Logical Methods in Computer Science
Vol. 2 (3:3) 2006, pp. 1–22
www.lmcs-online.org

Submitted Dec. 16, 2005
Published Jul. 27, 2006

RELATIONAL PARAMETRICITY AND CONTROL

MASAHITO HASEGAWA

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502 Japan, and PRESTO, Japan Science and Technology Agency
e-mail address: hassei@kurims.kyoto-u.ac.jp

Abstract. We study the equational theory of Parigot's second-order $\lambda\mu$-calculus in connection with a call-by-name continuation-passing style (CPS) translation into a fragment of the second-order λ-calculus. It is observed that the relational parametricity on the target calculus induces a natural notion of equivalence on the $\lambda\mu$-terms. On the other hand, the unconstrained relational parametricity on the $\lambda\mu$-calculus turns out to be inconsistent. Following these facts, we propose to formulate the relational parametricity on the $\lambda\mu$-calculus in a constrained way, which might be called "focal parametricity".

Dedicated to Prof. Gordon Plotkin on the occasion of his sixtieth birthday

1. Introduction

The *$\lambda\mu$-calculus*, introduced by Parigot [26], has been one of the representative term calculi for classical natural deduction, and widely studied from various aspects. Although it still is an active research subject, it can be said that we have some reasonable understanding of the first-order propositional $\lambda\mu$-calculus: we have good reduction theories, well-established CPS semantics and the corresponding operational semantics, and also some canonical equational theories enjoying semantic completeness [16, 24, 25, 36, 39]. The last point cannot be overlooked, as such complete axiomatizations provide deep understanding of equivalences between proofs and also of the semantic structure behind the syntactic presentation.

The *second-order $\lambda\mu$-calculus* ($\lambda\mu 2$), again due to Parigot [27], has been studied in depth as a calculus for second-order classical natural deduction. In particular, strong normalization results of $\lambda\mu 2$ [27, 23] and its extensions, e.g. with inductive types [21], have been a central research topic, because of the proof-theoretical importance of strong normalization. However, for $\lambda\mu 2$, it seems that there are few attempts of giving an equational theory supported by some fine semantic structure. This situation is rather frustrating, since without such equational and semantic accounts, we cannot discuss e.g. the correctness of the impredicative encoding of the datatypes in $\lambda\mu 2$. For the second-order λ-calculus $\lambda 2$ (system

2000 ACM Subject Classification: F.3.2.
Key words and phrases: polymorphism, parametricity, continuations.

DOI:10.2168/LMCS-2 (3:3) 2006

F) [9, 33], a subsystem of $\lambda\mu 2$, there are several beautiful results on the *relational parametricity* [34] and the universal properties of impredicative constructions [1, 15, 31, 40, 44], e.g. that $\mu X.\sigma = \forall X.(\sigma \to X) \to X$ (where σ covariant in X) gives an initial algebra of the functor $\Lambda X.\sigma$ in a suitable sense. We certainly wish to have such a story for $\lambda\mu 2$ too.

This work is an attempt to identify such an equational theory which is backed up by certain semantic structures. Specifically, we propose a relational parametricity principle which is sound and sufficiently powerful for deriving such equivalences on the $\lambda\mu$-terms.

1.1. **Parametric CPS semantics.** We first consider the semantics of $\lambda\mu 2$ given by a CPS-translation into a fragment of $\lambda 2$ — that of the second-order existential types $\exists X.\tau$, conjunction types $\tau_1 \wedge \tau_2$, and arrow types $\tau \to R$ into a distinguished type R (this choice of the target calculus is due to a recent work of Fujita [7]). The translation $(-)^\circ$ sends a type variable X to X, arrow type $\sigma_1 \to \sigma_2$ to $(\sigma_1^\circ \to R) \wedge \sigma_2^\circ$, and the universal type $\forall X.\sigma$ to $\exists X.\sigma^\circ$ — while a term $M : \sigma$ is sent to $[\![M]\!] : \sigma^\circ \to R$. It can be considered as a natural extension of Streicher's call-by-name CPS translation [36, 39, 43]. It follows that this translation already gives a reasonable equational theory on $\lambda\mu 2$, in that it validates the standard $\beta\eta$-equalities[1]. In fact, this is a consequence of a fibred version of the "category of continuations" construction [16, 36, 39].

However, this is just a starting point; we observe that, if some of the impredicative constructions in the target calculus satisfy certain universal properties (e.g. $\exists X.X$ is a terminal object) which follow from the relational parametricity, then so do the impredicative constructions in the source $\lambda\mu 2$-calculus — but not quite in the way that we first might expect. For instance, the type $\bot = \forall X.X$ does *not* give an initial object (cf. [37]) — instead it plays the role of the falsity type (or the "answer type"); in fact, we have a *double-negation elimination* from $(\sigma \to \bot) \to \bot$ to σ for any σ which actually is an algebra of the double-negation monad $((-) \to \bot) \to \bot$. As another major example, $\forall X.(\sigma \to X) \to X$ does *not* give an initial algebra of $\Lambda X.\sigma$; it gives an initial algebra of $\Lambda X.(\sigma \to \bot) \to \bot$ — not with respect to all terms but to a certain class of terms (the "focal terms", to be mentioned below). In particular, if X is not free in σ, $\forall X.(\sigma \to X) \to X$ is isomorphic not to σ but to $(\sigma \to \bot) \to \bot$. In short, impredicative encodings in $\lambda\mu 2$ get extra double negations, and the relational parametricity of $\lambda 2$ is not consistent with the equational theory of $\lambda\mu 2$ induced by the CPS semantics. As a consequence, we cannot encode cartesian products in $\lambda\mu 2$, though they can be added easily. Also we cannot express the classical disjunctions [36], though they can be added without changing the target of the CPS translation.

1.2. **Focal parametricity.** These results suggest that the CPS translation into parametric target calculus gives a reasonable semantic foundation and equational theory for $\lambda\mu 2$, which is sufficient for obtaining various interesting results. However, here the parametricity is used rather indirectly, via the CPS translation; we also wish to have a decent notion of parametricity directly within $\lambda\mu 2$. To figure out what sort of parametricity principle can be expected for $\lambda\mu 2$, recall the following fact on $\lambda 2$ with parametricity: given a polymorphic term $M : \forall X.F[X] \to G[X]$ (with X covariant in F and G) and types σ_1, σ_2, the instances

[1] We can say more – we can show that this CPS-semantics is sound and complete with respect to the $\beta\eta$-theory of $\lambda\mu 2$. This result, together with further syntactic analysis of this CPS translation, will appear in a forthcoming paper with Ken-etsu Fujita.

$M\,\sigma_1 : F[\sigma_1] \to G[\sigma_1]$ and $M\,\sigma_2 : F[\sigma_2] \to G[\sigma_2]$ obey the *naturality*, in that the following diagram

$$\begin{tikzcd} F[\sigma_1] \arrow[r, "F[f]"] \arrow[d, "M\,\sigma_1"'] & F[\sigma_2] \arrow[d, "M\,\sigma_2"] \\ G[\sigma_1] \arrow[r, "G[f]"'] & G[\sigma_2] \end{tikzcd}$$

commutes for *any* $f : \sigma_1 \to \sigma_2$. This is no longer true for $\lambda\mu 2$. For example, let $F[\sigma] = (\sigma \to \bot) \to \bot$, $G[\sigma] = \sigma$ and M be the double-negation elimination (which does not exist in $\lambda 2$); then the naturality for arbitrary maps implies inconsistency — we get $\sigma \simeq (\sigma \to \bot) \to \bot$ for every σ by letting f be the obvious map from σ to $(\sigma \to \bot) \to \bot$, which is enough to kill the theory [20]. Similar result can be observed for other "classical" proofs, e.g. of the Peirce law.

To this end, we look at the *focus* [36] (*centre* [32, 42], *C-maps* [16]) of $\lambda\mu 2$; a focal map is no other than an algebra morphism between the the double-negation monad mentioned above, i.e., a map making the naturality diagram for the double-negation elimination commute. It follows that a notion of relational parametricity on $\lambda\mu 2$ in which the construction of the graph relations is allowed only for focal maps is consistent, as there are nontrivial models. Together with the definability (fullness) of the CPS translation, we see that it is at least as powerful as the parametricity on the CPS target calculus which we have mentioned above, thus gives a powerful principle for deriving the equivalences of terms in $\lambda\mu 2$. (We actually conjecture that these two notions of parametricity do agree, but it is open as of writing this article.) This principle, which we shall call *focal parametricity*, should be a natural notion of parametricity for $\lambda\mu 2$. We will sketch some use of focal parametricity for deriving "free theorems" for $\lambda\mu 2$ syntactically.

1.3. **Towards parametricity for computational effects.** At the conceptual and abstract level, this story closely resembles to the study of *linear parametricity* and recursion [4, 30]. In the case of linear parametricity, the graph relations are allowed to be constructed only from the linear maps, and a linear map is an algebra map w.r.t. the lifting monad. We claim that, just like the linear parametricity gives a solution of accommodating non-termination and recursion in the polymorphic setting (as advocated by Plotkin [30]), the focal parametricity provides a way of accommodating control features in the polymorphic setting. In short:

$$\frac{\text{linear parametricity}}{\text{non-termination}} = \frac{\text{focal parametricity}}{\text{first-class control}}$$

As future work, it would be an interesting challenge to find a unifying framework of linear parametricity and focal parametricity; it should be useful to have parametric polymorphism, recursion, and control at once, as in the realistic programming languages (cf. [14, 18, 19]). More ambitiously, we are keen to see an adequate notion of parametricity for fairly general "effectful" settings. Possible starting points for this direction might include the "parametricity graphs" approach [5] which allows us to deal with parametricity at a general level (including the linear parametricity as an instance), and the "category of linear continuations construction" [12] which induces both the CPS translation and Girard translation as special cases. See Section 7 for further discussions related to this issue.

1.4. **Construction of this paper.** The rest of this paper is organised as follows. In section 2 and 3 we introduce the calculi and CPS-translation which are the subject of this study. In section 4 we consider the implications of the relational parametricity on the CPS-target calculus. The focal parametricity is introduced in section 5, followed by examples in section 6, including focally initial algebras and the type of Church numerals. Section 7 gives an alternative characterisation of focus, which suggests a generalisation of this work to a theory of parametricity for general effects. We then give some concluding remarks in Section 8.

2. The calculi

2.1. **The second-order $\lambda\mu$calculus.** The second-order $\lambda\mu$-calculus, $\lambda\mu 2$, is given as follows. We essentially follow Parigot's formulation [27] (with some flavour from Selinger's [36]). The types are the same as those of the second-order λ-calculus $\lambda 2$:

$$\sigma \ ::= \ X \mid \sigma \to \sigma \mid \forall X.\sigma$$

In a typing judgement $\Gamma \vdash M : \sigma \mid \Delta$, Γ stands for the typing context of variables, while Δ for the context of names (continuation variables).

$$\frac{}{\Gamma, x:\sigma, \Gamma' \vdash x : \sigma \mid \Delta}$$

$$\frac{\Gamma, x:\sigma_1 \vdash M : \sigma_2 \mid \Delta}{\Gamma \vdash \lambda x^{\sigma_1}.M : \sigma_1 \to \sigma_2 \mid \Delta}$$

$$\frac{\Gamma \vdash M : \sigma_1 \to \sigma_2 \mid \Delta \quad \Gamma \vdash N : \sigma_1 \mid \Delta}{\Gamma \vdash M\,N : \sigma_2 \mid \Delta}$$

$$\frac{\Gamma \vdash M : \sigma \mid \Delta \quad (X \notin FTV(\Gamma, \Delta))}{\Gamma \vdash \Lambda X.M : \forall X.\sigma \mid \Delta}$$

$$\frac{\Gamma \vdash M : \forall X.\sigma_1 \mid \Delta}{\Gamma \vdash M\,\sigma_2 : \sigma_1[\sigma_2/X] \mid \Delta}$$

$$\frac{\Gamma \vdash M : \sigma_2 \mid \alpha:\sigma_1, \Delta \quad (\beta:\sigma_2 \in \alpha:\sigma_1, \Delta)}{\Gamma \vdash \mu\alpha^{\sigma_1}.[\beta]M : \sigma_1 \mid \Delta}$$

The axioms for the equational theory are again the standard ones — note that we consider the extensional theory, i.e. with the η-axioms.

$$
\begin{array}{rcl}
(\lambda x^{\sigma}.M)\,N & = & M[N/x] \\
\lambda x^{\sigma}.M\,x & = & M\ (x \notin FV(M)) \\
(\Lambda X.M)\,\sigma & = & M[\sigma/X] \\
\Lambda X.M\,X & = & M\ (X \notin FTV(M))
\end{array}
$$

$$
\begin{array}{rcl}
\mu\alpha.[\beta](\mu\gamma.M) & = & \mu\alpha.M[\beta/\gamma] \\
\mu\alpha^{\sigma}.[\alpha]M & = & M\ (\alpha \notin FN(M)) \\
(\mu\alpha^{\sigma_1\to\sigma_2}.M)\,N & = & \mu\beta^{\sigma_2}.M[[\beta](-\,N)/[\alpha](-)] \\
(\mu\alpha^{\forall X.\sigma_1}.M)\,\sigma_2 & = & \mu\beta^{\sigma_1[\sigma_2/X]}.M[[\beta](-\,\sigma_2)/[\alpha](-)]
\end{array}
$$

In the last two axioms, we make uses of so-called "mixed substitution"; for instance, $M[[\beta](-\,N)/[\alpha](-)]$ means replacing occurances of the form $[\alpha]L$ in M by $[\beta](L\,N)$ recursively.

In the sequel, we frequently use the following syntactic sugar. First, we let $\bot$ be the type $\forall X.X$ — the type of falsity. We may also write $\neg\sigma$ for $\sigma \to \bot$. Using $\bot$, we define the "named term"

$$[\beta]M \ \equiv\ \mu\alpha^{\bot}.[\beta]M : \bot$$

(where $M : \sigma$, $\beta : \sigma$, with α fresh) and the μ-abstraction

$$\boldsymbol{\mu}\alpha^{\sigma}.M : \sigma \ \equiv\ \mu\alpha^{\sigma}.[\alpha](M\,\sigma)$$

for $M : \bot$. It follows that $\boldsymbol{\mu}\alpha^{\sigma}.[\beta]M = \mu\alpha^{\sigma}.[\beta]M$ holds.

With this $\bot$, we can express the double-negation elimination in $\lambda\mu 2$ by making use of both the polymorphic and classical features:

$$C_{\sigma} = \lambda m^{\neg\neg\sigma}.\boldsymbol{\mu}\alpha^{\sigma}.m\,(\lambda x^{\sigma}.[\alpha]x) : \neg\neg\sigma \to \sigma$$

As expected, we have $C_{\sigma}\,(\lambda k^{\neg\sigma}.k\,M) = M$. The properties of $\bot$ and C_{σ} will be further studied later under parametricity assumptions.

2.2. **Target: the $\{\exists, \wedge, \neg\}$calculus.** In tthe literature, the second-order λ-calculus ($\lambda 2$) is often taken as the target of the CPS translation for $\lambda\mu 2$. Fujita observed that it actually suffices to consider a fragment of $\lambda 2$ with negations, conjunctions and existential types as a target [7]. In this paper we follow this insight.

$$\tau \quad ::= \quad X \mid R \mid \neg\tau \mid \tau \wedge \tau \mid \exists X.\tau$$

$\neg\tau$ can be considered as a shorthand of $\tau \to R$. The type R can be replaced by $\exists X.\neg X \wedge X$, but for simplicity we keep R as a type constant. The syntax of terms is a fairly standard one, though for conjunctions we employ a slightly less familiar elimination rule (with let-binding) so that it parallels that of the existential types.

$$\frac{}{\Gamma, x : \tau, \Gamma' \vdash x : \tau}$$

$$\frac{\Gamma, x : \tau \vdash M : R}{\Gamma \vdash \lambda x^{\tau}.M : \neg\tau}$$

$$\frac{\Gamma \vdash M : \neg\tau \quad \Gamma \vdash N : \tau}{\Gamma \vdash M\,N : R}$$

$$\frac{\Gamma \vdash M : \tau_1 \quad \Gamma \vdash N : \tau_2}{\Gamma \vdash \langle M, N\rangle : \tau_1 \wedge \tau_2}$$

$$\frac{\Gamma \vdash M : \tau_1 \wedge \tau_2 \quad \Gamma, x : \tau_1, y : \tau_2 \vdash N : \tau_3}{\Gamma \vdash \mathsf{let}\ \langle x^{\tau_1}, y^{\tau_2}\rangle\ \mathsf{be}\ M\ \mathsf{in}\ N : \tau_3}$$

$$\frac{\Gamma \vdash M : \tau_1[\tau_2/X]}{\Gamma \vdash \langle \tau_2, M\rangle : \exists X.\tau_1}$$

$$\frac{\Gamma \vdash M : \exists X.\tau_1 \quad \Gamma, x : \tau_1 \vdash N : \tau_2 \quad X \notin FTV(\Gamma, \tau_2)}{\Gamma \vdash \mathsf{let}\ \langle X, x^{\tau_1}\rangle\ \mathsf{be}\ M\ \mathsf{in}\ N : \tau_2}$$

Again, we employ the standard $\beta\eta$-axioms.

$$\begin{array}{rcl}
(\lambda x^{\sigma}.M)\,N & = & M[N/x] \\
\lambda x^{\sigma}.M\,x & = & M\ (x \notin FV(M)) \\
\mathsf{let}\ \langle x, y\rangle\ \mathsf{be}\ \langle L, M\rangle\ \mathsf{in}\ N & = & N[L/x, M/y] \\
\mathsf{let}\ \langle x, y\rangle\ \mathsf{be}\ M\ \mathsf{in}\ N[\langle x, y\rangle/z] & = & N[M/z] \\
\mathsf{let}\ \langle X, x\rangle\ \mathsf{be}\ \langle \tau, M\rangle\ \mathsf{in}\ N & = & N[\tau/X, M/x] \\
\mathsf{let}\ \langle X, x\rangle\ \mathsf{be}\ M\ \mathsf{in}\ N[\langle X, x\rangle/z] & = & N[M/z]
\end{array}$$

3. CPS translation

3.1. **The CPS translation.** We present a call-by-name CPS translation which can be considered as an extension of that introduced by Streicher [16, 39, 36] (rather than the translations by Plotkin [29], Parigot [27] or Fujita [7] which introduce extra negations and do not respect extensionality).

$$\begin{array}{rcl}
X^{\circ} & = & X \\
(\sigma_1 \to \sigma_2)^{\circ} & = & \neg\sigma_1^{\circ} \wedge \sigma_2^{\circ} \\
(\forall X.\sigma)^{\circ} & = & \exists X.\sigma^{\circ}
\end{array}$$

$$\begin{array}{rcl}
[\![x^{\sigma}]\!] & = & x^{\neg\sigma^{\circ}} \\
[\![\lambda x^{\sigma_1}.M^{\sigma_2}]\!] & = & \lambda\langle x^{\neg\sigma_1^{\circ}}, k^{\sigma_2^{\circ}}\rangle.[\![M]\!]\,k \\
[\![M^{\sigma_1\to\sigma_2}\,N^{\sigma_1}]\!] & = & \lambda k^{\sigma_2^{\circ}}.[\![M]\!]\,\langle [\![N]\!], k\rangle \\
[\![\Lambda X.M^{\sigma}]\!] & = & \lambda\langle X, k^{\sigma^{\circ}}\rangle.[\![M]\!]\,k \\
[\![M^{\forall X.\sigma_1}\,\sigma_2]\!] & = & \lambda k^{\sigma_1[\sigma_2/X]^{\circ}}.[\![M]\!]\,\langle \sigma_2^{\circ}, k\rangle \\
[\![\mu\alpha^{\sigma_1}.[\beta^{\sigma_2}]M^{\sigma_2}]\!] & = & \lambda\alpha^{\sigma_1^{\circ}}.[\![M]\!]\,\beta
\end{array}$$

where

$$\begin{array}{rcl}
\lambda\langle x^{\sigma}, y^{\tau}\rangle.M & \equiv & \lambda z^{\sigma\wedge\tau}.\mathsf{let}\ \langle x^{\sigma}, y^{\tau}\rangle\ \mathsf{be}\ z\ \mathsf{in}\ M \\
\lambda\langle X, y^{\tau}\rangle.M & \equiv & \lambda z^{\exists X.\tau}.\mathsf{let}\ \langle X, y^{\tau}\rangle\ \mathsf{be}\ z\ \mathsf{in}\ M
\end{array}$$

3.2. **Soundness.** The type soundness follows from a straightforward induction.

Proposition 3.1 (type soundness)**.**

$$\Gamma \vdash M : \sigma \mid \Delta \implies \neg\Gamma^\circ, \Delta^\circ \vdash [\![M]\!] : \neg\sigma^\circ$$

where $\neg\Gamma^\circ$ is $x_1 : \neg\sigma_1^\circ, \ldots, x_m : \neg\sigma_m^\circ$ when Γ is $x_1 : \sigma_1, \ldots, x_m : \sigma_m$, and $\Delta^\circ = \alpha_1 : \sigma_1^\circ, \ldots, \alpha_n : \sigma_n^\circ$ for $\Delta = \alpha_1 : \sigma_1, \ldots, \alpha_n : \sigma_n$. □

Note that $(\sigma[\tau/X])^\circ \equiv \sigma^\circ[\tau^\circ/X]$, $[\![M[N/x]]\!] \equiv [\![M]\!][[\![N]\!]/x]$, and also $[\![M[\sigma/X]]\!] \equiv [\![M]\!][\sigma^\circ/X]$ hold. Then we have the equational soundness:

Proposition 3.2 (equational soundness)**.**

$$\Gamma \vdash M = N : \sigma \mid\Delta \implies \neg\Gamma^\circ, \Delta^\circ \vdash [\![M]\!] = [\![N]\!] : \neg\sigma^\circ$$

□

In addition, we have the definability result:

Proposition 3.3 (fullness)**.**

$$\neg\Gamma^\circ, \Delta^\circ \vdash N : \neg\sigma^\circ \Longrightarrow N = [\![M]\!] \text{ for some } \Gamma \vdash M : \sigma \mid \Delta$$ □

This can be proved by providing an inverse translation of the CPS translation, so that

$$\begin{array}{rcl}
\neg\Gamma^\circ, \Delta^\circ \vdash P : \neg\sigma^\circ & \Longrightarrow & \Gamma \vdash P^{-1} : \sigma \mid \Delta \\
\neg\Gamma^\circ, \Delta^\circ \vdash C : \sigma^\circ & \Longrightarrow & \Gamma \vdash C^{-1}[-^\sigma] : \bot \mid \Delta \\
\neg\Gamma^\circ, \Delta^\circ \vdash A : R & \Longrightarrow & \Gamma \vdash A^{-1} : \bot \mid \Delta
\end{array}$$

hold, where

$$\begin{array}{llcl}
\text{Program} : \neg\sigma^\circ & P & ::= & x \mid \lambda k.A \\
\text{Continuation} : \sigma^\circ & C & ::= & k \mid \langle P, C\rangle \mid \langle \sigma^\circ, C\rangle \mid \mathsf{let}\ \langle x, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ C \mid \mathsf{let}\ \langle X, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ C \\
\text{Answer} : R & A & ::= & P\,C \mid \mathsf{let}\ \langle x, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ A \mid \mathsf{let}\ \langle X, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ A
\end{array}$$

as follows.

$$\begin{array}{rcl}
x^{-1} & = & x \\
(\lambda k^{\sigma^\circ}.A)^{-1} & = & \boldsymbol{\mu} k^\sigma.A^{-1} \\
\\
k^{-1} & = & [k][-] \\
\langle P, C\rangle^{-1} & = & C^{-1}[-\,P^{-1}] \\
\langle \sigma^\circ, C\rangle^{-1} & = & C^{-1}[-\,\sigma] \\
(\mathsf{let}\ \langle x, k\rangle\ \mathsf{be}\ C_1\ \mathsf{in}\ C_2)^{-1} & = & C_1^{-1}[\lambda x.\mu k.C_2^{-1}[-]] \\
(\mathsf{let}\ \langle X, k\rangle\ \mathsf{be}\ C_1\ \mathsf{in}\ C_2)^{-1} & = & C_1^{-1}[\Lambda X.\mu k.C_2^{-1}[-]] \\
\\
(P\,C)^{-1} & = & C^{-1}[P^{-1}] \\
(\mathsf{let}\ \langle x, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ A)^{-1} & = & C^{-1}[\lambda x.\boldsymbol{\mu} k.A^{-1}] \\
(\mathsf{let}\ \langle X, k\rangle\ \mathsf{be}\ C\ \mathsf{in}\ A)^{-1} & = & C^{-1}[\Lambda X.\boldsymbol{\mu} k.A^{-1}]
\end{array}$$

This can be considered as a "continuation-grabbing style transformation" in the sense of Sabry [35]. It follows that for any $\neg\Gamma^\circ, \Delta^\circ \vdash M : \neg\sigma^\circ$ there exists $\neg\Gamma^\circ, \Delta^\circ \vdash P : \neg\sigma^\circ$ generated by this grammar such that $P = M$ — it suffices to take the β-normal form [7]. Moreover we can routinely show that $[\![P^{-1}]\!] = P$. Thus the CPS translation enjoys fullness: all terms are definable modulo the provable equality. This definability is important for relating the parametricity principles for the source and target calculi.

3.3. **A semantic explanation.** Here is a short explanation of why this CPS translation works, intended for readers with category theoretic background — on the "categories of continuations" construction [16, 39, 36], and on fibrations for polymorphic type theories [17]. As a response category $\mathcal{C}$ with a response object R induces a control category $R^{\mathcal{C}}$ with $R^{\mathcal{C}}(X,Y) = \mathcal{C}(R^X, R^Y)$, a fibred response category with finite products and simple coproducts (for existential quantifiers) induces a fibred control category with finite products and simple products (for universal quantifiers). Let us write $\mathcal{C}_\Gamma$ for the response category over the type-context Γ. We assume that the weakening functor $\pi^* : \mathcal{C}_\Gamma \to \mathcal{C}_{\Gamma\times A}$ has a left adjoint $\exists_A : \mathcal{C}_{\Gamma\times A} \to \mathcal{C}_\Gamma$ subject to the Beck-Chevalley condition. Thus

$$\mathcal{C}_{\Gamma\times A}(X, \pi^*(Y)) \simeq \mathcal{C}_\Gamma(\exists_A(X), Y)$$

We then have

$$\begin{array}{rcl} R^{\mathcal{C}_{\Gamma\times A}}(\pi^*(X), Y) & = & \mathcal{C}_{\Gamma\times A}(R^{\pi^*(X)}, R^Y) \\ & \simeq & \mathcal{C}_{\Gamma\times A}(Y, R^{R^{\pi^*(X)}}) \\ & \simeq & \mathcal{C}_{\Gamma\times A}(Y, \pi^*(R^{R^X})) \\ & \simeq & \mathcal{C}_\Gamma(\exists_A(Y), R^{R^X}) \\ & \simeq & \mathcal{C}_\Gamma(R^X, R^{\exists_A(Y)}) \\ & = & R^{\mathcal{C}_\Gamma}(X, \exists_A(Y)) \end{array}$$

Hence π^*, regarded as the weakening functor from $R^{\mathcal{C}_\Gamma}$ to $R^{\mathcal{C}_{\Gamma\times A}}$, has a right adjoint given by $\exists_A$, which can be used for interpreting the universal quantifier. Our CPS transformation is essentially a syntactic interpretation of this semantic construction.

4. CPS semantics with the parametric target calculus

4.1. **Parametricity for the target calculus.** As the target calculus can be seen as a subset of $\lambda 2$ (via the standard encoding of the conjunctions and existential types), we can define the relational parametricity for the target calculus in the same way as for $\lambda 2$, e.g. logic for parametricity [31, 40], system R [1], or system P [5]. One may directly define the parametricity principle (often called the simulation principle) for the existential type, see for example [31].

In this paper we only consider the relations constructed from the graphs of terms-in-context, identity, and σ^*'s obtained by the following construction, which we shall call "admissible relations".

Among admissible relations, the most fundamental are the graph relations. Given a term $f(x) : \tau_2$ with a free variable $x : \tau_1$ we define its graph relation $\langle x \vdash f(x)\rangle : \tau_1 \leftrightarrow \tau_2$ ($\langle f\rangle$ for short) by $u\,\langle f\rangle\,v$ iff $f(u) = v$.

Given a type τ whose free type variables are included in $X_1, \ldots, X_n$ and admissible relations $s_1 : \tau_1 \leftrightarrow \tau_1', \ldots, s_n : \tau_n \leftrightarrow \tau_n'$, we define an admissible relation τ^* as follows.

- $X_i^* = s_i : \tau_i \leftrightarrow \tau_i'$
- R^* is the identity relation on the terms of type R
- $(\neg\tau)^* : \neg\tau[\tau_1/X_1, \ldots] \leftrightarrow \neg\tau[\tau_1'/X_1, \ldots]$ is the relation so that $f\,(\neg\tau)^*\,g$ iff $x\,\tau^*\,y$ implies $f\,x\,R^*\,g\,y$ (hence $f\,x = g\,y$)
- $(\tau\wedge\tau')^* : (\tau\wedge\tau')[\tau_1/X_1, \ldots] \leftrightarrow (\tau\wedge\tau')[\tau_1'/X_1, \ldots]$ is the relation so that $u\,(\tau\wedge\tau')^*\,v$ iff $u = \langle x, x'\rangle$, $v = \langle y, y'\rangle$ and $x\,\tau^*\,y$, $x'\,\tau'^*\,y'$
- $(\exists X.\tau)^* : \exists X.\tau[\tau_1/X_1, \ldots] \leftrightarrow \exists X.\tau[\tau_1'/X_1, \ldots]$ is the relation so that $u\,(\exists X.\tau)^*\,v$ iff $u = \langle \tau', x\rangle$, $v = \langle \tau'', y\rangle$ and $x\,\tau[r/X]^*\,y$ for some admissible $r : \tau' \leftrightarrow \tau''$

In the last case, the relation $\tau[r/X]^* : \tau[\tau'/X] \leftrightarrow \tau[\tau''/X]$ is defined as τ^* with $X^* = r$. One may further define admissible relations $\neg r$, $r \wedge s$ and $\exists X.r$ for admissible r, s, so that $(\neg\tau)^* = \neg\tau^*$, $(\tau \wedge \tau')^* = \tau^* \wedge \tau'^*$ and $(\exists X.\tau)^* = \exists X.\tau^*$ hold.

Let $id_\tau : \tau \leftrightarrow \tau$ be the identity relation on the terms of type τ. The relational parametricity asserts that, for any τ whose free type variables are included in $X_1, \ldots, X_n$ and $\tau_1, \ldots, \tau_n$, $M : \tau[\tau_1/X_1, \ldots, \tau_n/X_n]$ implies $M\,\tau^*\,M$ with $s_i = id_{\tau_i}$.

Its consistency follows immediately from that of the parametricity for $\lambda 2$.

Proposition 4.1. As consequences of the parametricity, we can derive:

(1) $\exists X.X$ gives a terminal object $\top$ with a unique inhabitant $*$, so that for any $M : \top$ we have $M = *$.
(2) $\exists X.\neg(\tau \wedge X) \wedge X$ (which could be rewritten as $\exists X.(X \to \neg\tau) \wedge X$) gives a final coalgebra $\nu X.\neg\tau$ of $\Lambda X.\neg\tau$ where X only occurs negatively in τ.
(3) (as an instance of the last case) the isomorphism $\exists X.\neg(\tau \wedge X) \wedge X \simeq \neg\tau$ holds if X does not occur freely in τ. □

Their proofs are standard, cf. papers cited above [31, 1, 40, 5].

Below we will see the implications of these parametricity results on the target calculus. We refer to the $\lambda\mu 2$-theory induced by the CPS translation into this parametric target calculus as $\lambda\mu 2^P$.

4.2. **The falsity type.** As a first example, let us consider the falsity type $\bot = \forall X.X$ in $\lambda\mu 2$. We have

$$\bot^\circ = (\forall X.X)^\circ = \exists X.X \simeq \top$$

and

$$(\sigma \to \bot)^\circ = \neg(\sigma^\circ \wedge \bot^\circ) \simeq \neg(\sigma^\circ \wedge \top) \simeq \neg\sigma^\circ$$

Since $\exists X.X$ is terminal (with a unique inhabitant $*$) in the parametric target calculus, we obtain $[\![\boldsymbol{\mu}\alpha^\sigma.M]\!] = \lambda\alpha^{\sigma^\circ}.[\![M]\!]*$ and $[\![[\beta]M]\!] = \lambda u^{\exists X.X}.[\![M]\!]\beta$, which coincide with Streicher's translation. As a consequence, the following equations on the named terms and $\boldsymbol{\mu}$-abstractions are all validated in $\lambda\mu 2^P$.

$$\begin{array}{rcl} (\boldsymbol{\mu}\alpha^{\sigma_1\to\sigma_2}.M)\,N &=& \boldsymbol{\mu}\beta^{\sigma_2}.M[[\beta](-\,N)/[\alpha](-)] \\ (\boldsymbol{\mu}\alpha^{\forall X.\sigma_1}.M)\,\sigma_2 &=& \boldsymbol{\mu}\beta^{\sigma_1[\sigma_2/X]}.M[[\beta](-\,\sigma_2)/[\alpha](-)] \\ [\alpha'](\boldsymbol{\mu}\alpha^\sigma.M) &=& M[\alpha'/\alpha] \\ [\alpha^\bot]M &=& M \end{array}$$

Thus the type $\bot$ serves as the falsity type as found in some formulation of the $\lambda\mu$-calculus. In addition, we can show that $(\sigma, C_\sigma : ((\sigma \to \bot) \to \bot) \to \sigma)$ is an algebra of the double-negation monad $((-) \to \bot) \to \bot$ on the term model.

4.3. **Initial algebra?** A more substantial example is the "initial algebra" $\mu X.F[X] = \forall X.(F[X] \to X) \to X$, with X positive in $F[X]$ (here we see an unfortunate clash of μ's for the name-binding and for the fixed-point on types, but this should not cause any serious problem). We calculate:

$$\begin{array}{rcl}(\mu X.F[X])^\circ & = & (\forall X.(F[X] \to X) \to X)^\circ \\ & = & \exists X.\neg(\neg F[X]^\circ \wedge X) \wedge X \\ & \simeq & \nu X.\neg\neg F[X]^\circ \\ & \simeq & \neg\neg F[X]^\circ[\nu X.\neg\neg F[X]^\circ/X] \\ & \simeq & \neg\neg F[X]^\circ[(\mu X.F[X])^\circ/X] \\ & = & \neg\neg(F[\mu X.F[X]])^\circ \\ & \simeq & ((F[\mu X.F[X]] \to \bot) \to \bot)^\circ\end{array}$$

This suggests that $\mu X.F[X]$ is isomorphic not to $F[\mu X.F[X]]$ but to its double negation $(F[\mu X.F[X]] \to \bot) \to \bot$. One might think that this contradicts the standard experience on $\lambda 2$ with parametricity, where we have an isomorphism $\mathsf{in} : F[\mu X.F[X]] \to \mu X.F[X]$. Since $\lambda\mu 2$ subsumes $\lambda 2$, we have this in in $\lambda\mu 2$ too; however, it should not be an isomorphism, regarding the CPS interpretation above (otherwise it causes a degeneracy). The truth is that, in $\lambda\mu 2^P$, the term

$$\mathsf{in}^\sharp = \lambda m.\boldsymbol{\mu}\alpha.m\,(\lambda x.[\alpha](\mathsf{in}\,x)) \;:\; ((F[\mu X.F[X]] \to \bot) \to \bot) \to \mu X.F[X]$$

is an isomorphism. It still is not an initial algebra of $(F[-] \to \bot) \to \bot$; we shall further consider this issue later. For now, we shall emphasize that the parametricity principle for $\lambda 2$ should not be used for $\lambda\mu 2$, at least without certain constraint — otherwise in would be an isomorphism, hence a degeneracy follows (because we have $(\sigma \to \bot) \to \bot \simeq \sigma$ for every σ).

4.4. **Other impredicative encodings.** Recall other impredicative encodings of logical connectives:

$$\begin{array}{rcl}\top & = & \forall X.X \to X \\ \sigma_1 \wedge \sigma_2 & = & \forall X.(\sigma_1 \to \sigma_2 \to X) \to X \\ \sigma_1 \vee \sigma_2 & = & \forall X.(\sigma_1 \to X) \to (\sigma_2 \to X) \to X \\ \exists X.\sigma & = & \forall Y.(\forall X.(\sigma \to Y)) \to Y\end{array}$$

Their CPS translations into the parametric target calculus satisfy:

$$\begin{array}{rcl}\top^\circ & \simeq & R \\ (\sigma_1 \wedge \sigma_2)^\circ & \simeq & \neg(\neg{\sigma_1}^\circ \wedge \neg\sigma_2^\circ) \\ (\sigma_1 \vee \sigma_2)^\circ & \simeq & \neg\neg\sigma_1^\circ \wedge \neg\neg\sigma_2^\circ \\ (\exists X.\sigma)^\circ & \simeq & \neg\exists X.\neg\sigma^\circ\end{array}$$

As easily seen, these defined logical connectives in the source calculus do not obey the standard universal properties as in the parametric models of $\lambda 2$. In short, they are all "double-negated", hence amount to some classical encodings:

- $\sigma_1 \wedge \sigma_2$ is not a cartesian product of σ_1 and σ_2, but isomorphic to $(\sigma_1 \to \sigma_2 \to \bot) \to \bot$. It is possible to add cartesian product types $\sigma_1 \times \sigma_2$ to $\lambda\mu 2$, but then we also need to add coproduct types $\tau_1 + \tau_2$ to the target calculus, so that $(\sigma_1 \times \sigma_2)^\circ = \sigma_1^\circ + \sigma_2^\circ$ and $\sigma_1 \wedge \sigma_2 \simeq \neg\neg(\sigma_1 \times \sigma_2)$.
- $\top$ is not a terminal object, but isomorphic to $\bot \to \bot$. We can add a terminal object 1 to $\lambda\mu 2$ and an initial object 0 to the target, so that $1^\circ = 0$ and $\top \simeq \neg\neg 1$.
- $\sigma_1 \vee \sigma_2$ is not a coproduct of σ_1 and σ_2, but isomorphic to $(\sigma_1 \to \bot) \to (\sigma_2 \to \bot) \to \bot$. If there is a coproduct $\sigma_1 + \sigma_2$, then it should follow that $\sigma \vee \tau \simeq \neg\neg(\sigma + \tau)$. On the other hand, it is not possible to enrich $\lambda\mu 2$ with an initial object without a degeneracy, cf. Selinger's note on control categories [37]. Alternatively we might

add the "classical disjunction types" $\sigma_1 \wp \sigma_2$ [36] with $(\sigma_1 \wp \sigma_2)^\circ = \sigma_1^\circ \wedge \sigma_2^\circ$ — hence $\sigma_1 \to \sigma_2 \simeq \neg\sigma_1 \wp \sigma_2$ and $\sigma_1 \vee \sigma_2 \simeq \neg\neg\sigma_1 \wp \neg\neg\sigma_2$. We note that $\bot = \forall X.X$ serves as the unit of this classical disjunction.

- $\exists X.\sigma$ does not work as the existential type; it is isomorphic to $\neg\forall X.\neg\sigma$.

4.5. **Answer-type polymorphism.** Note that the answer type R has been considered just as a constant with no specific property. In fact we could have used any type for R — Everything is defined polymorphically regarding R. Thus we can apply the "answer-type polymorphism" principle (cf. [43]): in particular, a closed term of type σ in $\lambda\mu 2$ can be considered to be sent to a $\lambda 2$-term of type $\forall R.\neg\sigma^\circ$. This way of reasoning goes behind the parametricity principle for our target calculus, but it is justified by the parametricity of $\lambda 2$.

For instance, consider the type $\top = \forall X.X \to X$ of $\lambda\mu 2$. We have

$$\begin{array}{rcl}\forall R.\neg\top^\circ & \simeq & \forall R.\neg R\\ & \simeq & \forall R.R \to R\\ & \simeq & 1\end{array}$$

in $\lambda 2$ with parametricity. This means that, although $\top$ is not a terminal object in $\lambda\mu 2$, it has a unique closed inhabitant. Similarly, we have $\forall R.\neg\bot^\circ \simeq \forall R.R \simeq 0$, thus we see that there is no closed inhabitant of $\bot$ in $\lambda\mu 2$.

However, such reasonings based on the answer-type polymorphism become much harder for more complicated types. The force of answer-type polymorphism in this setting seems still not very obvious.

5. Focal parametricity

We have seen that the CPS semantics with respect to the target calculus with relational parametricity induces a reasonable equational theory $\lambda\mu 2^P$. However, here the parametricity is used rather indirectly, via the CPS translation. We now consider a notion of parametricity which is directly available within $\lambda\mu 2$.

5.1. **CPS translating relations.** The key of formulating the relational parametricity is the use of graph relations of terms (considered as representing a functional relation): without graph relations, relational parametricity reduces to just the basic lemma of the (second-order) logical relations. On the other hand, it does not have to allow all terms to be used for constructing relations. In fact, in linear parametricity [30] only linear (or strict) maps are allowed to be used for constructing graph relations, and this choice allows a weaker notion of parametricity which can accommodate recursion. Naturally, we are led to look for a characterisation of $\lambda\mu$-terms which can be used for graph relations without breaking the soundness with respect to the CPS semantics into the parametric target calculus.

Now suppose that we are allowed to use the graph relation $\langle f\rangle : \sigma_1 \leftrightarrow \sigma_2$ of a term $x : \sigma_1 \vdash f(x) : \sigma_2$. To ensure the soundness of the use of this graph relation, we shall consider the CPS translation of such relations. For instance, we hope that $\langle f\rangle$ will be sent to a relation between types σ_2° and σ_1° in the target calculus. However, since $x : \neg\sigma_1^\circ \vdash [\![f(x)]\!] : \neg\sigma_2^\circ$, we have some relation $\sigma_2^\circ \leftrightarrow \sigma_1^\circ$ only when $[\![f(x)]\!] = \lambda k.x\,(g\,(k))$ for some $k : \sigma_2^\circ \vdash g(k) : \sigma_1^\circ$ in the target calculus. If there is such g, we can complete the translation of the relations and reduce the parametricity principle on $\lambda\mu 2$ to the parametricity on the target calculus.

Fortunately, there is a way to characterise such "translatable" f's in the $\lambda\mu$-calculus without performing the CPS-translation (modulo a technical assumption on the CPS-target, known as "equalising requirement" [22]). It is the notion of "focus", which we now recall below.

5.2. **Focus.**

Definition 5.1. *A $\lambda\mu 2$-term $M : \sigma_1 \to \sigma_2$ is called* focal *if it is an algebra morphism from (σ_1, C_{σ_1}) to (σ_2, C_{σ_2}), i.e. the following diagram commutes.*

$$\begin{tikzcd}
(\sigma_1 \to \bot) \to \bot \arrow[r, "(M\to\bot)\to\bot"] \arrow[d, "C_{\sigma_1}"'] & (\sigma_2 \to \bot) \to \bot \arrow[d, "C_{\sigma_2}"] \\
\sigma_1 \arrow[r, "M"'] & \sigma_2
\end{tikzcd}$$

That is:

$$M\,(\boldsymbol{\mu}\alpha^{\sigma_1}.k\,(\lambda x^{\sigma_1}.[\alpha]x)) = \boldsymbol{\mu}\beta^{\sigma_2}.k\,(\lambda x^{\sigma_1}.[\beta](M\,x)) : \sigma_2$$

holds for any $k : (\sigma_1 \to \bot) \to \bot$.[2]

In any $\lambda\mu 2$-theory, focal terms compose, and the identity $\lambda x^\sigma.x$ is obviously focal. So, the (equivalence classes of) focal maps form a category. Hereafter we shall call it the *focus* of the $\lambda\mu 2$-theory.

While this characterisation of focal maps is concise and closely follows the semantic considerations in [36, 19], there is a subtle problem; the $\beta\eta$-axioms of $\lambda\mu 2$ are too weak to establish the focality of some important terms. This is because we have used the polymorphic feature of $\lambda\mu 2$ for expressing C_σ — it involves the falsity type $\bot = \forall X.X$, but the axioms of $\lambda\mu 2$ do not guarantee that $\bot$ does work properly. If there were not sufficiently many focal maps, the parametricity principle restricted on focal maps would be useless.

To see this issue more clearly, we shall look at another "classical" combinator (the Peirce law)

$$P_{\sigma_1,\sigma_2} = \lambda m.\mu\alpha^{\sigma_1}.[\alpha](m\,(\lambda x^{\sigma_1}.\mu\beta^{\sigma_2}.[\alpha]x)) \;:\; ((\sigma_1 \to \sigma_2) \to \sigma_1) \to \sigma_1$$

which does not make use of polymorphism, and the "abort" map (Ex Falso Quodlibet)

$$A_\sigma = \lambda x^\bot.x\,\sigma : \bot \to \sigma$$

which is defined without the classical feature. It is well known that the double-negation elimination is as expressible as the Peirce law together with Ex Falso Quodlibet, see e.g. [2]. This is also the case at the level of (uniformity of) proofs. Let us say that $M : \sigma_1 \to \sigma_2$ is *repeatable* if

$$\begin{tikzcd}
(\sigma_1 \to \sigma_3) \to \sigma_1 \arrow[r, "(M\to\sigma_3)\to M"] \arrow[d, "P_{\sigma_1,\sigma_3}"'] & (\sigma_2 \to \sigma_3) \to \sigma_2 \arrow[d, "P_{\sigma_2,\sigma_3}"] \\
\sigma_1 \arrow[r, "M"'] & \sigma_2
\end{tikzcd}$$

[2] In [16], a focal map from σ_1 to $\bot$ is called a "C-term of type σ_1". C-terms of type σ_1 with a free name of σ_2 correspond to focal maps from σ_1 to σ_2, thus these notions (and the associated constructions of the CPS target categories via C-terms (C-maps) [16] and via focus [36]) are essentially the same.

commutes for each σ_3; and *discardable* if

$$\begin{tikzcd} & \perp \arrow[dl, "A_{\sigma_1}"'] \arrow[dr, "A_{\sigma_2}"] & \\ \sigma_1 \arrow[rr, "M"'] & & \sigma_2 \end{tikzcd}$$

commutes.

Proposition 5.2. In a $\lambda\mu 2$-theory, $M : \sigma_1 \to \sigma_2$ is focal if and only if it is both repeatable and discardable. □

We note that the corresponding result in the call-by-value setting has been observed by Führmann [8] as the characterisation of algebraic values as repeatable discardable expressions. Here we follow his terminology.

This reformulation allows us to see that only the second diagram of A's involves the polymorphically defined $\perp$ and needs to be justified by additional conditions.[3] On the other hand, the first diagram of P's is not problematic, as it does not make use of polymorphism at all.

5.3. **Additional axioms.** To this end, we add more axioms to $\lambda\mu 2$ before thinking about parametricity. They are

(1) $\lambda x^{\sigma_1 \to \sigma_2}.x\, N : (\sigma_1 \to \sigma_2) \to \sigma_2$ is discardable for any $N : \sigma_1$
(2) $\lambda x^{\forall X.\sigma}.x\, \sigma_1 : \forall X.\sigma \to \sigma[\sigma_1/X]$ is discardable for any σ and σ_1
(3) $\lambda x^{\sigma}.[\alpha]x : \sigma \to \perp$ is discardable for any $\alpha : \sigma$

which are equivalent to asking

(1) $M\,(\sigma_1 \to \sigma_2)\, N = M\, \sigma_2$ for any $M : \perp$ and $N : \sigma_1$
(2) $M\,(\forall X.\sigma)\, \sigma_1 = M\,(\sigma[\sigma_1/X])$ for $M : \perp$
(3) $[\alpha](M\,\sigma) = M$ for $M : \perp$ and $\alpha : \sigma$ ($\alpha \notin FN(M)$)

and also equivalent to

(1) $(\boldsymbol{\mu}\alpha^{\sigma_1 \to \sigma_2}.M)\, N = \boldsymbol{\mu}\beta^{\sigma_2}.M[[\beta](-\,N)/[\alpha](-)]$
(2) $(\boldsymbol{\mu}\alpha^{\forall X.\sigma_1}.M)\, \sigma_2 = \boldsymbol{\mu}\beta^{\sigma_1[\sigma_2/X]}.M[[\beta](-\,\sigma_2)/[\alpha](-)]$
(3) $[\alpha'](\boldsymbol{\mu}\alpha^{\sigma}.M) = M[\alpha'/\alpha]$

Note that $\lambda\mu 2^P$ discussed in Section 4 satisfies these conditions. Also we shall note that $\lambda x^{\sigma_1 \to \sigma_2}.x\, N$, $\lambda x^{\forall X.\sigma}.x\, \sigma_1$, $\lambda x^{\perp}.[\alpha](x\, \sigma)$ are all repeatable in $\lambda\mu 2$. Together with these additional axioms, they become focal. (Alternatively, we could have $\perp$ as a type constant and assume the standard axiomatization of $\lambda\mu$-calculus with the falsity type [16, 36] — in that case C is defined without polymorphism, and this problem disappears.)

Below we develop the focal parametricity principle on top of $\lambda\mu 2$ with these additional axioms.

---

[3] This problem was overlooked in the preliminary version of this paper [13] where it was wrongly assumed that repeatability alone would imply focality.

5.4. **A parametricity principle for $\boldsymbol{\lambda\mu 2}$.** Given a focal $f : \sigma_1 \to \sigma_2$ we define its graph relation $\langle f \rangle : \sigma_1 \leftrightarrow \sigma_2$ by $u \langle f \rangle v$ iff $f\,u = v$. Also, let $id_\sigma : \sigma \leftrightarrow \sigma$ be the identity relation on the terms of type σ. In this paper we only consider the relations given by the graphs of focal maps, identity, and σ^*'s obtained by the following construction, which we shall call "focal relations".

Given a type σ whose free type variables are included in $X_1, \dots, X_n$ and focal relations $s_1 : \sigma_1 \leftrightarrow \sigma'_1, \dots, s_n : \sigma_n \leftrightarrow \sigma'_n$, we define a focal relation σ^* as follows.

- $X_i^* = s_i : \sigma_i \leftrightarrow \sigma'_i$
- $(\sigma \to \sigma')^* : (\sigma \to \sigma')[\sigma_1/X_1, \dots] \leftrightarrow (\sigma \to \sigma')[\sigma'_1/X_1, \dots]$ is the relation so that $f\,(\sigma \to \sigma')^*\,g$ iff $x\,\sigma^*\,y$ implies $(f\,x)\,\sigma'^*\,(g\,y)$
- $(\forall X.\sigma)^* : \forall X.\sigma[\sigma_1/X_1, \dots] \leftrightarrow \forall X.\sigma[\sigma'_1/X_1, \dots]$ is the relation so that $u\,(\forall X.\sigma)^*\,v$ iff $(u\,\sigma')\,\sigma[r/X]^*\,(v\,\sigma'')$ holds for any focal relation $r : \sigma' \leftrightarrow \sigma''$

The focal relational parametricity asserts that, for any σ whose free type variables are included in $X_1, \dots, X_n$, $M : \sigma[\sigma_1/X_1, \dots, \sigma_n/X_n]$ implies $M\,\sigma^*\,M$ with $s_i = id_{\sigma_i}$.

Thus the only departure from the standard parametricity principle is the condition that the graph relation construction is allowed only on focal maps. Note that this restriction is necessary; if we apply parametricity to polymorphic terms $\Lambda X.C_X$ or $\Lambda X.P_{X,\sigma}$, we will get the naturality diagrams above for any term which is allowed to be used for the graph relation construction.

5.5. **On consistency and soundness.** The consistency of focal parametricity (in the sense that the equational theory of $\lambda\mu 2$ with focal parametricity is not trivial) follows from the fact that there are non-trivial parametric models of $\lambda 2$ in which there is an object R so that the continuation monad $T\tau = R^{R^\tau}$ satisfies the "equalising requirement" [22], i.e. each component $\eta_\tau : \tau \to T\tau$ of its unit is an equaliser of $\eta_{T\tau}$ and $T\eta_\tau$. (Here we employ the syntax of the CPS target calculus as an internal language for such models, where the CPS translation is considered to give a semantic interpretation.) In such models, for any focal term $f : \sigma_1 \to \sigma_2$, there exists a unique $y : \sigma_2^\circ \vdash g(y) : \sigma_1^\circ$ such that $[\![f\,x]\!] = \lambda y.x\,(g(y))$ (cf. [36]).

Using this fact, given a focal relation $r : \sigma_1 \leftrightarrow \sigma_2$, we construct an admissible relation $r^\circ : \sigma_2^\circ \leftrightarrow \sigma_1^\circ$ as follows. For a graph relation $\langle f \rangle : \sigma_1 \leftrightarrow \sigma_2$, we let $\langle f \rangle^\circ = \langle g \rangle : \sigma_2^\circ \leftrightarrow \sigma_1^\circ$ where g is the unique map as given above. For σ^*, $\sigma^{*\circ}$ is defined by straightforward induction: $(\sigma \to \sigma')^{*\circ} = \neg\sigma^{*\circ} \wedge \sigma'^{*\circ}$, $(\forall X.\sigma)^{*\circ} = \exists X.\sigma^{*\circ}$ (where the parameter relations s_i are replaced by s_i°).

Theorem 5.3. In such a model, given a focal relation $r : \sigma_1 \leftrightarrow \sigma_2$, $M\,r\,N$ implies $[\![N]\!]\,\neg r^\circ\,[\![M]\!]$. □

Theorem 5.4 (consistency)**.** Focal parametricity is consistent. □

We do not know if the term model of the parametric target calculus satisfies the equalising requirement — if so, by the definability result, the parametricity on the target and the focal parametricity on $\lambda\mu 2$ should agree. Alternatively we should consider a refined target calculus with a construct ensuring the equalising requirement, as detailed in Taylor's work on sober space ("a lambda calculus for sobriety" [41]). For now, we only know that one direction is true (thanks to the definability).

Theorem 5.5. An equality derivable in $\lambda\mu 2^P$ is also derivable in $\lambda\mu 2$ with focal parametricity. □

6. Examples

We show that certain impredicative encodings in $\lambda\mu 2$ satisfy universal properties with respect to the focus using the focal parametricity principle.

6.1. Focal decomposition. We start with a remark on the following "focal decomposition" [36] (analogous to the linear decomposition $\sigma_1 \to \sigma_2 = !\sigma_1 \multimap \sigma_2$ [10]): there is a bijective correspondence between terms of $\sigma_1 \to \sigma_2$ and focal terms of $\neg\neg\sigma_1 \to \sigma_2$ natural in σ_1 and focal σ_2.

$$\frac{f : \neg\neg\sigma_1 \to \sigma_2 \text{ focal}}{f^\flat = f \circ \eta_{\sigma_1} = \lambda x^{\sigma_1}.f\,(\lambda k.k\,x) : \sigma_1 \to \sigma_2}$$

$$\frac{g : \sigma_1 \to \sigma_2}{g^\sharp = C_{\sigma_2} \circ \neg\neg g = \lambda m.\boldsymbol{\mu}\beta^{\sigma_2}.m\,(\lambda x^{\sigma_1}.[\beta](g\,x)) : \neg\neg\sigma_1 \to \sigma_2 \text{ focal}}$$

Proposition 6.1. $g^{\sharp^\flat} = g$ for any $g : \sigma_1 \to \sigma_2$, while $f^{\flat^\sharp} = f$ holds for $f : \neg\neg\sigma_1 \to \sigma_2$ if and only if f is focal. $\square$

6.2. Falsity as a focally initial object. Now we shall proceed to reason about impredicative encodings in $\lambda\mu 2$. The first example is the falsity $\bot = \forall X.X$.

First, we note that $A_\sigma = \lambda x^\bot.x\,\sigma : \bot \to \sigma$ is focal. The parametricity on $\bot$ says $x\,\bot^*\,x$ for any $x : \bot$. Since $\langle A_\sigma\rangle : \bot \leftrightarrow \sigma$, we have $x\,\bot\,\langle A_\sigma\rangle\,x\,\sigma$, i.e. $A_\sigma\,(x\,\bot) = x\,\bot\,\sigma = x\,\sigma$. By extensionality we get $x = x\,\bot$ for $x : \bot$.

Now suppose that $g : \bot \to \sigma$ is focal. Again by the parametricity on $\bot$ we know $x\,\bot^*\,x$ for any $x : \bot$, hence $x\,\bot\,\langle g\rangle\,x\,\sigma$. Thus $g\,(x\,\bot) = x\,\sigma$; but $x = x\,\bot$, so we have $g\,x = x\,\sigma$, hence $g = \lambda x^\bot.x\,\sigma = A_\sigma$.

So we conclude that A_σ is the unique focal map from $\bot$ to σ. This means that $\bot$ is initial in the focus.

6.3. Focally initial algebra. As in $\lambda 2$, there is a fairly standard encoding

$$\begin{array}{rcl} \mu X.F[X] & = & \forall X.(F[X] \to X) \to X \\ \mathsf{fold}_\sigma & = & \lambda a^{F[\sigma]\to\sigma}.\lambda x^{\mu X.F[X]}.x\,\sigma\,a : (F[\sigma] \to \sigma) \to \mu X.F[X] \to \sigma \\ \mathsf{in} & = & \lambda y.\Lambda X.\lambda k^{F[X]\to X}.k\,(F[\mathsf{fold}_X\,k]\,y) : F[\mu X.F[X]] \to \mu X.F[X] \end{array}$$

for which the following diagram commutes (just by β-axioms).

$$\begin{tikzcd} F[\mu X.F[X]] \arrow[r, "\mathsf{in}"] \arrow[d, "F[\mathsf{fold}\,a]"'] & \mu X.F[X] \arrow[d, "\mathsf{fold}\,a"] \\ F[\sigma] \arrow[r, "a"'] & \sigma \end{tikzcd}$$

Therefore in is a *weak* initial F-algebra. However, as we noted before, in is not an initial F-algebra — in fact it is not even an isomorphism. By applying the focal decomposition above, we obtain the commutative diagram

$$\begin{tikzcd}
\neg\neg F[\mu X.F[X]] \arrow[r, "\mathsf{in}^\sharp"] \arrow[d, "\neg\neg F[\mathsf{fold}\, a^\flat]"'] & \mu X.F[X] \arrow[d, "\mathsf{fold}\, a^\flat"] \\
\neg\neg F[\sigma] \arrow[r, "a"'] & \sigma
\end{tikzcd}$$

for any focal $a : \neg\neg F[\sigma] \to \sigma$. We show that $\mathsf{fold}\, a^\flat$ is the unique *focal* map making this diagram commute, thus $\mathsf{in}^\sharp$ is an initial $\neg\neg F[-]$-algebra in the focus.

We sketch a proof which is fairly analogous to that for the corresponding result in parametric $\lambda 2$ as given in [1]. First, from the parametricity on $\mu X.F[X]$ we obtain that

$$\begin{tikzcd}
F[\sigma_1] \arrow[r, "a"] \arrow[d, "F[h]"'] & \sigma_1 \arrow[d, "h"] \\
F[\sigma_2] \arrow[r, "b"'] & \sigma_2
\end{tikzcd}
\quad \text{implies} \quad
\begin{tikzcd}
\mu X.F[X] \arrow[r, "\mathsf{fold}\, a"] \arrow[d, "||"'] & \sigma_1 \arrow[d, "h"] \\
\mu X.F[X] \arrow[r, "\mathsf{fold}\, b"'] & \sigma_2
\end{tikzcd}$$

whenever h is focal. We also have $M\,(\mu X.F[X])\,\mathsf{in} = M$ for any $M : \mu X.F[X]$ as a corollary (thanks to extensionality). By combining these observations, now we have the desired result. That is, if $h : \mu X.F[X] \to \sigma$ is focal and satisfies $h \circ \mathsf{in}^\sharp = a \circ \neg\neg F[h]$, then

$$\begin{array}{rcl}
\mathsf{fold}_\sigma\, a^\flat\, x & = & h\,(\mathsf{fold}_{\mu X.F[X]}\,\mathsf{in}\, x) \\
& = & h\,(x\,(\mu X.F[X])\,\mathsf{in}) \\
& = & h\, x
\end{array}$$

so by extensionality we conclude $\mathsf{fold}_\sigma\, a^\flat = h$. This also implies that $\mathsf{in}^\sharp$ is an isomorphism, with the inverse given by $\mathsf{fold}_{\neg\neg\mu X.F[X]}\,(\neg\neg F[\mathsf{in}])$.

As a special case, by letting F be a constant functor, we obtain isomorphisms between $(\sigma \to \bot) \to \bot$ and $\forall X.(\sigma \to X) \to X$ where X is not free in σ. With some further calculation we see that $\mathsf{in}^\sharp = \lambda m.\Lambda X.\lambda k^{\sigma\to X}.\boldsymbol{\mu}\alpha^X.m\,(\lambda x^\sigma.[\alpha](k\,x))$ is the inverse of $\lambda n.n\,\bot : (\forall X.(\sigma \to X) \to X) \to (\sigma \to \bot) \to \bot$. We will see more about this isomorphism in Section 7.

6.4. **The type of Church numerals.** We conclude this section by a remark on the type of Church numerals $\mathbf{N} = \forall X.X \to (X \to X) \to X$. Recall that, in $\lambda 2$ with parametricity, $\mathbf{N}$ is an initial algebra of $\forall X.X \to (- \to X) \to X \simeq 1 + (-)$, i.e. a natural numbers object, whose closed inhabitants are equal to the Church numerals $\mathsf{S}^n\,\mathsf{O}$ which can be given by, as usual,

$$\begin{array}{rclcl}
\mathsf{O} & = & \Lambda X.\lambda x^X f^{X\to X}.x & : & \mathbf{N} \\
\mathsf{S} & = & \lambda n^{\mathbf{N}}.\Lambda X.\lambda x^X f^{X\to X}.f\,(n\,X\,x\,f) & : & \mathbf{N} \to \mathbf{N}
\end{array}$$

It is no longer true in $\lambda\mu 2$, as observed by Parigot, as there are closed inhabitants which are not equal to Church numerals, e.g.

$$\mu\alpha^{\mathbf{N}}.[\alpha](\mathsf{S}\,(\mu\beta^{\mathbf{N}}.[\alpha]\mathsf{O})) = \Lambda X.\lambda x^X f^{X\to X}.\mu\alpha^X.[\alpha](f\,(\mu\beta^X.[\alpha]x)) : \mathbf{N}$$

In contrast, $\mathbf{N}$ in $\lambda\mu 2$ with focal parametricity is a focally initial algebra of $\forall X.X \to (- \to X) \to X \simeq \bot \to (- \to \bot) \to \bot$; this can be shown in the same way as the case of focally initial algebras. Spelling this out, we have a focal map $\mathsf{in} : (\bot \to (\mathbf{N} \to \bot) \to \bot) \to \mathbf{N}$, and for any focal $g : (\bot \to (\sigma \to \bot) \to \bot) \to \sigma$ there exists a unique focal $\mathsf{fold}_\sigma\, g : \mathbf{N} \to \sigma$ making the following diagram commute.

$$\begin{tikzcd}
\bot \to (\mathbf{N} \to \bot) \to \bot \arrow[r, "\mathsf{in}"] \arrow[d, "\bot\to(\mathsf{fold}_\sigma\, g \to \bot)\to\bot"'] & \mathbf{N} \arrow[d, "\mathsf{fold}_\sigma\, g"] \\
\bot \to (\sigma \to \bot) \to \bot \arrow[r, "g"'] & \sigma
\end{tikzcd}$$

To see this, it is useful to observe the following bijective correspondence (a variant of the focal decomposition): given focal $g : (\bot \to (\sigma \to \bot) \to \bot) \to \sigma$ we have

$$\begin{array}{lcllcl} g_o & = & g\,(\lambda x^{\bot} k^{\sigma\to\bot}.x) & : & \sigma \\ g_s & = & \lambda y^{\sigma}.g\,(\lambda x^{\bot} k^{\sigma\to\bot}.k\,y) & : & \sigma \to \sigma \end{array}$$

and conversely, for $a : \sigma$ and $f : \sigma \to \sigma$ we have a focal map

$$\varphi_{a,f} \;=\; \lambda m^{\bot\to(\sigma\to\bot)\to\bot}.\mu\alpha^{\sigma}.m\,([\alpha]a)\,(\lambda y^{\sigma}.[\alpha](f\,y)) \;:\; (\bot \to (\sigma \to \bot) \to \bot) \to \sigma$$

It follows that $(\varphi_{a,f})_o = a$ and $(\varphi_{a,f})_s = f$ hold for any a and f, while $\varphi_{g_o,g_s} = g$ for any focal g. Now we define

$$\begin{array}{lcllcl} \mathsf{fold}_A\, g & = & \lambda n^{\mathbf{N}}.n\,A\,g_o\,g_s & : & \mathbf{N} \to A \\ \mathsf{in} & = & \varphi_{\mathsf{O},\mathsf{S}} & : & (\bot \to (\mathbf{N} \to \bot) \to \bot) \to \mathbf{N} \end{array}$$

It then follows that the diagram above commutes — and the focal parametricity implies that $\mathsf{fold}_A\, g$ is the unique such focal map.

7. A general characterisation

So far, we concentrated on the relational parametricity for $\lambda\mu2$. One may feel that this story is very specific to the case of $\lambda\mu2$, or of the first-class continuations, and is not immediately applicable to other computational effects.

In this section we describe an alternative characterisation of the focus, which makes sense in any extension of $\lambda2$. Namely, we show that, any $\lambda2$-theory is equipped with a monad L, such that each type is equipped with an algebra structure — and then see that, in the case of $\lambda\mu2$ with focal parametricity, this monad L is isomorphic to the double-negation (continuation) monad, and focal maps are precisely the algebra maps of the monad L. This suggests a natural generalisation of this work to a theory of parametricity for general computational effects.

7.1. **A monad on $\lambda2$.** Let $L\sigma = \forall X.(\sigma \to X) \to X$ (with no free X in σ), and define

$$\begin{array}{lcllcl} \eta_\sigma & = & \lambda x^{\sigma}.\Lambda X.\lambda k^{\sigma\to X}.k\,x & : & \sigma \to L\sigma \\ \mu_\sigma & = & \lambda z^{L^2\sigma}.\Lambda X.\lambda k^{\sigma\to X}.z\,X\,(\lambda y^{L\sigma}.y\,X\,k) & : & L^2\sigma \to L\sigma \\ L(f) & = & \lambda y^{L\sigma_1}.\Lambda X.\lambda h^{\sigma_2\to X}.y\,X\,(h\circ f) & : & L\sigma_1 \to L\sigma_2\ (f : \sigma_1 \to \sigma_2) \end{array}$$

Proposition 7.1. On the term model of any $\lambda2$-theory, (L, η, μ) forms a monad. $\square$

One might think that this is trivial as $L\sigma$ is isomorphic to σ when we assume the standard parametricity. This is not always the case however, as we have already seen, $L\sigma \simeq (\sigma \to \bot) \to \bot$ in the focally parametric $\lambda\mu 2$.

Proposition 7.2. $\alpha_\sigma = \lambda y^{L\sigma}.y\,\sigma\,(\lambda x^\sigma.x) : L\sigma \to \sigma$ is an algebra of the monad (L, η, μ). □

Thus each σ is canonically equipped with an algebra structure α_σ. Again one may think that this is trivial, as under the standard parametricity α_σ is just an isomorphism with η_σ being an inverse. However, again it is not the case in a non-trivial $\lambda\mu 2$-theory.

Now we define the notion of linear maps in terms of the monad L and the canonical algebras α_σ — this is close to what we do in (axiomatic) domain theory for characterising the strict maps, and also in control categories for characterising the focal maps.

Definition 7.3. *$f : \sigma_1 \to \sigma_2$ is* linear *when it is an algebra morphism from α_{σ_1} to α_{σ_2}, i.e. $f \circ \alpha_{\sigma_1} = \alpha_{\sigma_2} \circ L(f)$ holds.*

$$\begin{tikzcd} L\sigma_1 \arrow[r, "L(f)"] \arrow[d, "\alpha_{\sigma_1}"'] & L\sigma_2 \arrow[d, "\alpha_{\sigma_2}"] \\ \sigma_1 \arrow[r, "f"'] & \sigma_2 \end{tikzcd}$$

That is, f is linear when

$$f\,(M\,\sigma_1\,(\lambda x^{\sigma_1}.x)) = M\,\sigma_2\,f$$

holds for any $M : L\sigma_1$. We may write $f : \sigma_1 \multimap \sigma_2$ for a linear $f : \sigma_1 \to \sigma_2$. Under the standard parametricity every $f : \sigma_1 \to \sigma_2$ is linear, while for focal parametricity on $\lambda\mu 2$ we have that linear maps are precisely the focal maps (see below). In passing, we note the following interesting observation.

Proposition 7.4. In a $\lambda 2$-theory, the following conditions are equivalent.

(1) algebras on $\sigma_1 \to \sigma_2$ and $\forall X.\sigma$ are determined in the pointwise manner, i.e.

$$\begin{array}{rcl} \alpha_{\sigma_1\to\sigma_2} & = & \lambda f^{L(\sigma_1\to\sigma_2)}.\lambda x^{\sigma_1}.\alpha_{\sigma_2}\,(L(\lambda g^{\sigma_1\to\sigma_2}.g\,x)\,f) \\ \alpha_{\forall X.\sigma} & = & \lambda x^{L(\forall X.\sigma)}.\Lambda X.\alpha_\sigma\,(L(\lambda y^{\forall X.\sigma}.y\,X)\,x) \end{array}$$

(2) $\lambda x^{\sigma_1\to\sigma_2}.x\,N$ is linear for any $N : \sigma_1$, and $\lambda x^{\forall X.\sigma}.x\,\sigma_1$ is linear for any σ and σ_1. □

Note that they are very close to the "additional axioms" for $\lambda\mu 2$ discussed in Section 5. Also note that, if a $\lambda 2$-theory satisfies one of these conditions, μ_σ and $\alpha_{L\sigma}$ agree for every σ. and we have a "linear decomposition" correspondence between the maps of $\sigma_1 \to \sigma_2$ and the linear maps of $L\sigma_1 \to \sigma_2$.

7.2. **Focal maps as algebra maps.** Now we shall consider the double-negation monad $\neg\neg\sigma = (\sigma \to \bot) \to \bot$ on $\lambda\mu 2$ with focal parametricity.

Proposition 7.5. In a focally parametric $\lambda\mu 2$-theory, $C_\sigma : \neg\neg\sigma \to \sigma$ is an algebra of the double-negation monad. □

Corollary 7.6. *$f : \sigma_1 \to \sigma_2$ is focal if and only if it is an algebra map from C_{σ_1} to C_{σ_2}.* □

Proposition 7.7. The monad (L, η, μ) is isomorphic to the double negation monad in the focally parametric $\lambda\mu 2$, with $\lambda x^{L\sigma}.x\,\bot : L\sigma \xrightarrow{\simeq} \neg\neg\sigma$. □

Proposition 7.8. The following diagram commutes in a focally parametric $\lambda\mu 2$-theory:

$$\begin{tikzcd} L\sigma \arrow[rr, "\lambda x^{L\sigma}.x\,\bot"] \arrow[dr, "\alpha_\sigma"'] & & \neg\neg\sigma \arrow[dl, "C_\sigma"] \\ & \sigma & \end{tikzcd}$$

□

Corollary 7.9. *$f : \sigma_1 \to \sigma_2$ is linear if and only if it is focal.* □

Thus a focal map in a focally parametric $\lambda\mu 2$-theory can be characterised just in terms of the monad L which is defined for arbitrary $\lambda 2$-theory.

We believe that the monad L deserves much attention. It has been considered trivial, but now we know that it does characterise an essential notion (focus) in the case of relational parametricity under the presence of control feature. In fact, the story does not end here; under the presence of non-termination or recursion, L behaves like a *lifting monad* — indeed it is a lifting in the theory of linear parametricity, because

$$L\sigma \;=\; \forall X.(\sigma \to X) \to X \;=\; \forall X.!(!\sigma \multimap X) \multimap X \;\simeq\; !\sigma$$

where the last isomorphism follows from the fact that $\forall X.!(F[X] \multimap X) \multimap X$ gives an initial algebra of F, cf. [4].

These observations suggest that there exists a general framework similar to (axiomatic or synthetic) domain theory where the lifting monad can be replaced by any strong monad — a continuation monad for example — on which a theory of parametricity for general computational effects can be built. Recently, Alex Simpson has made a progress in this direction, by developing a two-level polymorphic type theory (for interpreting types and algebras of a monad) in a constructive universe [38]. His work fits very well with the case of linear parametricity for recursion; it is plausible that it also explains the case of focal parametricity for first-class control.

8. Conclusion and future work

We have studied the relational parametricity for $\lambda\mu 2$, first by considering the CPS translation into a parametric fragment of $\lambda 2$, and then by directly giving a constrained parametricity for $\lambda\mu 2$. The later, which we call "focal parametricity", seems to be a natural parametricity principle under the presence of first-class controls — in the same sense that linear parametricity works under the presence of recursion and non-termination.

There remain many things to be addressed in future. In the previous section, we already discussed a research direction towards a relational parametricity for general effects. Below we shall briefly mention some future work more closely related to the main development of this paper.

Firstly, we are yet to complete the precise comparison between focal parametricity on $\lambda\mu 2$ and the parametricity on the CPS target calculus. This involves some subtle interaction between parametricity and a technical condition (equalising requirement).

Secondly, we should study focal parametricity for extensions of $\lambda\mu 2$. As we observed, $\lambda\mu 2$ with focal parametricity does not have many popular datatypes, e.g. cartesian products, and classical disjunction types which however can be added with no problem. Adding general initial algebras is problematic (having an initial object already means inconsistency), but it might be safe to add certain carefully chosen instances. On the other hand, final coalgebras seem less problematic, though a generic account for them in $\lambda\mu 2$ is still missing.

Perhaps we also need to consider the CPS translation of such datatypes (cf. [3]) in a systematic way.

An interesting topic we have not discussed in this paper is the Filinski-Selinger duality [6, 36] between call-by-name and call-by-value calculi with control primitives. In fact it is straightforward to consider its second-order extension: in short, universal quantifiers in call-by-name (as studied in this paper) amount to existential quantifiers in call-by-value. We are not sure if the call-by-value calculus with existential quantifiers itself is of some interest. However, it can be a good starting point to understand the call-by-value parametric polymorphism (possibly with computational effects), from both syntactic and semantic aspects. In particular, it should provide new insights on the famous difficulty of accommodating first-class continuations in ML type system [11].

Finally, we also should consider if there is a better (ideally semantic) formulation of focal relations. In this paper we only consider those coming from focal maps, but it seems natural to regard a subalgebra (of the double-negation monad) of $C_{\sigma_1 \times \sigma_2}$ as a focal relation between σ_1 and σ_2, where we assume the presence of cartesian product $\sigma_1 \times \sigma_2$. This looks very closely related to Pitts' $\top\top$-closed relations for $\lambda 2$ with recursion [28].

Acknowledgement

I thank Ken-etsu Fujita for discussions and cooperations related to this work. I am also grateful to Ryu Hasegawa, Paul-André Melliès and Alex Simpson for comments, discussions and encouragements.

References

[1] M. Abadi, L. Cardelli, and P.-L. Curien. Formal parametric polymorphism. *Theoret. Comput. Sci.* 121:9–58, 1993.

[2] Z.M. Ariola and H. Herbelin. Minimal classical logic and control operators. In *Proc. Automata, Languages and Programming*, Springer Lecture Notes in Comput. Sci. 2719, pp.871–885, 2003.

[3] G. Barthe and T. Uustalu. CPS translating inductive and coinductive types. In *Proc. Partial Evaluation and Semantics-Based Program Manipulation*, pp.131–142, 2002.

[4] G.M. Bierman, A.M. Pitts, and C.V. Russo. Operational properties of Lily, a polymorphic linear lambda calculus with recursion. In *Proc. Higher Order Operational Techniques in Semantics*, Electronic Notes in Theoretical Computer Science 41, 2000.

[5] B.P. Dunphy. *Parametricity as a Notion of Uniformity in Reflexive Graphs.* PhD thesis, University of Illinois, 2002.

[6] A. Filinski. Declarative continuations: an investigation of duality in programming language semantics. In *Proc. Category Theory and Computer Science*, Springer Lecture Notes in Comput. Sci. 389, pp.224–249, 1989.

[7] K. Fujita. Galois embedding from polymorphic types into existential types. In *Proc. Typed Lambda Calculi and Applications*, Springer Lecture Notes in Comput. Sci. 3461, pp.194–208, 2005.

[8] C. Führmann. Varieties of effects. In *Proc. Foundations of Software Science and Computation Structures*, Springer Lecture Notes in Comput. Sci. 2303, pp.144–159, 2002.

[9] J.-Y. Girard. *Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur.* Thèse d'Etat, Université Paris VII, 1972.

[10] J.-Y. Girard. Linear logic. *Theoret. Comp. Sci.* 50:1–102, 1987.

[11] B. Harper, B.F. Duba and D. MacQueen. Typing first-class continuations in ML. *J. Funct. Programming* 3(4):465–484.

[12] M. Hasegawa. Semantics of linear continuation-passing in call-by-name. In *Proc. Functional and Logic Programming*, Springer Lecture Notes in Comput. Sci. 2998, pp.229–243, 2004.

[13] M. Hasegawa, Relational parametricity and control (extended abstract). In *Proc. Logic in Computer Science*, pp.72–81, 2005.
[14] M. Hasegawa and Y. Kakutani. Axioms for recursion in call-by-value. *Higher-Order and Symbolic Comput.* 15(2/3):235–264, 2002.
[15] R. Hasegawa. Categorical data types in parametric polymorphism. *Math. Struct. Comp. Sci.* 4(1):71–109, 1994.
[16] M. Hofmann and T. Streicher. Completeness of continuation models for $\lambda\mu$-calculus. *Inf. Comput.* 179(2):332–355, 2002.
[17] B. Jacobs. *Categorical Logic and Type Theory.* Elsevier, 1999.
[18] Y. Kakutani. Duality between call-by-name recursion and call-by-value iteration. In *Proc. Computer Science Logic*, Springer Lecture Notes in Comput. Sci. 2471, pp.506–521, 2002.
[19] Y. Kakutani and M. Hasegawa. Parameterizations and fixed-point operators on control categories. *Fundam. Inform.* 65(1/2):153–172, 2005.
[20] J. Lambek and P. Scott (1986) *Introduction to Higher-order Categorical Logic*. Cambridge University Press.
[21] R. Matthes. Parigot's second order $\lambda\mu$-calculus and inductive types. In *Proc. Typed Lambda Calculi and Applications*, Springer Lecture Notes in Comput. Sci. 2044, pp.329–343, 2001.
[22] E. Moggi. Notions of computation and monads. *Inf. Comput.* 93(1):55–92, 1991.
[23] K. Nakazawa and M. Tatsuta. Strong normalization proof with CPS-translation for second order classical natural deduction. *J. Symb. Log.* 68(3):851–859, 2003. Corrigendum: *J. Symb. Log.* 68(4):1415–1416, 2003.
[24] C.-H.L. Ong. A semantic view of classical proofs: type-theoretic, categorical, and denotational characterizations (preliminary extended abstract). In *Proc. Logic in Computer Science*, pp.230–241, 1996.
[25] C.-H.L. Ong and C.A. Stewart. A Curry-Howard foundation for functional computation with control. In *Proc. Principles of Programming Languages*, pp.215–227, 1997.
[26] M. Parigot. $\lambda\mu$-calculus: an algorithmic interpretation of classical natural deduction. In *Proc. Logic Programming and Automated Reasoning*, Springer Lecture Notes in Comput. Sci. 624, pp.190–201, 1992.
[27] M. Parigot. Proofs of strong normalisation for second order classical natural deduction. *J. Symb. Log.* 62(4):1461–1479, 1997.
[28] A. Pitts, Parametric polymorphism and operational equivalence. *Math. Struct. Comp. Sci.* 10:321–359, 2000.
[29] G.D. Plotkin. Call-by-name, call-by-value, and the λ-calculus. *Theoret. Comput. Sci.* 1(1):125–159, 1975.
[30] G.D. Plotkin. Type theory and recursion (extended abstract). In *Proc. Logic in Computer Science*, pp.374, 1993.
[31] G.D. Plotkin and M. Abadi. A logic for parametric polymorphism. In *Proc. Typed Lambda Calculi and Applications*, Springer Lecture Notes in Comput. Sci. 664, pp.361–375, 1993.
[32] A.J. Power and E. Robinson. Premonoidal categories and notions of computation. *Math. Struct. Comp. Sci.* 7(5):453–468, 1997.
[33] J.C. Reynolds. Towards a theory of type structure. In *Proc. Colloque sur la Programmation*, Springer Lecture Notes in Comput. Sci. 19, pp.408–425, 1974.
[34] J.C. Reynolds. Types, abstraction and parametric polymorphism. In *Proc. IFIP 9th World Computer Congress*, Information Processing 83, pp.513–523, 1983.
[35] A. Sabry. Note on axiomatizing the semantics of control operators. Tech. Rep. CIS-TR-96-03, University of Oregon, 1996.
[36] P. Selinger. Control categories and duality: on the categorical semantics of the lambda-mu calculus. *Math. Struct. Comp. Sci.* 11(2):207–260, 2001.
[37] P. Selinger. Some remarks on control categories. Manuscript, 2003.
[38] A. Simpson. Relational parametricity for computational effects. Manuscript, 2006.
[39] T. Streicher and B. Reus. Classical logic, continuation semantics and abstract machines. *J. Funct. Program.* 8(6):543–572, 1998.
[40] I. Takeuti. An axiomatic system of parametricity. *Fundam. Inform.* 33(4):397–432, 1998.
[41] P. Taylor. Sober spaces and continuations. *Theory and Applications of Categories* 10(12):248–300, 2002.

[42] H. Thielecke. *Categorical Structure of Continuation Passing Style*. PhD thesis, University of Edinburgh, 1997.

[43] H. Thielecke. Answer type polymorphism in call-by-name continuation passing. In *Proc. European Symposium on Programming*, Springer Lecture Notes in Comput. Sci. 2986, pp.279–293, 2004.

[44] P. Wadler. Theorems for free! In *Proc. Functional Programming Languages and Computer Architecture*, pp.347–359, 1989.