{\mathrm{n}}\langle\!\langle \lambda x.e' \rangle\!\rangle
\end{array}
$$

case $e \equiv e_0\, e_1$:

$$
\begin{array}{rcl}
(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_0\, e_1 \rangle\!\rangle & = & \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle \mathscr{T}\langle\!\langle e_0 \rangle\!\rangle\,(\mathit{delay}\ \mathscr{T}\langle\!\langle e_1 \rangle\!\rangle) \rangle\!\rangle \\
& = & \lambda k.(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_0 \rangle\!\rangle\,(\lambda y_0.(\lambda k.k\,(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_1 \rangle\!\rangle)\,(\lambda y_1.y_0\, y_1\, k)) \\
& \longrightarrow_{\beta_i} & \lambda k.(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_0 \rangle\!\rangle\,(\lambda y_0.(\lambda y_1.y_0\, y_1\, k)\,(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_1 \rangle\!\rangle) \\
& \longrightarrow_{\beta_i} & \lambda k.(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_0 \rangle\!\rangle\,(\lambda y_0.y_0\,(\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T})\langle\!\langle e_1 \rangle\!\rangle\, k) \\
& =_{\beta_i} & \lambda k.\mathscr{C}_{\mathrm{n}}\langle\!\langle e_0 \rangle\!\rangle\,(\lambda y_0.y_0\,\mathscr{C}_{\mathrm{n}}\langle\!\langle e_1 \rangle\!\rangle\, k) \qquad \textit{...by the ind. hyp.} \\
& = & \mathscr{C}_{\mathrm{n}}\langle\!\langle e_0\, e_1 \rangle\!\rangle
\end{array}
$$

□

This theorem implies that the diagram in the abstract commutes up to β_i-equivalence, i.e. indifferently up to β_{v}- and β-equivalence. Note that $\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T}$ and $\mathscr{C}_{\mathrm{n}}$ only differ by administrative reductions. In fact, if we consider optimizing versions of $\mathscr{C}_{\mathrm{n}}$ and $\mathscr{C}_{\mathrm{v}}$ that remove administrative redexes, then the diagram commutes up to identity (i.e. up to α-equivalence).

Figures 5 and 6 present two such optimizing transformations $\mathscr{C}_{\mathrm{n}.\mathit{opt}}$ and $\mathscr{C}_{\mathrm{v}.\mathit{opt}}$. The output of $\mathscr{C}_{\mathrm{n}.\mathit{opt}}$ is $\beta_{\mathrm{v}}\eta_{\mathrm{v}}$ equivalent to the output of $\mathscr{C}_{\mathrm{n}}$, and similarly for $\mathscr{C}_{\mathrm{v}.\mathit{opt}}$ and $\mathscr{C}_{\mathrm{v}}$, as shown by Danvy and Filinski (Danvy and Filinski, 1992, pp. 387, 367). Both are applied to the identity continuation. In Figures 5 and 6, they are presented in a two-level language *à la* Nielson and Nielson (1992). Operationally, the overlined λ's and @'s correspond to functional abstractions and applications in the program implementing the translation, while the underlined λ's and @'s represent abstract-syntax constructors. It is simple to transcribe $\mathscr{C}_{\mathrm{n}.\mathit{opt}}$ and $\mathscr{C}_{\mathrm{v}.\mathit{opt}}$ into functional programs.

The optimizing transformation $\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}$ is obtained from $\mathscr{C}_{\mathrm{v}.\mathit{opt}}$ by adding the following definitions:

$$
\begin{array}{rcl}
\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}\langle\!\langle \mathit{force}\ e \rangle\!\rangle & = & \overline{\lambda} k.\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}\langle\!\langle e \rangle\!\rangle\ \overline{@}\ (\overline{\lambda} y_0.y_0\ \underline{@}\ (\underline{\lambda} y_1.k\ \overline{@}\ y_1)) \\
\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}\langle \mathit{delay}\ e \rangle & = & \underline{\lambda} k.\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}\langle\!\langle e \rangle\!\rangle\ \overline{@}\ (\overline{\lambda} y.k\ \underline{@}\ y)
\end{array}
$$

Taking an operational view of these two-level specifications, the following theorem states that, for all $e \in \Lambda$, the result of applying $\mathscr{C}_{\mathrm{v}.\mathit{opt}}^{+}$ to $\mathscr{T}\langle\!\langle e \rangle\!\rangle$ (with an initial continuation $\overline{\lambda} a.a$) is α-equivalent to the result of applying $\mathscr{C}_{\mathrm{n}.\mathit{opt}}$ to e (with an initial continuation $\overline{\lambda} a.a$).

$$
\begin{array}{rcl}
\mathscr{C}_{\mathrm{n}.opt}\llbracket\cdot\rrbracket & : & \Lambda \to (\Lambda \to \Lambda) \to \Lambda \\
\mathscr{C}_{\mathrm{n}.opt}\llbracket v\rrbracket & = & \overline{\lambda} k.k \mathbin{\overline{@}} \mathscr{C}_{\mathrm{n}.opt}\langle v\rangle \\
\mathscr{C}_{\mathrm{n}.opt}\llbracket x\rrbracket & = & \overline{\lambda} k.x \mathbin{\underline{@}} (\underline{\lambda} y.k \mathbin{\overline{@}} y) \\
\mathscr{C}_{\mathrm{n}.opt}\llbracket e_0\, e_1\rrbracket & = & \overline{\lambda} k.\mathscr{C}_{\mathrm{n}.opt}\llbracket e_0\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y_0.y_0 \mathbin{\underline{@}} (\underline{\lambda} k.\mathscr{C}_{\mathrm{n}.opt}\llbracket e_1\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y_1.k \mathbin{\underline{@}} y_1)) \mathbin{\underline{@}} (\underline{\lambda} y_2.k \mathbin{\overline{@}} y_2)) \\
 & & \\
\mathscr{C}_{\mathrm{n}.opt}\langle\cdot\rangle & : & Values_{\mathrm{n}}[\Lambda] \to \Lambda \\
\mathscr{C}_{\mathrm{n}.opt}\langle b\rangle & = & b \\
\mathscr{C}_{\mathrm{n}.opt}\langle \lambda x.e\rangle & = & \underline{\lambda} x.\underline{\lambda} k.\mathscr{C}_{\mathrm{n}.opt}\llbracket e\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y.k \mathbin{\underline{@}} y)
\end{array}
$$

Fig. 5. Optimizing call-by-name CPS transformation.

$$
\begin{array}{rcl}
\mathscr{C}_{\mathrm{v}.opt}\llbracket\cdot\rrbracket & : & \Lambda \to (\Lambda \to \Lambda) \to \Lambda \\
\mathscr{C}_{\mathrm{v}.opt}\llbracket v\rrbracket & = & \overline{\lambda} k.k \mathbin{\overline{@}} \mathscr{C}_{\mathrm{v}.opt}\langle v\rangle \\
\mathscr{C}_{\mathrm{v}.opt}\llbracket e_0\, e_1\rrbracket & = & \overline{\lambda} k.\mathscr{C}_{\mathrm{v}.opt}\llbracket e_0\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y_0.\mathscr{C}_{\mathrm{v}.opt}\llbracket e_1\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y_1.y_0 \mathbin{\underline{@}} y_1 \mathbin{\underline{@}} (\underline{\lambda} y_2.k \mathbin{\overline{@}} y_2))) \\
 & & \\
\mathscr{C}_{\mathrm{v}.opt}\langle\cdot\rangle & : & Values_{\mathrm{v}}[\Lambda] \to \Lambda \\
\mathscr{C}_{\mathrm{v}.opt}\langle b\rangle & = & b \\
\mathscr{C}_{\mathrm{v}.opt}\langle x\rangle & = & x \\
\mathscr{C}_{\mathrm{v}.opt}\langle \lambda x.e\rangle & = & \underline{\lambda} x.\underline{\lambda} k.\mathscr{C}_{\mathrm{v}.opt}\llbracket e\rrbracket \mathbin{\overline{@}} (\overline{\lambda} y.k \mathbin{\underline{@}} y)
\end{array}
$$

Fig. 6. Optimizing call-by-value CPS transformation.

Theorem 2
For all $e \in \Lambda$, $(\mathscr{C}^{+}_{\mathrm{v}.opt} \circ \mathscr{T})\llbracket e\rrbracket \mathbin{\overline{@}} (\overline{\lambda} a.a) \equiv \mathscr{C}_{\mathrm{n}.opt}\llbracket e\rrbracket \mathbin{\overline{@}} (\overline{\lambda} a.a)$.

Proof
A simple structural induction similar to the one required in the proof of Theorem 1. We show only the case for identifiers (the others are similar). The overlined constructs are computed at translation time, and thus simplifying overlined constructs using β-conversion yields equivalent specifications.

case $e \equiv x$:

$$
\begin{array}{rcl}
(\mathscr{C}^{+}_{\mathrm{v}.opt} \circ \mathscr{T})\llbracket x\rrbracket & = & \overline{\lambda} k.(\overline{\lambda} k.k \mathbin{\overline{@}} x) \mathbin{\overline{@}} (\overline{\lambda} y_0.y_0 \mathbin{\underline{@}} (\underline{\lambda} y_1.k \mathbin{\overline{@}} y_1)) \\
 & = & \overline{\lambda} k.(\overline{\lambda} y_0.y_0 \mathbin{\underline{@}} (\underline{\lambda} y_1.k \mathbin{\overline{@}} y_1)) \mathbin{\overline{@}} x \\
 & = & \overline{\lambda} k.x \mathbin{\underline{@}} (\underline{\lambda} y_1.k \mathbin{\overline{@}} y_1) \\
 & = & \mathscr{C}_{\mathrm{n}.opt}\llbracket x\rrbracket
\end{array}
$$

□

4 Revisiting Plotkin's correctness properties

Figure 7 presents single-step evaluation rules specifying the call-by-name and call-by-value operational semantics of Λ programs (closed terms). The (partial) evaluation functions $eval_{\mathrm{n}}$ and $eval_{\mathrm{v}}$ are defined in terms of the reflexive, transitive closure (denoted $\longmapsto^{*}$) of the single-step evaluation rules.

$$eval_{\mathrm{n}}(e) = v \quad \text{iff} \quad e \longmapsto^{*}_{\mathrm{n}} v$$

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

Call-by-name:

$$(\lambda x.e_0)\, e_1 \longmapsto_{\mathrm{n}} e_0[x := e_1] \qquad \frac{e_0 \longmapsto_{\mathrm{n}} e_0'}{e_0\, e_1 \longmapsto_{\mathrm{n}} e_0'\, e_1}$$

Call-by-value:

$$(\lambda x.e)\, v \longmapsto_{\mathrm{v}} e[x := v] \qquad \frac{e_0 \longmapsto_{\mathrm{v}} e_0'}{e_0\, e_1 \longmapsto_{\mathrm{v}} e_0'\, e_1} \qquad \frac{e_1 \longmapsto_{\mathrm{v}} e_1'}{(\lambda x.e_0)\, e_1 \longmapsto_{\mathrm{v}} (\lambda x.e_0)\, e_1'}$$

Fig. 7. Single-step evaluation rules.

$$eval_{\mathrm{v}}(e) = v \quad \text{iff} \quad e \longmapsto_{\mathrm{v}}^{*} v$$

The evaluation rules for Λ_τ are obtained by adding the following rules to both the call-by-name and call-by-value evaluation rules of Figure 7.

$$\frac{e \longmapsto e'}{force\ e \longmapsto force\ e'} \qquad force\ (delay\ e) \longmapsto e$$

For a language l, $Programs[l]$ denotes the closed terms in l. For meta-language expressions E_1, E_2, we write $E_1 \simeq E_2$ when E_1 and E_2 are both undefined, or else both are defined and denote α-equivalent terms. We will also write $E_1 \simeq_r E_2$ when E_1 and E_2 are both undefined, or else are both defined and denote r-convertible terms for the convertibility relation generated by some notion of reduction r.

Plotkin expressed the correctness of his simulations $\mathscr{C}_{\mathrm{n}}$ and $\mathscr{C}_{\mathrm{v}}$ via three properties: **Indifference**, **Simulation** and **Translation**. **Indifference** states that call-by-name and call-by-value evaluation coincide on terms in the image of the CPS transformation. **Simulation** states that the desired evaluation strategy is properly simulated. **Translation** states how the transformation relates program calculi for each evaluation strategy (e.g. $\lambda\beta$, $\lambda\beta_{\mathrm{v}}$). Let us restate these properties for Plotkin's original presentation of $\mathscr{C}_{\mathrm{n}}$ (hereby noted $\mathscr{P}_{\mathrm{n}}$) (Plotkin, 1975, p. 153), that only differs from Figure 1 at the line for identifiers.

$$\mathscr{P}_{\mathrm{n}}\langle\!\langle x \rangle\!\rangle \quad = \quad x$$

Theorem 3
[Plotkin, 1975] For all $e \in Programs[\Lambda]$,

1. **Indifference:** $eval_{\mathrm{v}}(\mathscr{P}_{\mathrm{n}}\langle\!\langle e \rangle\!\rangle\, I) \simeq eval_{\mathrm{n}}(\mathscr{P}_{\mathrm{n}}\langle\!\langle e \rangle\!\rangle\, I)$
2. **Simulation:** $\mathscr{P}_{\mathrm{n}}\langle eval_{\mathrm{n}}(e)\rangle \simeq eval_{\mathrm{v}}(\mathscr{P}_{\mathrm{n}}\langle\!\langle e \rangle\!\rangle\, I)$

where I denotes the identity function and is used as the initial continuation.

Plotkin also claimed the following **Translation** property:

Claim 1
[Plotkin, 1975] For all $e_1, e_2 \in \Lambda$,

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

$$\begin{array}{rll}\textbf{Translation:}\quad \lambda\beta \vdash e_1 = e_2 & \text{iff} & \lambda\beta_v \vdash \mathscr{P}_n\langle\![e_1]\!\rangle = \mathscr{P}_n\langle\![e_2]\!\rangle \\ & \text{iff} & \lambda\beta \vdash \mathscr{P}_n\langle\![e_1]\!\rangle = \mathscr{P}_n\langle\![e_2]\!\rangle \\ & \text{iff} & \lambda\beta_v \vdash \mathscr{P}_n\langle\![e_1]\!\rangle\, I = \mathscr{P}_n\langle\![e_2]\!\rangle\, I \\ & \text{iff} & \lambda\beta \vdash \mathscr{P}_n\langle\![e_1]\!\rangle\, I = \mathscr{P}_n\langle\![e_2]\!\rangle\, I\end{array}$$

The **Translation** property purports to show that β-equivalence classes are preserved and reflected by $\mathscr{P}_n$. The property, however, does not hold because

$$\lambda\beta \vdash e_1 = e_2 \;\not\Rightarrow\; \lambda\beta_i \vdash \mathscr{P}_n\langle\![e_1]\!\rangle = \mathscr{P}_n\langle\![e_2]\!\rangle.$$

The proof breaks down at the statement 'It is straightforward to show that $\lambda\beta \vdash e_1 = e_2$ implies $\lambda\beta_v \vdash \mathscr{P}_n\langle\![e_1]\!\rangle = \mathscr{P}_n\langle\![e_2]\!\rangle$ ...' (Plotkin, 1975, p. 158). In some cases, η_v is needed to establish the equivalence of the CPS-images of two β-convertible terms. For example, $\lambda x.(\lambda z.z)\, x \longrightarrow_\beta \lambda x.x$ but

$$\begin{array}{rcll}\mathscr{P}_n\langle\![\lambda x.(\lambda z.z)\, x]\!\rangle & = & \lambda k.k\,(\lambda x.\lambda k.(\lambda k.k\,(\lambda z.z))\,(\lambda y.y\,x\,k)) & (1)\\ & \longrightarrow_{\beta_v} & \lambda k.k\,(\lambda x.\lambda k.(\lambda y.y\,x\,k)\,(\lambda z.z)) & (2)\\ & \longrightarrow_{\beta_v} & \lambda k.k\,(\lambda x.\lambda k.(\lambda z.z)\,x\,k) & (3)\\ & \longrightarrow_{\beta_v} & \lambda k.k\,(\lambda x.\lambda k.x\,k) & (4)\\ & \longrightarrow_{\eta_v} & \lambda k.k\,(\lambda x.x) \qquad \textit{...}\eta_v \textit{ is needed for this step} & (5)\\ & = & \mathscr{P}_n\langle\![\lambda x.x]\!\rangle. & (6)\end{array}$$

Since the two distinct terms at lines (4) and (5) are β_i-normal, confluence of β_i implies $\lambda\beta_i \nvdash \mathscr{P}_n\langle\![e_1]\!\rangle = \mathscr{P}_n\langle\![e_2]\!\rangle$.

In practice, though, η_v reductions such as those required in the example above are unproblematic if they are embedded in proper CPS contexts (e.g. contexts in the language of terms in the image of $\mathscr{P}_n$ closed under β_i reductions). When $\lambda k.k\,(\lambda x.\lambda k.x\,k)$ is embedded in a CPS context, x will always bind to a term of the form $\lambda k.e$ during evaluation. In this case, the η_v reduction can be expressed by a β_v reduction. If the term, however, is not embedded in a CPS context (e.g. $[\cdot]\,(\lambda y.y\,b)$), the η_v reduction is unsound, i.e., it fails to preserve *operational equivalence* as defined by Plotkin (Plotkin, 1975, pp. 144, 147). Such reductions are unsound due to 'improper' uses of basic constants. For example, $\lambda x.b\,x \longrightarrow_{\eta_v} b$ but $\lambda x.b\,x \not\approx_v b$ (take $C = [\cdot]$) where $\approx_v$ is the call-by-value operational equivalence relation defined by Plotkin (Hatcliff and Danvy, 1995, p. 9). Note, finally, that a simple typing discipline eliminates improper uses of basic constants, and consequently give soundness for η_v.

The simplest solution for recovering the **Translation** property is to change the translation of identifiers from $\mathscr{P}_n\langle\![x]\!\rangle = x$ to $\lambda k.x\,k$ – obtaining the translation $\mathscr{C}_n$ given in Figure 1.†

For the example above, the modified translation gives

$$\lambda\beta_i \vdash \mathscr{C}_n\langle\![\lambda x.(\lambda z.z)\, x]\!\rangle = \mathscr{C}_n\langle\![\lambda x.x]\!\rangle.$$

† In the context of Parigot's $\lambda\mu$-calculus (Parigot, 1992), de Groote independently noted the problem with Plotkin's **Translation** theorem and proposed a similar correction (de Groote, 1994).

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

The following theorem gives the correctness properties for $\mathscr{C}_{\mathrm{n}}$.

Theorem 4
For all $e \in \mathit{Programs}[\Lambda]$ and $e_1, e_2 \in \Lambda$,

1. **Indifference:** $\mathit{eval}_{\mathrm{v}}(\mathscr{C}_{\mathrm{n}}\langle\!\langle e\rangle\!\rangle\, I) \simeq \mathit{eval}_{\mathrm{n}}(\mathscr{C}_{\mathrm{n}}\langle\!\langle e\rangle\!\rangle\, I)$
2. **Simulation:** $\mathscr{C}_{\mathrm{n}}\langle \mathit{eval}_{\mathrm{n}}(e)\rangle \simeq_{\beta_i} \mathit{eval}_{\mathrm{v}}(\mathscr{C}_{\mathrm{n}}\langle\!\langle e\rangle\!\rangle\, I)$
3. **Translation:** $\lambda\beta \vdash e_1 = e_2$ iff $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle = \mathscr{C}_{\mathrm{n}}\langle\!\langle e_2\rangle\!\rangle$
   iff $\lambda\beta \vdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle = \mathscr{C}_{\mathrm{n}}\langle\!\langle e_2\rangle\!\rangle$
   iff $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle\, I = \mathscr{C}_{\mathrm{n}}\langle\!\langle e_2\rangle\!\rangle\, I$
   iff $\lambda\beta \vdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle\, I = \mathscr{C}_{\mathrm{n}}\langle\!\langle e_2\rangle\!\rangle\, I$

The **Indifference** and **Translation** properties remain the same. The **Simulation** property, however, holds up to β_i-equivalence while Plotkin's **Simulation** for $\mathscr{P}_{\mathrm{n}}$ holds up to α-equivalence. For example,

$$\mathscr{C}_{\mathrm{n}}\langle \mathit{eval}_{\mathrm{n}}((\lambda z.\lambda y.z)\, b)\rangle = \lambda y.\lambda k.k\, b$$

whereas

$$\mathit{eval}_{\mathrm{v}}(\mathscr{C}_{\mathrm{n}}\langle\!\langle [(\lambda z.\lambda y.z)\, b]\rangle\!\rangle\, I) = \lambda y.\lambda k.(\lambda k.k\, b)\, k.$$

In fact, proofs of **Indifference**, **Simulation** and most of the **Translation** can be derived from the correctness properties of $\mathscr{C}_{\mathrm{v}}^{+}$ and $\mathscr{T}$ (see section 5). All that remains of **Translation** is to show that $\lambda\beta \vdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e_1\rangle\!\rangle\, I = \mathscr{C}_{\mathrm{n}}\langle\!\langle e_2\rangle\!\rangle\, I$ implies $\lambda\beta \vdash e_1 = e_2$ and this follows in a straightforward manner from Plotkin's original proof for $\mathscr{P}_{\mathrm{n}}$ (Hatcliff and Danvy, 1995, p. 31). The following theorem gives the **Indifference**, **Simulation**, and **Translation** properties for $\mathscr{C}_{\mathrm{v}}$:

Theorem 5
[Plotkin, 1975] For all $e \in \mathit{Programs}[\Lambda]$ and $e_1, e_2 \in \Lambda$,

1. **Indifference:** $\mathit{eval}_{\mathrm{n}}(\mathscr{C}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle\, I) \simeq \mathit{eval}_{\mathrm{v}}(\mathscr{C}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle\, I)$
2. **Simulation:** $\mathscr{C}_{\mathrm{v}}\langle \mathit{eval}_{\mathrm{v}}(e)\rangle \simeq \mathit{eval}_{\mathrm{n}}(\mathscr{C}_{\mathrm{v}}\langle\!\langle e\rangle\!\rangle\, I)$
3. **Translation:** If $\lambda\beta_{\mathrm{v}} \vdash e_1 = e_2$ then $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{v}}\langle\!\langle e_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}\langle\!\langle e_2\rangle\!\rangle$
   Also $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{v}}\langle\!\langle e_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}\langle\!\langle e_2\rangle\!\rangle$ iff $\lambda\beta \vdash \mathscr{C}_{\mathrm{v}}\langle\!\langle e_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}\langle\!\langle e_2\rangle\!\rangle$

The **Translation** property states that β_{v}-convertible terms are also convertible in the image of $\mathscr{C}_{\mathrm{v}}$. In contrast to the theory $\lambda\beta$ appearing in the **Translation** property for $\mathscr{C}_{\mathrm{n}}$ (Theorem 4), the theory $\lambda\beta_{\mathrm{v}}$ is *incomplete* in the sense that it cannot prove the equivalence of some terms whose CPS images are provably equivalent using $\lambda\beta$ or $\lambda\beta_{\mathrm{v}}$ (Sabry and Felleisen, 1993). The properties of $\mathscr{C}_{\mathrm{v}}$ as stated in Theorem 5 can be extended to the transformation $\mathscr{C}_{\mathrm{v}}^{+}$ defined on the language T – the set of terms in the image of $\mathscr{T}$ closed under $\beta_i\tau$ reduction. It is straightforward to show that the following grammar generates exactly the set of terms T (Hatcliff and Danvy, 1995, pp. 32, 33).

$$t \quad ::= \quad b \quad | \quad \mathit{force}\; x \quad | \quad \mathit{force}\,(\mathit{delay}\; t) \quad | \quad \lambda x.t \quad | \quad t_0\,(\mathit{delay}\; t_1)$$

Theorem 6
For all $t \in \mathit{Programs}[T]$ and $t_1, t_2 \in T$,

1. **Indifference:** $eval_{\mathrm{n}}(\mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t\rangle\!\rangle\, I) \simeq eval_{\mathrm{v}}(\mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t\rangle\!\rangle\, I)$
2. **Simulation:** $\mathscr{C}_{\mathrm{v}}^{+}\langle eval_{\mathrm{v}}(t)\rangle \simeq eval_{\mathrm{n}}(\mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t\rangle\!\rangle\, I)$
3. **Translation:** If $\lambda\beta_{\mathrm{v}}\tau \vdash t_1 = t_2$ then $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_2\rangle\!\rangle$
   Also $\lambda\beta_{\mathrm{v}} \vdash \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_2\rangle\!\rangle$ iff $\lambda\beta \vdash \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_1\rangle\!\rangle = \mathscr{C}_{\mathrm{v}}^{+}\langle\!\langle t_2\rangle\!\rangle$

Proof
For **Indifference** and **Simulation** it is only necessary to extend Plotkin's colon-translation proof technique and definition of *stuck terms* to account for *delay* and *force*. The proofs then proceed along the same lines as Plotkin's original proofs for $\mathscr{C}_{\mathrm{v}}$ (Plotkin, 1975, pp. 148–152). **Translation** follows from the **Translation** component of Theorem 5 and Property 1 (Hatcliff and Danvy, 1995, p. 39). □

Thunks are sufficient for establishing a call-by-name simulation satisfying all of the correctness properties of the continuation-passing simulation $\mathscr{C}_{\mathrm{n}}$. Specifically, we prove the following theorem which recasts the correctness theorem for $\mathscr{C}_{\mathrm{n}}$ (Theorem 4) in terms of $\mathscr{T}$. The last two assertions of the **Translation** component of Theorem 4 do not appear here, since the identity function as the initial continuation only plays a rôle in CPS evaluation.

Theorem 7
For all $e \in Programs[\Lambda]$ and $e_1, e_2 \in \Lambda$,

1. **Indifference:** $eval_{\mathrm{v}}(\mathscr{T}\langle\!\langle e\rangle\!\rangle) \simeq eval_{\mathrm{n}}(\mathscr{T}\langle\!\langle e\rangle\!\rangle)$
2. **Simulation:** $\mathscr{T}\langle\!\langle eval_{\mathrm{n}}(e)\rangle\!\rangle \simeq_{\tau} eval_{\mathrm{v}}(\mathscr{T}\langle\!\langle e\rangle\!\rangle)$
3. **Translation:** $\lambda\beta \vdash e_1 = e_2$ iff $\lambda\beta_{\mathrm{v}}\tau \vdash \mathscr{T}\langle\!\langle e_1\rangle\!\rangle = \mathscr{T}\langle\!\langle e_2\rangle\!\rangle$
   iff $\lambda\beta\tau \vdash \mathscr{T}\langle\!\langle e_1\rangle\!\rangle = \mathscr{T}\langle\!\langle e_2\rangle\!\rangle$

Proof
The proof of **Indifference** is trivial: one can intuitively see from the grammar for T (which includes the set of terms in the image of $\mathscr{T}$ closed under evaluation steps) that call-by-name and call-by-value evaluation will coincide since all function arguments are values.

The proof of **Simulation** is somewhat involved. It begins by inductively defining a relation $\overset{\tau}{\sim} \subseteq \Lambda \times \Lambda_{\tau}$ such that $e \overset{\tau}{\sim} t$ holds exactly when $\lambda\tau \vdash \mathscr{T}\langle\!\langle e\rangle\!\rangle = t$. The crucial step is then to show that for all $e \in Programs[\Lambda]$ and $t \in Programs[\Lambda_{\tau}]$ such that $e \overset{\tau}{\sim} t$, $e \longmapsto_{\mathrm{n}} e'$ implies that there exists a t' such that $t \longmapsto_{\mathrm{v}}^{+} t'$ and $e' \overset{\tau}{\sim} t'$ (Hatcliff and Danvy, 1995, Sect. 2.3.2).

Translation is established by first defining a translation $\mathscr{T}^{-1} : \Lambda_{\tau} \to \Lambda$ that simply removes *delay* and *force* constructs. One then shows that $\mathscr{T}$ and $\mathscr{T}^{-1}$ establish an equational correspondence (Sabry and Felleisen, 1993) between theories $\lambda\beta$ and $\lambda\beta_{\mathrm{v}}\tau$ (and $\lambda\beta$ and $\lambda\beta\tau$). **Translation** follows as a corollary of this stronger result (Hatcliff and Danvy, 1995, Sect. 2.3.3). □

Representing thunks via abstract suspension operators *delay* and *force* simplifies the technical presentation and enables the connection between $\mathscr{C}_{\mathrm{n}}$ and $\mathscr{C}_{\mathrm{v}}$ presented in section 3. Elsewhere (Hatcliff, 1994) we show that the *delay*/*force* representation

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

$$
\begin{array}{rcll}
\mathcal{T}_{\mathcal{L}} & : & \Lambda \rightarrow \Lambda & \\
\mathcal{T}_{\mathcal{L}}\langle\!\![b]\!\!\rangle & = & b & \\
\mathcal{T}_{\mathcal{L}}\langle\!\![x]\!\!\rangle & = & x\,b & \textit{...for some arbitrary basic constant } b \\
\mathcal{T}_{\mathcal{L}}\langle\!\![\lambda x.e]\!\!\rangle & = & \lambda x.\mathcal{T}_{\mathcal{L}}\langle\!\![e]\!\!\rangle & \\
\mathcal{T}_{\mathcal{L}}\langle\!\![e_0\, e_1]\!\!\rangle & = & \mathcal{T}_{\mathcal{L}}\langle\!\![e_0]\!\!\rangle\,(\lambda z.\mathcal{T}_{\mathcal{L}}\langle\!\![e_1]\!\!\rangle) & \textit{...where } z \notin FV(e_1)
\end{array}
$$

Fig. 8. Thunk introduction implemented in Λ.

of thunks and associated properties (i.e. reduction properties and translation into CPS) are not arbitrary, but are determined by the relationship between strictness and continuation monads (Moggi, 1991).

Figure 8 presents the transformation $\mathcal{T}_{\mathcal{L}}$ that implements thunks directly in Λ using what Plotkin described as the 'protecting by a λ' technique (Plotkin, 1975, p. 147). An expression is delayed by wrapping it in an abstraction with a dummy parameter. A thunk is forced by applying it to a dummy argument.

The following theorem recasts the correctness theorem for $\mathcal{C}_{\mathrm{n}}$ (Theorem 4) in terms of $\mathcal{T}_{\mathcal{L}}$:

Theorem 8
For all $e \in \mathit{Programs}[\Lambda]$ and $e_1, e_2 \in \Lambda$,

1. **Indifference:** $\mathit{eval}_{\mathrm{v}}(\mathcal{T}_{\mathcal{L}}\langle\!\![e]\!\!\rangle) \simeq \mathit{eval}_{\mathrm{n}}(\mathcal{T}_{\mathcal{L}}\langle\!\![e]\!\!\rangle)$
2. **Simulation:** $\mathcal{T}_{\mathcal{L}}\langle\!\![\mathit{eval}_{\mathrm{n}}(e)]\!\!\rangle \simeq_{\beta_i} \mathit{eval}_{\mathrm{v}}(\mathcal{T}_{\mathcal{L}}\langle\!\![e]\!\!\rangle)$
3. **Translation:** $\lambda\beta \vdash e_1 = e_2$ iff $\lambda\beta_{\mathrm{v}} \vdash \mathcal{T}_{\mathcal{L}}\langle\!\![e_1]\!\!\rangle = \mathcal{T}_{\mathcal{L}}\langle\!\![e_2]\!\!\rangle$
   iff $\lambda\beta \vdash \mathcal{T}_{\mathcal{L}}\langle\!\![e_1]\!\!\rangle = \mathcal{T}_{\mathcal{L}}\langle\!\![e_2]\!\!\rangle$

Proof
Follows the same pattern as the proof of Theorem 7 (Hatcliff and Danvy, 1995, Sect. 2.4). □

5 Applications

5.1 Deriving correctness properties of $\mathcal{C}_{\mathrm{n}}$

When working with CPS, one often needs to establish technical properties for both a call-by-name and a call-by-value CPS transformation. This requires two sets of proofs that both involve CPS. By appealing to the factoring property, however, often only one set of proofs over call-by-value CPS terms is necessary. The second set of proofs deals with thunked terms which have a simpler structure. For instance, **Indifference** and **Simulation** for $\mathcal{C}_{\mathrm{n}}$ follow from **Indifference** and **Simulation** for $\mathcal{C}_{\mathrm{v}}^{+}$ and $\mathcal{T}$ and Theorem 1. Here we show only the results where evaluation is undefined or results in a basic constant b. See Hatcliff and Danvy (1995, p. 31) for a derivation of $\mathcal{C}_{\mathrm{n}}$ **Simulation** for arbitrary results.

For **Indifference**, let $e, b \in \Lambda$ where b is a basic constant. Then

$$
\begin{array}{cll}
 & \mathit{eval}_{\mathrm{v}}(\mathcal{C}_{\mathrm{n}}\langle\!\![e]\!\!\rangle\,(\lambda y.y)) = b & \\
\Leftrightarrow & \mathit{eval}_{\mathrm{v}}((\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\![e]\!\!\rangle\,(\lambda y.y)) = b & \textit{...Theorem 1 and the soundness of } \beta_{\mathrm{v}} \\
\Leftrightarrow & \mathit{eval}_{\mathrm{n}}((\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\![e]\!\!\rangle\,(\lambda y.y)) = b & \textit{...Theorem 6 } \textbf{(Indifference)} \\
\Leftrightarrow & \mathit{eval}_{\mathrm{n}}(\mathcal{C}_{\mathrm{n}}\langle\!\![e]\!\!\rangle\,(\lambda y.y)) = b & \textit{...Theorem 1 and the soundness of } \beta
\end{array}
$$

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

For **Simulation**, let $e, b \in \Lambda$ where b is a basic constant. Then

$$\begin{array}{lll} & eval_{\mathrm{n}}(e) = b & \\ \Leftrightarrow & eval_{\mathrm{v}}(\mathcal{T}\langle\!\llbracket e \rrbracket\!\rangle) = b & \text{...}\textit{Theorem 7}\ \textbf{(Simulation)} \\ \Leftrightarrow & eval_{\mathrm{n}}((\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\llbracket e \rrbracket\!\rangle\,(\lambda y.y)) = b & \text{...}\textit{Theorem 6}\ \textbf{(Simulation)} \\ \Leftrightarrow & eval_{\mathrm{v}}((\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\llbracket e \rrbracket\!\rangle\,(\lambda y.y)) = b & \text{...}\textit{Theorem 6}\ \textbf{(Indifference)} \\ \Leftrightarrow & eval_{\mathrm{v}}(\mathcal{C}_{\mathrm{n}}\langle\!\llbracket e \rrbracket\!\rangle\,(\lambda y.y)) = b & \text{...}\textit{Theorem 1 and the soundness of } \beta_{\mathrm{v}} \end{array}$$

For **Translation**, it is not possible to establish Theorem 4 (**Translation** for $\mathcal{C}_{\mathrm{n}}$) in the manner above since Theorem 6 (**Translation** for $\mathcal{C}_{\mathrm{v}}^{+}$) is weaker in comparison. However, the following weaker version can be derived: Let $e_1, e_2 \in \Lambda$. Then

$$\begin{array}{lll} & \lambda\beta \vdash e_1 = e_2 & \\ \Leftrightarrow & \lambda\beta_{\mathrm{v}}\tau \vdash \mathcal{T}\langle\!\llbracket e_1 \rrbracket\!\rangle = \mathcal{T}\langle\!\llbracket e_2 \rrbracket\!\rangle & \text{...}\textit{Theorem 7}\ \textbf{(Translation)} \\ \Rightarrow & \lambda\beta_i \vdash (\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\llbracket e_1 \rrbracket\!\rangle = (\mathcal{C}_{\mathrm{v}}^{+} \circ \mathcal{T})\langle\!\llbracket e_2 \rrbracket\!\rangle & \text{...}\textit{Theorem 6}\ \textbf{(Translation)} \\ \Leftrightarrow & \lambda\beta_i \vdash \mathcal{C}_{\mathrm{n}}\langle\!\llbracket e_1 \rrbracket\!\rangle = \mathcal{C}_{\mathrm{n}}\langle\!\llbracket e_2 \rrbracket\!\rangle & \text{...}\textit{Theorem 1} \\ \Rightarrow & \lambda\beta_i \vdash \mathcal{C}_{\mathrm{n}}\langle\!\llbracket e_1 \rrbracket\!\rangle\, I = \mathcal{C}_{\mathrm{n}}\langle\!\llbracket e_2 \rrbracket\!\rangle\, I & \text{...}\textit{compatibility of } =_{\beta_i} \end{array}$$

5.2 Deriving a CPS transformation directed by strictness information

Strictness information indicates arguments that may be safely evaluated eagerly (i.e., without being delayed) – in effect, reducing the number of thunks needed in a program and the overhead associated with creating and evaluating suspensions (Bloss *et al.*, 1988; Mycroft, 1981; Okasaki *et al.*, 1994). In an earlier work (Danvy and Hatcliff, 1993), we gave a transformation $\mathcal{T}_s$ that optimizes thunk introduction based on strictness information. We then used the factorization of $\mathcal{C}_{\mathrm{n}}$ by $\mathcal{C}_{\mathrm{v}}^{+}$ and $\mathcal{T}$ to derive an optimized CPS transformation $\mathcal{C}_s$ for strictness-analysed call-by-name terms. This staged approach can be contrasted with Burn and Le Métayer's monolithic strategy (Burn & Le Métayer, 1996).

The resulting transformation $\mathcal{C}_s$ yields both call-by-name-like and call-by-value-like continuation-passing terms. Due to the factorization, the proof of correctness for the optimized transformation follows as a corollary of the correctness of the strictness analysis and the correctness of $\mathcal{T}$ and $\mathcal{C}_{\mathrm{v}}^{+}$.

Amtoft (1993) and Steckler and Wand (1994) have proven the correctness of transformations which optimize the introduction of thunks based on strictness information.

5.3 Deriving a call-by-need CPS transformation

Okasaki, Lee and Tarditi (1994) have also applied the factorization to obtain a 'call-by-need CPS transformation' $\mathcal{C}_{need}$. The lazy evaluation strategy characterizing call-by-need is captured with memo-thunks (Bloss *et al.*, 1988). $\mathcal{C}_{need}$ is obtained by extending $\mathcal{C}_{\mathrm{v}}^{+}$ to transform memo-thunks to CPS terms with store operations (which are used to implement the memoization) and composing it with the memo-thunk introduction.

Okasaki *et al.* (1994) optimize $\mathcal{C}_{need}$ by using strictness information along the lines discussed above. They also use sharing information to detect where memo-thunks can be replaced by ordinary thunks. In both cases, optimizations are achieved by working with simpler thunked terms as opposed to working directly with CPS terms.

5.4 Alternative CPS transformations

Thunks can be used to factor a variety of call-by-name CPS transformations. In addition to those discussed here, one can factor a variant of Reynolds's CPS transformation directed by strictness information (Hatcliff, 1994; Reynolds, 1974), as well as a call-by-name analogue of Fischer's call-by-value CPS transformation (Fischer, 1993; Sabry and Felleisen, 1993).

Obtaining the desired call-by-name CPS transformation via $\mathscr{C}_{\mathrm{v}}^{+}$ and $\mathscr{T}$ depends on the representation of thunks. For example, if one works with $\mathscr{T}_{\mathscr{L}}$ (see Figure 8) instead of $\mathscr{T}$, $\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}}$ still gives a valid CPS simulation of call-by-name by call-by-value. However, β_i equivalence with $\mathscr{C}_{\mathrm{n}}$ is not obtained (i.e. $\lambda\beta_i \nvdash \mathscr{C}_{\mathrm{n}}\langle\!\langle e\rangle\!\rangle = (\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}})\langle\!\langle e\rangle\!\rangle$), as shown by the following derivations:

$$
\begin{aligned}
(\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}})\langle\!\langle x\rangle\!\rangle &= \mathscr{C}_{\mathrm{v}}\langle\!\langle x\, b\rangle\!\rangle \\
&= \lambda k.(x\, b)\, k
\end{aligned}
$$

$$
\begin{aligned}
(\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}})\langle\!\langle e_0\, e_1\rangle\!\rangle &= \mathscr{C}_{\mathrm{v}}\langle\!\langle \mathscr{T}_{\mathscr{L}}\langle\!\langle e_0\rangle\!\rangle\, (\lambda z.\mathscr{T}_{\mathscr{L}}\langle\!\langle e_1\rangle\!\rangle)\rangle\!\rangle \\
&= \lambda k.(\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}})\langle\!\langle e_0\rangle\!\rangle\, (\lambda y.(y\, (\lambda z.(\mathscr{C}_{\mathrm{v}} \circ \mathscr{T}_{\mathscr{L}})\langle\!\langle e_1\rangle\!\rangle))\, k)
\end{aligned}
$$

The representation of thunks given by $\mathscr{T}_{\mathscr{L}}$ is too concrete in the sense that the delaying and forcing of computation is achieved using specific instances of the more general abstraction and application constructs. When composed with $\mathscr{T}_{\mathscr{L}}$, $\mathscr{C}_{\mathrm{v}}$ treats the specific instances of thunks in their full generality, and the resulting CPS terms contain a level of inessential encoding of *delay* and *force*.

5.5 The factorization holds for types

Plotkin's continuation-passing transformations were originally stated in terms of untyped λ-calculi. These transformations have been shown to preserve well-typedness of terms (Griffin, 1990; Harper and Lillibridge, 1993; Meyer and Wand, 1985; Murthy, 1990). The thunk transformation $\mathscr{T}$ also preserves well-typedness of terms, and the relationship between $\mathscr{C}_{\mathrm{v}}^{+} \circ \mathscr{T}$ and $\mathscr{C}_{\mathrm{n}}$ is reflected in transformations on types (Hatcliff and Danvy, 1995, Sect. 4).

6 Related work

Ingerman (1961), in his work on the implementation of Algol 60, gave a general technique for generating machine code implementing procedure parameter passing. The term *thunk* was coined to refer to the compiled representation of a delayed expression as it gets pushed on the control (Raymond, 1992). Since then, the term *thunk* has been applied to other higher-level representations of delayed expressions and we have followed this practice.

Bloss, Hudak and Young (1988) study thunks as the basis of an implementation of lazy evaluation. Optimizations associated with lazy evaluation (e.g. overwriting a forced expression with its resulting value) are encapsulated in the thunk. They give several representations with differing effects on space and time overhead.

Riecke (1991) has used thunks to obtain fully abstract translations between versions of PCF with differing evaluation strategies. In effect, he establishes a fully abstract version of the **Simulation** property for thunks. The **Indifference** property is also immediate for Riecke, since all function arguments are values in the image of his translation (and this property is maintained under reductions). The thunk translation required for full abstraction is much more complicated than our transformation $\mathscr{T}$ and consequently it cannot be used to factor $\mathscr{C}_{\rm n}$. In addition, since Riecke's translation is based on typed-indexed retractions, it does not seem possible to use it (and the corresponding results) in an untyped setting as we require here.

Asperti and Curien formulate thunks in a categorical setting (Asperti, 1992; Curien, 1986). Two combinators *freeze* and *unfreeze*, which are analogous to *delay* and *force* but have slightly different equational properties, are used to implement lazy evaluation in the Categorical Abstract Machine. In addition, *freeze* and *unfreeze* can be elegantly characterized using a co-monad.

In his original paper (Plotkin, 1975, p. 147), Plotkin acknowledges that thunks provide some simulation properties but states that "...these 'protecting by a λ' techniques do not seem to be extendable to a complete simulation and it is fortunate that the technique of continuations is available." (Plotkin, 1975, p. 147). By 'protecting by a λ', Plotkin refers to a representation of thunks as λ-abstractions with a dummy parameter, as in Figure 8. In a set of unpublished notes, however, he later showed that the 'protecting by a λ' technique is sufficient for a complete simulation (Plotkin, 1978).

An earlier version of section 3 appeared in the proceedings of WSA'92 (Danvy and Hatcliff, 1992). Most of these proofs have been checked in Elf (Pfenning, 1991) by Niss and the first author (Niss and Hatcliff, 1995). Elsewhere (Hatcliff, 1994), we also consider an optimizing version of $\mathscr{T}$ that does not introduce thunks for identifiers occurring as function arguments:

$$\mathscr{T}_{opt}\langle\!\![e\,x]\!\!\rangle \quad = \quad \mathscr{T}_{opt}\langle\!\![e]\!\!\rangle\; x$$

$\mathscr{T}_{opt}$ generates a language T_{opt} which is more refined than T (referred to in Theorem 6).

Finally, Lawall and Danvy (1993) investigate staging the call-by-value CPS transformation into conceptually different passes elsewhere.

7 Conclusion

We have connected the traditional thunk-based simulation $\mathscr{T}$ of call-by-name under call-by-value and Plotkin's continuation-based simulations $\mathscr{C}_{\rm n}$ and $\mathscr{C}_{\rm v}$ of call-by-name and call-by-value. Almost all of the technical properties Plotkin established for $\mathscr{C}_{\rm n}$ follow from the properties of $\mathscr{T}$ and $\mathscr{C}_{\rm v}^{+}$ (the extension of $\mathscr{C}_{\rm v}$ to thunks). When reasoning about $\mathscr{C}_{\rm n}$ and $\mathscr{C}_{\rm v}$, it is thus often sufficient to reason about $\mathscr{C}_{\rm v}^{+}$

https://doi.org/10.1017/S0956796897002748 Published online by Cambridge University Press

and the simpler simulation $\mathscr{T}$. We have also given several applications involving deriving optimized continuation-based simulations for call-by-name and call-by-need languages and performing CPS transformation after static program analysis.

Acknowledgements

Andrzej Filinski, Sergey Kotov, Julia Lawall, Henning Niss and David Schmidt gave helpful comments on earlier drafts of this paper. Thanks are also due to Dave Sands for several useful discussions. Special thanks to Gordon Plotkin for enlightening conversations at the LDPL'95 workshop and for subsequently mailing us his unpublished course notes. Finally, we are grateful to the reviewers for their lucid comments and their exhortation to be more concise, and to our editors, for their encouragement and direction.

The commuting diagram was drawn with Kristoffer Rose's XY-pic package.

References

Amtoft, T. (1993) Minimal thunkification. In: Cousot, P., Falaschi, M., Filè, G. and Rauzy, A. (editors), *Proceedings of the Third International Workshop on Static Analysis WSA'93: Lecture Notes in Computer Science 724*, pp. 218–229. Springer-Verlag.

Asperti, A. (1992) A categorical understanding of environment machines. *J. Functional Programming* **2**(1): 23–59.

Barendregt, H. (1984) *The Lambda Calculus: its syntax and semantics.* North-Holland.

Bloss, A., Hudak, P. and Young, J. (1988) Code optimization for lazy evaluation. *Lisp and Symbolic Computation* **1**: 147–164.

Burn, G. and Le Métayer, D. (1996) Proving the correctness of compiler optimisations based on a global program analysis. *J. Functional Programming* **6**(1).

Curien, P.-L. (1986) *Categorical combinators, sequential algorithms and functional programming*. Research Notes in Theoretical Computer Science, Vol. 1. Pitman.

Danvy, O. and Filinski, A. (1992) Representing control, a study of the CPS transformation. *Mathematical Structures in Computer Science* **2**(4): 361–391.

Danvy, O. and Hatcliff, J. (1992) Thunks (continued). In: *Proceedings of the Second International Workshop on Static Analysis WSA'92*, pp. 3–11. *Bigre Journal* **81–82**. Bordeaux, France: IRISA, Rennes, France.

Danvy, O. and Hatcliff, J. (1993) CPS transformation after strictness analysis. *ACM Letters On Programming Languages And Systems* **1**(3): 195–212.

de Groote, P. (1994) A CPS-translation of the $\lambda\mu$-calculus. In: Sophie Tison (editor), *19th Colloquium on Trees in Algebra and Programming (CAAP'94): Lecture Notes in Computer Science 787*, pp. 47–58. Springer-Verlag.

Fischer, M. J. (1993) Lambda-calculus schemata. In: C. L. Talcott (editor), *Special Issue on Continuations (Part I). Lisp and Symbolic Computation* **6**(3/4). (An earlier version appeared in an *ACM Conference on Proving Assertions about Programs, SIGPLAN Notices* **7**(1) January 1972.)

Griffin, T. G. (1990) A formulae-as-types notion of control. In: P. Hudak (editor), *Proceedings of the Seventeenth Annual ACM Symposium on Principles Of Programming Languages*, pp. 47–58. ACM Press.

Harper, B. and Lillibridge, M. (1993) Polymorphic type assignment and CPS conversion. In: C. L. Talcott (editor), *Special Issue on Continuations (Part I). Lisp and Symbolic Computation* **6**(3/4).

Hatcliff, J. (1994) *The structure of continuation-passing styles.* PhD thesis, Department of Computing and Information Sciences, Kansas State University.

Hatcliff, J. and Danvy, O. (1995) *Thunks and the λ-calculus.* Technical Report 95/3. DIKU, Computer Science Department, University of Copenhagen.

Ingerman, P. Z. (1961) Thunks, a way of compiling procedure statements with some comments on procedure declarations. *Communications of the ACM* **4**(1): 55–58.

Lawall, J. L. and Danvy, O. (1993) Separating stages in the continuation-passing style transformation. In: S. L. Graham (editor), *Proceedings of the Twentieth Annual ACM Symposium on Principles Of Programming Languages*, pp. 124–136. ACM Press.

Meyer, A. R. and Wand, M. (1985) Continuation semantics in typed lambda-calculi (summary). In: R. Parikh (editor), *Logics of Programs – proceedings: Lecture Notes in Computer Science 193*, pp. 219–224.

Moggi, E. (1991) Notions of computation and monads. *Information and Computation* **93**: 55–92.

Murthy, C. R. (1990) *Extracting constructive content from classical proofs.* PhD thesis, Department of Computer Science, Cornell University.

Mycroft, A. (1981) *Abstract interpretation and optimising transformations for applicative programs.* PhD thesis, University of Edinburgh.

Nielson, F. and Nielson, H. R. (1992) *Two-level functional languages.* Cambridge Tracts in Theoretical Computer Science, Vol. 34. Cambridge University Press.

Niss, H. and Hatcliff, J. (1995) Encoding operational semantics in logical frameworks: A critical review of LF/Elf. In: B. Nördstrom (editor), *Proceedings of the 1995 Nordic Workshop on Programming Language Theory*, pp. 57–81.

Okasaki, C., Lee, P. and Tarditi, D. (1994) Call-by-need and continuation-passing style. In: C. I. Talcott (editor), *Special Issue on Continuations (Part II). LISP and Symbolic Computation* **7**(1).

Parigot, M. (1992) $\lambda\mu$-calculus: an algorithmic interpretation of classical natural deduction. In: A. Voronkov(editor), *Proceedings of the International Conference on Logic Programming and Automated Reasoning. Lecture Notes in Artificial Intelligence 624*, pp. 190–201.

Pfenning, F. (1991) Logic programming in the LF logical framework. In: G. Huet and G.Plotkin (editors), *Logical Frameworks*, pp. 149–181. Cambridge University Press.

Plotkin, G. D. (1975) Call-by-name, call-by-value and the λ-calculus. *Theoretical Computer Science* **1**: 125–159.

Plotkin, G. D. (1978) *Course notes on operational semantics.* Unpublished manuscript.

Raymond, E. (editor) (1992) *The New Hacker's Dictionary.* MIT Press.

Reynolds, J. C. (1974) On the relation between direct and continuation semantics. In: J. Loeckx (editor), *2nd Colloquium on Automata, Languages and Programming: Lecture Notes in Computer Science 14*, pp. 141–156.

Riecke, J. G. (1991) Fully abstract translations between functional languages. In: R. Cartwright (editor), *Proceedings of the Eighteenth Annual ACM Symposium on Principles Of Programming Languages*, pp. 245–254. ACM Press.

Sabry, A. and Felleisen, M. (1993) Reasoning about programs in continuation-passing style. In: C. L. Talcott (editor), *Special Issue on Continuations (Part I). Lisp and Symbolic Computation* **6**(3/4) 289–360.

Sabry, A. and Wadler, P. (1996) Compiling with reflections. In: R. K. Dybvig (editor), *Proceedings of the 1996 ACM SIGPLAN International Conference on Functional Programming.* ACM Press.

Steckler, P. and Wand, M. (1994) Selective thunkification. In: B. Le Charlier (editor), *Proceedings of the First International Static Analysis Symposium: Lecture Notes in Computer Science 864*, pp. 162–178.

Steele Jr., G. L. (1978) *Rabbit: A compiler for Scheme*. Technical Report AI-TR-474. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Talcott, C. L. (editor) (1993) *Special Issue on Continuations (Part I): LISP and Symbolic Computation* **6**(3/4).