INFORMATION AND COMPUTATION 93, 55–92 (1991)

Notions of Computation and Monads

EUGENIO MOGGI*

Department of Computer Science, University of Edinburgh, Edinburgh EH9 3JZ, UK

The λ-calculus is considered a useful mathematical tool in the study of programming languages, since programs can be *identified* with λ-terms. However, if one goes further and uses $\beta\eta$-conversion to prove equivalence of programs, then a gross simplification is introduced (programs are identified with total functions from *values* to *values*) that may jeopardise the applicability of theoretical results. In this paper we introduce calculi, based on a categorical semantics for *computations*, that provide a correct basis for proving equivalence of programs for a wide range of *notions of computation*. © 1991 Academic Press, Inc.

INTRODUCTION

This paper is about logics for reasoning about programs, in particular for proving equivalence of programs. Following a consolidated tradition in theoretical computer science we identify programs with the closed λ-terms, possibly containing extra constants, corresponding to some features of the programming language under consideration. There are three semantics-based approaches to proving equivalence of programs:

- The *operational* approach starts from an *operational semantics*, e.g., a partial function mapping every program (i.e., closed term) to its resulting value (if any), which induces a congruence relation on open terms called *operational equivalence* (see e.g. Plotkin (1975)). Then the problem is to prove that two terms are operationally equivalent.

- The *denotational* approach gives an interpretation of the (programming) language in a mathematical structure, the *intended model*. Then the problem is to prove that two terms denote the same object in the intended model.

- The *logical* approach gives a class of *possible models* for the (programming) language. Then the problem is to prove that two terms denote the same object in all possible models.

The operational and denotational approaches give only a theory: the operational equivalence $\approx$ or the set *Th* of formulas valid in the intended model, respectively. On the other hand, the logical approach gives a conse-

* Research partially supported by EEC Joint Collaboration Contract ST2J-0374-C(EDB).

0890-5401/91 $3.00

quence relation $\vdash$, namely $Ax \vdash A$ iff the formula A is true in all models of the set of formulas Ax, which can deal with different programming languages (e.g. functional, imperative, non-deterministic) in a rather *uniform* way, by simply changing the set of axioms Ax, and possibly extending the language with new constants. Moreover, the relation $\vdash$ is often semidecidable, so it is possible to give a sound and complete formal system for it, while Th and $\approx$ are semidecidable only in oversimplified cases.

We do not take as a starting point for proving equivalence of programs the theory of $\beta\eta$-conversion, which identifies the denotation of a program (procedure) of type $A \to B$ with a total function from A to B, since this identification wipes out completely behaviours such as non-termination, non-determinism, and side-effects, that can be exhibited by real programs. Instead, we proceed as follows:

1. We take category theory as a general theory of functions and develop on top a *categorical semantics of computations* based on monads.

2. We consider simple formal systems matching the categorical semantics of computation.

3. We extend stepwise categorical semantics and formal system in order to interpret richer languages, in particular the λ-calculus.

4. We show that w.l.o.g. one may consider only (monads over) toposes, and we exploit this fact to establish conservative extension results.

The methodology outlined above is inspired by Scott (1980)[1], and it is followed in Rosolini (1986) and Moggi (1986) to obtain the λ_p-calculus. The view that "category theory comes, logically, before the λ-calculus" led us to consider a categorical semantics of computations first, rather than to modify the rules of $\beta\eta$-conversion directly to get a *correct* calculus.

Related Work

The operational approach to finding *correct* λ-calculi w.r.t. an operational equivalence was first considered in Plotkin (1975) for call-by-value and call-by-name operational equivalence. This approach was later extended, following a similar methodology, to other features of computations such as nondeterminism (see Sharma (1984)) side-effects, and continuations (see Felleisen *et al.* (1986, 1989)). The calculi based only on operational considerations, such as the λ_v-calculus, are sound and complete

[1] "I am trying to find out where λ-calculus *should* come from, and the fact that the notion of a cartesian closed category is a late developing one (Eilenberg and Kelly, 1966), is not relevant to the argument: I shall try to explain in my own words in the next section why we should look to it first."

w.r.t. the operational semantics, i.e., a program M has a value according to the operational semantics iff it is provably equivalent to a value (not necessarily the same) in the calculus, but they are too weak for proving equivalences of programs.

Previous work on axiom systems for proving equivalence of programs with side effects has shown the importance of the *let*-constructor (see Mason (1988) and Mason and Talcott (1989a, b)). In the framework of the computational lambda-calculus the importance of *let* becomes even more apparent.

The denotational approach may suggest important principles, e.g. fix-point induction (see Scott (1969) and Gordon, Milner, and Wadsworth (1979)), that can be found only after a semantics is developed based on mathematical structures rather than term models, but it does not give clear criteria to single out the general principles among the properties satisfied by the model. Moreover, the theory at the heart of Denotational Semantics, Domain Theory (see Gunter and Scott (1989) and Mosses (1989)), has focused on the mathematical structures for giving semantics to recursive definitions of types and functions (see Smith and Plotkin (1982)), while other structures, that might be relevant to a better understanding of programming languages, have been overlooked. This paper identifies one such structure, *monads*, but probably there are others just waiting to be discovered.

The categorical semantics of computations presented in this paper has been strongly influenced by the reformulation of Denotational Semantics based on the category of cpos, possibly without bottom, and partial continuous functions (see Plotkin (1985)) and the work on categories of partial morphisms in Rosolini (1986) and Moggi (1986). Our work generalises the categorical account of partiality to other notions of computation; indeed *partial cartesian closed categories* turn out to be a special case of λ_c-models (see Definition 3.9).

A type theoretic approach to partial functions and computations is proposed in Constable and Smith (1987, 1988) by introducing a type-constructor $\bar{A}$, whose intuitive meaning is the set of *computations* of type A. Our categorical semantics is based on a similar idea. Constable and Smith, however, do not adequately capture the general axioms for computations (as we do), since their notion of model, based on an untyped partial applicative structure, accounts only for partial computations.

1. A Categorical Semantics of Computations

The basic idea behind the categorical semantics below is that, in order to interpret a programming language in a category $\mathscr{C}$, we distinguish the object A of values (of type A) from the object TA of computations (of type

A), and take as denotations of programs (of type A) the *elements* of TA. In particular, we identify the type A with the object of values (of type A) and obtain the object of computations (of type A) by applying an unary type-constructor T to A. We call T a *notion of computation*, since it abstracts away from the type of values computations may produce. There are many choices for TA corresponding to different notions of computations.

EXAMPLE 1.1. We give few notions of computation in the category of sets:

- **partiality** $TA = A_\perp$ (i.e., $A + \{\perp\}$), where $\perp$ is the *diverging computation*
- **nondeterminism** $TA = \mathscr{P}_{\text{fin}}(A)$
- **side-effects** $TA = (A \times S)^S$, where S is a set of states, e.g. a set U^L of stores or a set of input/output sequences U^*
- **exceptions** $TA = (A + E)$, where E is the set of exceptions
- **continuations** $TA = R^{(R^A)}$, where R is the set of results
- **interactive input** $TA = (\mu\gamma . A + \gamma^U)$, where U is the set of characters; more explicitly TA is the set of U-branching trees with finite branches and A-labelled leaves
- **interactive output** $TA = (\mu\gamma . A + (U \times \gamma))$; more explicitly TA is (isomorphic to) $U^* \times A$.

Further examples (in a category of cpos) could be given based on the denotational semantics for various programming languages (see Schmidt (1986), Gunter and Scott (1989), and Mosses (1989)).

Rather than focusing on a specific T, we want to find the general properties common to all notions of computation; therefore we impose as the only requirement that *programs* should form a category. The aim of this section is to convince the reader, with a sequence of informal argumentations, that such a requirement amounts to saying that T is part of a Kleisli triple $(T, \eta, \_^*)$ and that the category of programs is the Kleisli category for such a triple.

DEFINITION 1.2 (Manes, 1976). A *Kleisli triple* over a category $\mathscr{C}$ is a triple $(T, \eta, \_^*)$, where $T\colon \mathrm{Obj}(\mathscr{C}) \to \mathrm{Obj}(\mathscr{C})$, $\eta_A\colon A \to TA$ for $A \in \mathrm{Obj}(\mathscr{C})$, $f^*\colon TA \to TB$ for $f\colon A \to TB$ and the following equations hold:

- $\eta_A^* = \mathrm{id}_{TA}$
- $\eta_A ; f^* = f$ for $f\colon A \to TB$
- $f^* ; g^* = (f ; g^*)^*$ for $f\colon A \to TB$ and $g\colon B \to TC$.

A Kleisli satisfies the *mono requirement* provided η_A is mono for $A \in \mathscr{C}$.

Intuitively η_A is the *inclusion* of values into computations (in several cases η_A is indeed a mono) and f^* is the *extension* of a function f from values to computations to a function from computations to computations, which first evaluates a computation and then applies f to the resulting value. In summary

$$a\colon A \xmapsto{\eta_A} [a]\colon TA$$

$$\frac{a\colon A \xmapsto{f} f(a)\colon TB}{c\colon TA \xmapsto{f^*} (\text{let } x \Leftarrow c \text{ in } f(x))\colon TB}.$$

In order to justify the axioms for a Kleisli triple we have first to introduce a category $\mathscr{C}_T$ whose morphisms correspond to programs. We proceed by analogy with the categorical semantics for terms, where types are interpreted by objects and terms of type B with a parameter (free variable) of type A are interpreted by morphisms from A to B. Since the denotation of programs of type B are supposed to be elements of TB, programs of type B with a parameter of type A ought to be interpreted by morphisms with codomain TB, but for their domain there are two alternatives, either A or TA, depending on whether parameters of type A are identified with values or computations of type A. We choose the first alternative, because it entails the second. Indeed computations of type A are the same as values of type TA. So we take $\mathscr{C}_T(A, B)$ to be $\mathscr{C}(A, TB)$. It remains to define composition and identities in $\mathscr{C}_T$ (and show that they satisfy the unit and associativity axioms for categories).

DEFINITION 1.3. Given a Kleisli triple $(T, \eta, -^*)$ over $\mathscr{C}$, the *Kleisli category* $\mathscr{C}_T$ is defined as follows:

- the objects of $\mathscr{C}_T$ are those of $\mathscr{C}$
- the set $\mathscr{C}_T(A, B)$ of morphisms from A to B in $\mathscr{C}_T$ is $\mathscr{C}(A, TB)$
- the identity on A in $\mathscr{C}_T$ is $\eta_A\colon A \to TA$
- $f \in \mathscr{C}_T(A, B)$ followed by $g \in \mathscr{C}_T(B, C)$ in $\mathscr{C}_T$ is $f; g^*\colon A \to TC$.

It is natural to take η_A as the identity on A in the category $\mathscr{C}_T$, since it maps a parameter x to $[x]$, i.e., to x viewed as a computation. Similarly composition in $\mathscr{C}_T$ has a simple explanation in terms of the intuitive meaning of f^*, in fact

$$\frac{x\colon A \xmapsto{f} f(x)\colon TB \quad y\colon B \xmapsto{g} g(y)\colon TC}{x\colon A \xmapsto{f;g^*} (\text{let } y \Leftarrow f(x) \text{ in } g(y))\colon TC},$$

i.e., f followed by g in $\mathcal{C}_T$ with parameter x is the program which first evaluates the program $f(x)$ and then feeds the resulting value as parameter to g. At this point we can give also a simple justification for the three axioms of Kleisli triples, namely that they are equivalent to the unit and associativity axioms for $\mathcal{C}_T$:

- $f; \eta_B^* = f$ for $f\colon A \to TB$
- $\eta_A; f^* = f$ for $f\colon A \to TB$
- $(f; g^*); h^* = f; (g; h^*)^*$ for $f\colon A \to TB$, $g\colon B \to TC$ and $h\colon C \to TD$.

EXAMPLE 1.4. We go through the notions of computation given in Example 1.1 and show that they are indeed part of suitable Kleisli triples.

- **partiality** $TA = A_\bot (= A + \{\bot\})$

η_A is the inclusion of A into $A_\bot$
if $f\colon A \to TB$, then $f^*(\bot) = \bot$ and $f^*(a) = f(a)$ (when $a \in A$)

- **nondeterminism** $TA = \mathcal{P}_{\mathrm{fin}}(A)$

η_A is the singleton map $a \mapsto \{a\}$
if $f\colon A \to TB$ and $c \in TA$, then $f^*(c) = \bigcup_{x \in c} f(x)$

- **side-effects** $TA = (A \times S)^S$

η_A is the map $a \mapsto (\lambda s\colon S.\langle a, s\rangle)$
if $f\colon A \to TB$ and $c \in TA$, then $f^*(c) = \lambda s\colon S.(\mathrm{let}\ \langle a, s'\rangle = c(s)\ \mathrm{in}\ f(a)(s'))$

exceptions $TA = (A + E)$

η_A is the injection map $a \mapsto \mathrm{inl}(a)$
if $f\colon A \to TB$ then $f^*(\mathrm{inr}(e)) = e$ (when $e \in E$) and $f^*(\mathrm{inl}(a)) = f(a)$ (when $a \in A$)

- **continuations** $TA = R^{(R^A)}$

η_A is the map $a \mapsto (\lambda k)\colon R^A.k(a))$
if $f\colon A \to TB$ and $c \in TA$, then $f^*(c) = (\lambda k\colon R^B.c(\lambda a\colon A.f(a)(k)))$

- **interactive input** $TA = (\mu\gamma.A + \gamma^U)$

η_A maps a to the tree consisting only of one leaf labelled with a
if $f\colon A \to TB$ and $c \in TA$, then $f^*(c)$ is the tree obtained by replacing leaves of c labelled by a with the tree $f(a)$

- **interactive output** $TA = (\mu\gamma.A + (U \times \gamma))$

η_A is the map $a \mapsto \langle \varepsilon, a\rangle$
if $f\colon A \to TB$, then $f^*(\langle s, a\rangle) = \langle s * s', b\rangle$, where $f(a) = \langle s', b\rangle$ and $s * s'$ is the concatenation of s followed by s'.

Kleisli triples are just an alternative description for monads. Although the former are easy to justify from a computational perspective, the latter are more widely used in the literature on category theory and have the advantage of being defined only in terms of functors and natural transformations, which make them more suitable for abstract manipulation.

DEFINITION 1.5 (MacLane, 1971). A *monad* over a category $\mathscr{C}$ is a triple (T, η, μ), where $T: \mathscr{C} \to \mathscr{C}$ is a functor, $\eta: \mathrm{Id}_{\mathscr{C}} \to T$, and $\mu: T^2 \to T$ are natural transformations and the following diagrams commute:

$$\begin{tikzcd} T^3A \arrow[r, "\mu_{TA}"] \arrow[d, "T\mu_A"'] & T^2A \arrow[d, "\mu_A"] \\ T^2A \arrow[r, "\mu_A"'] & TA \end{tikzcd} \qquad \begin{tikzcd} TA \arrow[r, "\eta_{TA}"] \arrow[dr, "\mathrm{id}_{TA}"'] & T^2A \arrow[d, "\mu_A"] & TA \arrow[l, "T\eta_A"'] \arrow[dl, "\mathrm{id}_{TA}"] \\ & TA & \end{tikzcd}$$

PROPOSITION 1.6 (Manes, 1976). *There is a one-one correspondence between Kleisli triples and monads.*

Proof. Given a Kleisli triple $(T, \eta, \_^*)$, the corresponding monad is (T, η, μ), where T is the extension of the function T to an endofunctor by taking $T(f) = (f; \eta_B)^*$ for $f: A \to B$ and $\mu_A = \mathrm{id}^*_{TA}$. Conversely, given a monad (T, η, μ), the corresponding Kleisli triple is $(T, \eta, \_^*)$, where T is the restriction of the functor T to objects and $f^* = (Tf); \mu_B$ for $f: A \to TB$. ∎

Remark 1.7. In general the categorical semantics of partial maps, based on a category $\mathscr{C}$ equipped with a *dominion* $\mathscr{M}$ (see Rosolini, 1986) cannot be reformulated in terms of a Kleisli triple over $\mathscr{C}$ satisfying some additional properties unless $\mathscr{C}$ has *lifting*; i.e., the inclusion functor from $\mathscr{C}$ into the category of partial maps $P(\mathscr{C}, \mathscr{M})$ has a right adjoint $\_{}_\perp$ characterised by the natural isomorphism

$$\mathscr{C}(A, B_\perp) \doteq\cong P(\mathscr{C}, \mathscr{M})(A, B).$$

This mismatch disappears when partial cartesian closed categories are considered.

2. SIMPLE LANGUAGES FOR MONADS

In this section we consider two formal systems motivated by different objectives: reasoning about programming languages and reasoning about programs in a fixed programming language. When reasoning about programming languages one has different monads (for simplicity we assume that they are over the same category), one for each programming language, and the main aim is to study how they relate to each other. So it is natural to base a formal system on a *metalanguage* for a category and treat monads as unary type-constructors. When reasoning about programs one has only one monad, because the programming language is fixed, and the main aim is to prove properties of programs. In this case the obvious choice for the term language is the *programming language* itself, which is more naturally interpreted in the Kleisli category.

Remark 2.1. We regard the metalanguage as more fundamental. In fact, its models are more general, as they do not have to satisfy the mono requirement, and the interpretation of programs (of some given programming language) can be defined simply by translation into (a suitable extension of) the metalanguage. It should be pointed out that the mono requirement cannot be axiomatised in the metalanguage, as we should need conditional equations $[x]_T = [y]_T \rightarrow x = y$, and that existence assertions cannot be translated into formulas of the metalanguage, as we would need existentially quantified formulas $(e \downarrow_\sigma)^\circ \equiv (\exists!\, x\colon \sigma . e^\circ = [x]_T)$.[2]

In Section 2.3 we will explain once for all the correspondence between theories of a simple programming language and categories with a monad satisfying the mono requirement. For other programming languages we will give only their translation in a suitable extension of the metalanguage. In this way, issues such as call-by-value versus call-by-name affect the translation, but not the metalanguage.

In categorical logic it is common practice to identify a *theory* $\mathcal{T}$ with a category $\mathcal{F}(\mathcal{T})$ with additional structure such that there is a one-one correspondence between *models* of $\mathcal{T}$ in a category $\mathcal{C}$ with additional structure and structure preserving functors from $\mathcal{F}(\mathcal{T})$ to $\mathcal{C}$ (see Kock and Reyes, 1977).[3] This identification was originally proposed by Lawvere, who also showed that algebraic theories can be viewed as categories with finite products.

In Section 2.2 we give a class of theories that can be viewed as categories with monads, so that any category with a monad is, up to *equivalence* (of categories with a monad), one of such theories. Such a reformulation in terms of theories is more suitable for formal manipulation and more appealing to those unfamiliar with category theory. However, there are other advantages in having an alternative presentation of monads. For instance, natural extensions of the syntax may suggest extensions of the categorical structure that may not be immediate to motivate and justify otherwise (we will exploit this in Section 3). In Section 2.3 we take a programming language perspective and establish a correspondence between theories (with equivalence and existence assertions) for a simple program-

[2] The uniqueness of x s.t. $e^0 = [x]_T$ follows from the mono requirement.

[3] In Lambek and Scott (1986) a stronger relation is sought between theories and categories with additional structure, namely an equivalence between the category of theories and translations and the category of small categories with additional structure and structure preserving functors. In the case of typed λ-calculus, for instance, such an equivalence between λ-theories and cartesian closed categories requires a modification in the definition of λ-theory, which allows not only equations between λ-terms but also equations between type expressions.

ming language and categories with a monad satisfying the mono requirement, i.e., η_A mono for every A.

As starting point we take *many sorted monadic equational logic*, because it is more primitive than many sorted equational logic; indeed, monadic theories are equivalent to categories without any additional structure.

2.1. *Many Sorted Monadic Equational Logic*

The language and formal system of many sorted monadic equational logic are parametric in a signature, i.e., a set of base types A and unary function symbols $f\colon A_1 \to A_2$. The language is made of types $\vdash A$ type, terms $x\colon A_1 \vdash e\colon A_2$, and equations $x\colon A_1 \vdash e_1 =_{A_2} e_2$ defined by the following formation rules:

$$A\ \frac{}{\vdash A \text{ type}}\ A \text{ base type}$$

$$\text{var}\ \frac{\vdash A \text{ type}}{x\colon A \vdash x\colon A}$$

$$f\ \frac{x\colon A \vdash e_1\colon A_1}{x\colon A \vdash f(e_1)\colon A_2}\ f\colon A_1 \to A_2$$

$$\text{eq}\ \frac{x\colon A_1 \vdash e_1\colon A_2 \quad x\colon A_1 \vdash e_2\colon A_2}{x\colon A_1 \vdash e_1 =_{A_2} e_2}.$$

TABLE 1

Interpretation of Many Sorted Monadic Equational Language

Rule	Syntax		Semantics
A			
	$\overline{\vdash A \text{ type}}$	$=$	$[\![A]\!]$
var			
	$\vdash A$ type	$=$	c
	$\overline{x\colon A \vdash x\colon A}$	$=$	id_c
$f\colon A_1 \to A_2$			
	$x\colon A \vdash e_1\colon A_1$	$=$	g
	$\overline{x\colon A \vdash f(e_1)\colon A_2}$	$=$	$g; [\![f]\!]$
eq			
	$x\colon A_1 \vdash e_1\colon A_2$	$=$	g_1
	$x\colon A_1 \vdash e_2\colon A_2$	$=$	g_2
	$\overline{x\colon A_1 \vdash e_1 =_{A_2} e_2}$	$\Leftrightarrow$	$g_1 = g_2$

Remark 2.2. Terms of (many sorted) monadic equational logic have exactly one free variable (the one declared in the context) which occurs exactly once, and equations are between terms with the same free variable.

An interpretation $[\![-]\!]$ of the language in a category $\mathscr{C}$ is parametric in an interpretation of the symbols in the signature and is defined by induction on the derivation of well-formedness for (types) terms and equations (see Table 1) according to the following general pattern:

- the interpretation $[\![A]\!]$ of a base type A is an object of $\mathscr{C}$
- the interpretation $[\![f]\!]$ of a unary function $f\colon A_1 \to A_2$ is a morphism from $[\![A_1]\!]$ to $[\![A_2]\!]$ in $\mathscr{C}$; similarly for the interpretation of a term $x\colon A_1 \vdash e\colon A_2$
- the interpretation of an assertion $x\colon A \vdash \phi$ (in this case just an equation) is either true or false.

Remark 2.3. The interpretation of equations is standard. However, if one want to consider more complex assertions, e.g. formulas of first order logic, then they should be interpreted by subobjects; in particular, equality $- = -: A$ should be interpreted by the diagonal $\Delta_{[\![A]\!]}$.

The formal consequence relation on the set of equations is generated by the inference rules for equivalences ((refl), (simm), and (trans)), congruence, and substitutivity (see Table 2). This formal consequence relation is *sound and complete* w.r.t interpretation of the language in categories; i.e., an equation is formally derivable from a set of equational axioms if and only if all the interpretations satisfying the axioms satisfy the equation.

TABLE 2

Inference Rules of Many Sorted Monadic Equational Logic

refl	$\dfrac{x\colon A \vdash e\colon A_1}{x\colon A \vdash e =_{A_1} e}$
symm	$\dfrac{x\colon A \vdash e_1 =_{A_1} e_2}{x\colon A \vdash e_2 =_{A_1} e_1}$
trans	$\dfrac{x\colon A \vdash e_1 =_{A_1} e_2 \quad x\colon A \vdash e_2 =_{A_1} e_3}{x\colon A \vdash e_2 =_{A_1} e_3}$
congr	$\dfrac{x\colon A \vdash e_1 =_{A_1} e_2}{x\colon A \vdash f(e_1) =_{A_2} f(e_2)}\, f\colon A_1 \to A_2$
subst	$\dfrac{x\colon A \vdash e\colon A_1 \quad x\colon A_1 \vdash \phi}{x\colon A \vdash [e/x]\phi}$

Soundness follows from the admissibility of the inference rules in any interpretation, while completeness follows from the fact that any theory $\mathcal{T}$ (i.e., a set of equations closed w.r.t. the inference rules) is the set of equations satisfied by *the canonical interpretation* in the category $\mathcal{F}(\mathcal{T})$, i.e., $\mathcal{T}$ viewed as a category.

DEFINITION. 2.4. Given a monadic equational theory $\mathcal{T}$, the category $\mathcal{F}(\mathcal{T})$ is defined as follows:

- objects are (base) types A,
- morphisms from A_1 to A_2 are equivalence classes $[x: A_1 \vdash e: A_2]_{\mathcal{T}}$ of terms w.r.t. the equivalence relation induced by the theory $\mathcal{T}$; i.e.,

$$(x: A_1 \vdash e_1 : A_2) \equiv (x: A_1 \vdash e_2 : A_2) \Leftrightarrow (x: A_1 \vdash e_1 =_{A_2} e_2) \in \mathcal{T}$$

- composition is substitution; i.e.,

$$[x: A_1 \vdash e_1 : A_2]_{\mathcal{T}}; [x: A_2 \vdash e_2 : A_3]_{\mathcal{T}} = [x: A_1 \vdash [e_1/x]e_2 : A_3]_{\mathcal{T}}$$

- identity over A is $[x: A \vdash x: A]_{\mathcal{T}}$.

There is also a correspondence in the opposite direction, namely every category $\mathcal{C}$ (with additional structure) can be viewed as a theory $\mathcal{T}_{\mathcal{C}}$ (i.e., the *theory* of $\mathcal{C}$ over the language for $\mathcal{C}$), so that $\mathcal{C}$ and $\mathcal{F}(\mathcal{T}_{\mathcal{C}})$ are equivalent as categories (with additional structure). Actually, in the case of monadic equational theories and categories, $\mathcal{C}$ and $\mathcal{F}(\mathcal{T}_{\mathcal{C}})$ are isomorphic.

In the sequel we consider other equational theories. They can be viewed as categories in the same way described above for monadic theories; moreover, these categories are equipped with additional structure, depending on the specific nature of the theories under consideration.

2.2. *The Simple Metalanguage*

We extend many sorted monadic equational logic to match categories equipped with a monad (or equivalently a Kleisli triple). Although we consider only one monad, it is conceptually straightforward to have several monads at once.

The first step is to extend the language. This could be done in several ways without affecting the correspondence between theories and monads. We choose a presentation inspired by Kleisli triples; more specifically we introduce a unary type-constructor T and the two term-constructors $[-]$ and let, used informally in Section 1. The definition of signature is slightly modified, since the domain and codomain of a unary function symbol f: $\tau_1 \to \tau_2$ can be any type, not just base types (the fact is that in many sorted monadic logic the only types are base types). An interpretation $[\![-]\!]$ of the

TABLE 3

Interpretation of the Simple Metalanguage

Rule	Syntax		Semantics
A			
	$\dfrac{}{\vdash_{\mathrm{ml}} A \text{ type}}$	$=$	$[\![A]\!]$
T			
	$\dfrac{\vdash_{\mathrm{ml}} \tau \text{ type}}{\vdash_{\mathrm{ml}} T\tau \text{ type}}$	$=$ $=$	c Tc
var			
	$\dfrac{\vdash_{\mathrm{ml}} \tau \text{ type}}{x\colon \tau \vdash_{\mathrm{ml}} x\colon \tau}$	$=$ $=$	c id_c
$f\colon \tau_1 \to \tau_2$			
	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1\colon \tau_1}{x\colon \tau \vdash_{\mathrm{ml}} f(e_1)\colon \tau_2}$	$=$ $=$	g $g; [\![f]\!]$
$[-]_T$			
	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e\colon \tau'}{x\colon \tau \vdash_{\mathrm{ml}} [e]_T\colon T\tau'}$	$=$ $=$	g $g; \eta_{[\![\tau']\!]}$
let			
	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1\colon T\tau_1 \quad x_1\colon \tau_1 \vdash_{\mathrm{ml}} e_2\colon T\tau_2}{x\colon \tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } e_2)\colon T\tau_2}$	$=$ $=$ $=$	g_1 g_2 $g_1; g_2^*$
eq			
	$\dfrac{x\colon \tau_1 \vdash_{\mathrm{ml}} e_1\colon \tau_2 \quad x\colon \tau_1 \vdash_{\mathrm{ml}} e_2\colon \tau_2}{x\colon \tau_1 \vdash_{\mathrm{ml}} e_1 =_{\tau_2} e_2}$	$=$ $=$ $\Leftrightarrow$	g_1 g_2 $g_1 = g_2$

language in a category $\mathscr{C}$ with a Kleisli triple $(T, \eta, -^*)$ is parametric in an interpretation of the symbols in the signature and is defined by induction on the derivation of well-formedness for types, terms, and equations (see Table 3). Finally we add to many sorted monadic equational logic appropriate inference rules capturing axiomatically the properties of the new type- and term-constructors after interpretation (see Table 4).

PROPOSITION 2.5. *Every theory $\mathscr{T}$ of the simple metalanguage, viewed as a category $\mathscr{F}(\mathscr{T})$, is equipped with a Kleisli triple $(T, \eta, -^*)$:*

- $T(\tau) = T\tau$,
- $\eta_\tau = [x\colon \tau \vdash_{\mathrm{ml}} [x]_T\colon T\tau]_{\mathscr{T}}$,
- $([x\colon \tau_1 \vdash_{\mathrm{ml}} e\colon T\tau_2]_{\mathscr{T}})^* = [x'\colon T\tau_1 \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x \Leftarrow x' \textit{ in } e)\colon T\tau_2]_{\mathscr{T}}$.

TABLE 4

Inference Rules of the Simple Metalanguage

$[-].\xi$	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1 =_{\tau_1} e_2}{x\colon \tau \vdash_{\mathrm{ml}} [e_1]_T =_{T\tau_1} [e_2]_T}$
let.ξ	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1 =_{T\tau_1} e_2 \quad x'\colon \tau_1 \vdash_{\mathrm{ml}} e'_1 =_{T\tau_2} e'_2}{x\colon \tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x' \Leftarrow e_1 \text{ in } e'_1) =_{T\tau_2} (\mathrm{let}_T\, x' \Leftarrow e_2 \text{ in } e'_2)}$
ass	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1\colon T\tau_1 \quad x_1\colon \tau_1 \vdash_{\mathrm{ml}} e_2\colon T\tau_2 \quad x_2\colon \tau_2 \vdash_{\mathrm{ml}} e_3\colon T\tau_3}{x\colon \tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_2 \Leftarrow (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } e_2) \text{ in } e_3) =_{T\tau_3} (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } (\mathrm{let}_T\, x_2 \Leftarrow e_2 \text{ in } e_3))}$
$T.\beta$	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1\colon \tau_1 \quad x_1\colon \tau_1 \vdash_{\mathrm{ml}} e_2\colon T\tau_2}{x\colon \tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_1 \Leftarrow [e_1]_T \text{ in } e_2) =_{T\tau_2} [e_1/x_1]e_2}$
$T.\eta$	$\dfrac{x\colon \tau \vdash_{\mathrm{ml}} e_1\colon T\tau_1}{x\colon \tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } [x_1]_T) =_{T\tau_1} e_1}$

Proof. We have to show that the three axioms for Kleisli triples are valid. The validity of each axiom amounts to the derivability of an equation. For instance, $\eta^*_\tau = \mathrm{id}_{T\tau}$ is valid provided $x'\colon T\tau \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x \Leftarrow x' \text{ in } [x]_T) =_{T\tau} x'$ is derivable, indeed it follows from $(T.\eta)$. The reader can check that the equations corresponding to the axioms $\eta_\tau; f^* = f$ and $f^*; g^* = (f; g^*)$ follow from $(T.\beta)$ and (ass) respectively. ∎

2.3. *A Simple Programming Language*

In this section we take a programming language perspective by introducing a simple programming language, whose terms are interpreted by morphisms of the Kleisli category for a monad. Unlike the metalanguage of Section 2.2, the programming language does not allow to consider more than one monad at once.

The interpretation in the Kleisli category can also be given indirectly via a translation in the simple metalanguage of Section 2.2 mapping *programs* of type τ into *terms* of type $T\tau$. If we try to establish a correspondence between *equational theories* of the simple programming language and categories with one monad (as done for the metalanguage), then we run into problems, since there is no way (in general) to recover $\mathscr{C}$ from $\mathscr{C}_T$. What we do instead is to establish a correspondence between theories with *equivalence* and *existence* assertions and categories with one monad satisfying the *mono requirement*; i.e., η_A is mono for every object A (note that η_{TA} is always a mono, because $\eta_{TA}; \mu_A = \mathrm{id}_{TA}$). The intended extension of the existence predicate on computations of type A is the set of computations of the form $[v]$ for some value of type A, so it is natural to require η_A to be mono and interpret the existence predicate as the *subobject* corresponding to η_A.

The simple programming language is parametric in a signature, i.e., a set of base types and unary command symbols. To stress that the interpretation is in $\mathscr{C}_T$ rather than $\mathscr{C}$, we use unary command symbols $p\colon \tau_1 \rightharpoonup \tau_2$ (instead of unary function symbols $f\colon \tau_1 \to \tau_2$), we call $x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2$ a program (instead of a term) and we write $- \equiv_\tau -$ (instead of $- =_{T\tau} -$)

TABLE 5

Interpretation of the Simple Programming Language

Rule	Syntax		Semantics
A			
	$\vdash_{\mathrm{pl}} A$ type	$=$	$[\![A]\!]$
T	$\vdash_{\mathrm{pl}} \tau$ type	$=$	c
	$\vdash_{\mathrm{pl}} T\tau$ type	$=$	Tc
var	$\vdash_{\mathrm{pl}} \tau$ type	$=$	c
	$x\colon \tau \vdash_{\mathrm{pl}} x\colon \tau$	$=$	η_c
$p\colon \tau_1 \rightharpoonup \tau_2$	$x\colon \tau \vdash_{\mathrm{pl}} e_1\colon \tau_1$	$=$	g
	$x\colon \tau \vdash_{\mathrm{pl}} p(e_1)\colon \tau_2$	$=$	$g; [\![p]\!]^*$
$[-]$	$x\colon \tau \vdash_{\mathrm{pl}} e\colon \tau'$	$=$	g
	$x\colon \tau \vdash_{\mathrm{pl}} [e]\colon T\tau'$	$=$	$g; \eta_{T[\![\tau']\!]}$
μ	$x\colon \tau \vdash_{\mathrm{pl}} e\colon T\tau'$	$=$	g
	$x\colon \tau \vdash_{\mathrm{pl}} \mu(e)\colon \tau'$	$=$	$g; \mu_{[\![\tau']\!]}$
let	$x\colon \tau \vdash_{\mathrm{pl}} e_1\colon \tau_1$	$=$	g_1
	$x_1\colon \tau_1 \vdash_{\mathrm{pl}} e_2\colon \tau_2$	$=$	g_2
	$x\colon \tau \vdash_{\mathrm{pl}} (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ e_2)\colon \tau_2$	$=$	$g_1; g_2^*$
eq	$x\colon \tau_1 \vdash_{\mathrm{pl}} e_1\colon \tau_2$	$=$	g_1
	$x\colon \tau_1 \vdash_{\mathrm{pl}} e_2\colon \tau_2$	$=$	g_2
	$x\colon \tau_1 \vdash_{\mathrm{pl}} e_1 \equiv_{\tau_2} e_2$	$\Leftrightarrow$	$g_1 = g_2$
ex	$x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2$	$=$	g
	$x\colon \tau_1 \vdash_{\mathrm{pl}} e\downarrow_{\tau_2}$	$\Leftrightarrow$	$\exists!\, h\colon [\![\tau_1]\!] \to [\![\tau_2]\!]$ s.t. $g = h; \eta_{[\![\tau_2]\!]}$

as equality of computations of type τ. Given a category $\mathscr{C}$ with a Kleisli triple $(T, \eta, -^*)$ satisfying the mono requirement, an interpretation $[\![-]\!]$ of the programming language is parametric in an interpretation of the symbols in the signature and is defined by induction on the derivation of well-formedness for types, terms, and equations (see Table 5) following the same pattern given for many sorted monadic equational logic, but with $\mathscr{C}$ replaced by $\mathscr{C}_T$, namely:

- the interpretation $[\![\tau]\!]$ of a (base) type τ is an object of $\mathscr{C}_T$, or equivalently an object of $\mathscr{C}$
- the interpretation $[\![p]\!]$ of a unary command $p\colon \tau_1 \rightharpoonup \tau_2$ is a morphism from $[\![\tau_1]\!]$ to $[\![\tau_2]\!]$ in $\mathscr{C}_T$, or equivalently a morphism from $[\![\tau_1]\!]$ to $T[\![\tau_2]\!]$ in $\mathscr{C}$; similarly for the interpretation of a program $x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2$.
- the interpretation of an equivalence or existence assertion is a truth value.

Remark 2.6. The let-constructor play a fundamental role: operationally it corresponds to sequential evaluation of programs and categoritcally it corresponds to composition in the Kleisli category $\mathscr{C}_T$ (while substitution corresponds to composition in $\mathscr{C}$). In the λ_v-calculus (let $x \Leftarrow e$ in e') is treated as syntactic sugar for $(\lambda x.e')e$. We think that this is not the right way to proceed, because it explains the let-constructor (i.e., sequential evaluation of programs) in terms of constructors available only in functional languages. On the other hand, (let $x \Leftarrow e$ in e') cannot be treated as syntactic sugar for $[e/x]e'$ (involving only the more primitive substitution) without collapsing computations to values.

The existence predicate $e\downarrow$ is inspired by the logic of partial terms/elements (see Fourman, 1977; Scott, 1979; Moggi, 1988); however, there are important differences, e.g.

$$\text{strict}\ \frac{x\colon \tau \vdash_{\mathrm{pl}} p(e)\downarrow_{\tau_2}}{x\colon \tau \vdash_{\mathrm{pl}} e\downarrow_{\tau_1}}\ p\colon \tau_1 \rightharpoonup \tau_2$$

is admissible for partial computations, but not in general. For certain notions of computation there may be other predicates on computations worth considering, or the existence predicate itself may have a more specialised meaning, for instance:

- a partial computation exists iff it terminates;
- a non-deterministic computation exists iff it gives exactly one result;
- a computation with side-effects exists iff it does not change the store.

Programs can be translated into terms of the metalanguage via a translation $-^\circ$ s.t. for every well-formed program $x: \tau_1 \vdash_{pl} e: \tau_2$ the term $x: \tau_1 \vdash_{ml} e^\circ: T\tau_2$ is well-formed and $[\![x: \tau_1 \vdash_{pl} e: \tau_2]\!] = [\![x: \tau_1 \vdash_{ml} e^\circ: T\tau_2]\!]$ (the proof of these properties is left to the reader).

DEFINITION 2.7. Given a signature Σ for the programming language, let Σ° be the signature for the metalanguage with the same base types and a function $p: \tau_1 \to T\tau_2$ for each command $p: \tau_1 \rightharpoonup \tau_2$ in Σ. The translation $-^\circ$ from programs over Σ to terms over Σ° is defined by induction on raw programs:

- $x^\circ \overset{\Delta}{\equiv} [x]_T$
- $(\text{let } x_1 \Leftarrow e_1 \text{ in } e_2)^\circ \overset{\Delta}{\equiv} (\text{let}_T\, x_1 \Leftarrow e_1^\circ \text{ in } e_2^\circ)$
- $p(e_1)^\circ \overset{\Delta}{\equiv} (\text{let}_T\, x \Leftarrow e_1^\circ \text{ in } p(x))$
- $[e]^\circ \overset{\Delta}{\equiv} [e^\circ]_T$
- $\mu(e)^\circ \overset{\Delta}{\equiv} (\text{let}_T\, x \Leftarrow e^\circ \text{ in } x)$.

The inference rules for deriving equivalence and existence assertions of the simple programming language can be partitioned as follows:

TABLE 6

General Inference Rules

refl	$\dfrac{x: \tau \vdash_{pl} e: \tau_1}{x: \tau \vdash_{pl} e \equiv_{\tau_1} e}$
symm	$\dfrac{x: \tau \vdash_{pl} e_1 \equiv_{\tau_1} e_2}{x: \tau \vdash_{pl} e_2 \equiv_{\tau_1} e_1}$
trans	$\dfrac{x: \tau \vdash_{pl} e_1 \equiv_{\tau_1} e_2 \quad x: \tau \vdash_{pl} e_2 \equiv_{\tau_1} e_3}{x: \tau \vdash_{pl} e_2 \equiv_{\tau_1} e_3}$
congr	$\dfrac{x: \tau \vdash_{pl} e_1 \equiv_{\tau_1} e_2}{x: \tau \vdash_{pl} p(e_1) \equiv_{\tau_2} p(e_2)}\ p: \tau_1 \rightharpoonup \tau_2$
$E.x$	$\dfrac{\vdash_{pl} \tau \text{ type}}{x: \tau \vdash_{pl} x \downarrow_\tau}$
$E.$congr	$\dfrac{x: \tau \vdash_{pl} e_1 \equiv_{\tau_1} e_2 \quad x: \tau \vdash_{pl} e_1 \downarrow_{\tau_1}}{x: \tau \vdash_{pl} e_2 \downarrow_{\tau_1}}$
subst	$\dfrac{x: \tau \vdash_{pl} e \downarrow_{\tau_1} \quad x: \tau_1 \vdash_{pl} \phi}{x: \tau \vdash_{pl} [e/x]\phi}$

- general rules (see Table 6) for terms denoting computations, but with variables ranging over values; these rules replace those of Table 2 for many sorted monadic equational logic
- rules capturing the properties of type- and term-constructors (see Table 7) after interpretation of the programming language; these rules replace the additional rules for the metalanguage given in Table 4.

Soundness and completeness of the formal consequence relation w.r.t. interpretation of the simple programming language in categories with a monad satisfying the mono requirement is established in the usual way (see Section 2.1). The only step which differs is how to view a theory $\mathcal{T}$ of the simple programming language (i.e., a set of equivalence and existence

TABLE 7

Inference Rules of the Simple Programming Language

$[-].\xi$	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 \equiv_{\tau_1} e_2}{x:\tau \vdash_{\mathrm{pl}} [e_1] \equiv_{T\tau_1} [e_2]}$
$E.[-]$	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 : \tau_1}{x:\tau \vdash_{\mathrm{pl}} [e_1] \downarrow_{T\tau_1}}$
$\mu.\xi$	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 \equiv_{T\tau_1} e_2}{x:\tau \vdash_{\mathrm{pl}} \mu(e_1) \equiv_{\tau_1} \mu(e_2)}$
$\mu.\beta$	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 : \tau_1}{x:\tau \vdash_{\mathrm{pl}} \mu([e_1]) \equiv_{\tau_1} e_1}$
$\mu.\eta$	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 \downarrow_{T\tau_1}}{x:\tau \vdash [\mu(e_1)] \equiv_{T\tau_1} e_1}$
let.ξ	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 \equiv_{\tau_1} e_2 \quad x':\tau_1 \vdash_{\mathrm{pl}} e'_1 \equiv_{\tau_2} e'_2}{x:\tau \vdash_{\mathrm{pl}} (\mathrm{let}\ x' \Leftarrow e_1\ \mathrm{in}\ e'_1) \equiv_{\tau_2} (\mathrm{let}\ x' \Leftarrow e_2\ \mathrm{in}\ e'_2)}$
unit	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 : \tau_1}{x:\tau \vdash_{\mathrm{pl}} (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ x_1) \equiv_{\tau_1} e_1}$
ass	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 : \tau_1 \quad x_1:\tau_1 \vdash_{\mathrm{pl}} e_2 : \tau_2 \quad x_2:\tau_2 \vdash_{\mathrm{pl}} e_3 : \tau_3}{x:\tau \vdash_{\mathrm{pl}} (\mathrm{let}\ x_2 \Leftarrow (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ e_2)\ \mathrm{in}\ e_3) \equiv_{\tau_3} (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ (\mathrm{let}\ x_2 \Leftarrow e_2\ \mathrm{in}\ e_3))}$
let.β	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 \downarrow_{\tau_1} \quad x_1:\tau_1 \vdash_{\mathrm{pl}} e_2 : \tau_2}{x:\tau \vdash_{\mathrm{pl}} (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ e_2) \equiv_{\tau_2} [e_1/x]e_2}$
let.p	$\dfrac{x:\tau \vdash_{\mathrm{pl}} e_1 : \tau_1}{x:\tau \vdash_{\mathrm{pl}} p(e_1) \equiv_{\tau_1} (\mathrm{let}\ x_1 \Leftarrow e_1\ \mathrm{in}\ p(x_1))}\ p:\tau_1 \rightharpoonup \tau_2$

assertions closed w.r.t. the inference rules) as a category $\mathcal{F}(\mathcal{T})$ with the required structure.

DEFINITION 2.8. Given a theory $\mathcal{T}$ of the simple programming language, the category $\mathcal{F}(\mathcal{T})$ is defined as follows:

- objects are types τ,
- morphisms from τ_1 to τ_2 are equivalence classes $[x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2]_{\mathcal{T}}$ of existing programs $x\colon \tau_1 \vdash_{\mathrm{pl}} e \downarrow_{\tau_2} \in \mathcal{T}$ w.r.t. the equivalence relation induced by the theory $\mathcal{T}$, i.e.,

$$(x\colon \tau_1 \vdash_{\mathrm{pl}} e_1 : \tau_2) \equiv (x\colon \tau_1 \vdash_{\mathrm{pl}} e_2 : \tau_2) \Leftrightarrow (x\colon \tau_1 \vdash_{\mathrm{pl}} e_1 \equiv_{\tau_2} e_2) \in \mathcal{T}$$

- composition is substitution, i.e.,

$$[x\colon \tau_1 \vdash_{\mathrm{pl}} e_1 : \tau_2]_{\mathcal{T}}; [x\colon \tau_2 \vdash_{\mathrm{pl}} e_2 : \tau_3]_{\mathcal{T}} = [x\colon \tau_1 \vdash_{\mathrm{pl}} [e_1/x]e_2 ; \tau_3]_{\mathcal{T}}$$

- identity over τ is $[x\colon \tau \vdash_{\mathrm{pl}} x\colon \tau]_{\mathcal{T}}$.

In order for composition in $\mathcal{F}(\mathcal{T})$ to be well-defined, it is essential to consider only equivalence classes of existing programs, since the simple programming language satisfies only a restricted form of substitutivity.

PROPOSITION 2.9. *Every theory $\mathcal{T}$ of the simple programming language, viewed as a category $\mathcal{F}(\mathcal{T})$, is equipped with a Kleisli triple $(T, \eta, -^*)$ satisfying the mono requirement*:

- $T(\tau) = T\tau$,
- $\eta_\tau = [x\colon \tau \vdash_{\mathrm{pl}} [x]\colon T\tau]_{\mathcal{T}}$,
- $([x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon T\tau_2]_{\mathcal{T}})^* = [x'\colon T\tau_1 \vdash_{\mathrm{pl}} [(\mathrm{let}\ x \Leftarrow \mu(x')\ \mathrm{in}\ \mu(e))]\colon T\tau_2]_{\mathcal{T}}$.

Proof. We have to show that the three axioms for Kleisli triples are valid. The validity of each axiom amounts to the derivability of an existence and equivalence assertion. For instance, $\eta_\tau^* = \mathrm{id}_{T\tau}$ is valid provided $x'\colon T\tau \vdash_{\mathrm{pl}} x' \downarrow_{T\tau}$ and $x'\colon T\tau \vdash_{\mathrm{pl}} [(\mathrm{let}\ x \Leftarrow \mu(x')\ \mathrm{in}\ \mu([x]))] \equiv_{T\tau} x'$ are derivable. The existence assertion follows immediately from $(E.x)$, while the equivalence is derived as follows:

- $x'\colon T\tau \vdash_{\mathrm{pl}} [(\mathrm{let}\ x \Leftarrow \mu(x')\ \mathrm{in}\ \mu([x]))] \equiv_{T\tau} [(\mathrm{let}\ x \Leftarrow \mu(x')\ \mathrm{in}\ x])$ by $(\mu.\beta)$, (refl), and $(\mathrm{let}.\xi)$

- $x'\colon T\tau \vdash_{\mathrm{pl}} [(\text{let } x \Leftarrow \mu(x') \text{ in } x)] \equiv_{T\tau} [\mu(x')]$ by (unit) and (let.ξ)
- $x'\colon T\tau \vdash_{\mathrm{pl}} [\mu(x')] \equiv_{T\tau} x'$ by $(E.x)$ and $(\mu.\eta)$
- $x'\colon T\tau \vdash_{\mathrm{pl}} [(\text{let } x \Leftarrow \mu(x') \text{ in } \mu([x]))] \equiv_{T\tau} x'$ by (trans).

We leave to the reader the derivation of the existence and equivalence assertions corresponding to the other axioms for Kleisli triples, and prove instead the mono requirement, i.e., that $f_1;\, \eta_\tau = f_2;\, \eta_\tau$ implies $f_1 = f_2$. Let f_i be $[x\colon \tau' \vdash_{\mathrm{pl}} e_i\colon \tau]_{\mathscr{T}}$; we have to derive $x\colon \tau' \vdash_{\mathrm{pl}} e_1 \equiv_\tau e_2$ from $x\colon \tau' \vdash_{\mathrm{pl}} [e_1] \equiv_{T\tau} [e_2]$ (and $x\colon \tau' \vdash_{\mathrm{pl}} e_i \downarrow_\tau$):

- $x\colon \tau' \vdash_{\mathrm{pl}} \mu([e_1]) \equiv_\tau \mu([e_2])$ by the first assumption and $(\mu.\xi)$
- $x\colon \tau' \vdash_{\mathrm{pl}} \mu([e_i]) \equiv_\tau e_i$ by $(\mu.\beta)$
- $x\colon \tau' \vdash_{\mathrm{pl}} e_1 \equiv_\tau e_2$ by (trans). ∎

Remark 2.10. One can show that the canonical interpretation of a program $x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2$ in the category $\mathscr{F}(\mathscr{T})$ is the morphism $[x\colon \tau_1 \vdash_{\mathrm{pl}} [e]\colon\; T\tau_2]_{\mathscr{T}}$. This interpretation establishes a one–one correspondence between morphisms from τ_1 to $T\tau_2$ in the category $\mathscr{F}(\mathscr{T})$, i.e., morphisms from τ_1 to τ_2 in the Kleisli category, and equivalence classes of programs $x\colon \tau_1 \vdash_{\mathrm{pl}} e\colon \tau_2$ (not necessarely existing). The inverse correspondence maps a morphism $[x\colon \tau_1 \vdash_{\mathrm{pl}} e'\colon T\tau_2]_{\mathscr{T}}$ to the equivalence class of $x\colon\; \tau_1 \vdash_{\mathrm{pl}} \mu(e')\colon\; \tau_2$. Indeed, $x\colon\; \tau_1 \vdash_{\mathrm{pl}} e \equiv_{\tau_2} \mu([e])$ and $x\colon \tau_1 \vdash_{\mathrm{pl}} e' \equiv_{\tau_2} [\mu(e')]$ are derivable provided $x\colon \tau_1 \vdash_{\mathrm{pl}} e' \downarrow_{T\tau_2}$.

3. Extending the Simple Metalanguage

So far we have considered only languages and formal systems for *monadic terms* $x\colon \tau_1 \vdash e\colon \tau_2$, having exactly one free variable (occurring once). In this section we want to extend these languages (and formal systems) by allowing *algebraic terms* $x_1\colon \tau_1, ..., x_n\colon \tau_n \vdash e\colon \tau$, having a finite number of free variables (occurring finitely many times) and investigate how this affects the interpretation and the structure on theories viewed as categories. For convenience in relating theories and categories with additional structure, we also allow types to be closed w.r.t. finite products;[4] in

[4] If the metalanguage does not have finite products, we conjecture that its theories would no longer correspond to categories with finite products and a strong monad (even by taking as objects contexts and/or the Karoubi envelope, used in Scott (1980) to associate a cartesian closed category to an untyped λ-theory), but instead to *multicategories* with a Kleisli triple. We felt the greater generality (of not having products in the metalanguage) was not worth the mathematical complications.

particular a typing context $x_1: \tau_1, \ldots, x_n: \tau_n$ can be identified with a type. In general, the interpretation of an algebraic term $x_1: \tau_1, \ldots, x_n: \tau_n \vdash e: \tau$ in a category (with finite products) is a morphism from $(\llbracket\tau_1\rrbracket \times \cdots \times \llbracket\tau_n\rrbracket)$ to $\llbracket\tau\rrbracket$.

The extension of monadic equational logic to algebraic terms is equational logic, whose theories correspond to categories with finite products. We will introduce the *metalanguage*, i.e., the extension of the simple metalanguage described in Section 2.2 to algebraic terms, and show that its theories correspond to categories with finite products and a *strong monad*, i.e., a monad and a natural transformation $t_{A,B}: A \times TB \to T(A \times B)$. Intuitively $t_{A,B}$ transforms a pair value-computation into a computation of a pair of values, as follows:

$$a: A, c: TB \xmapsto{t_{A,B}} (\text{let } y \Leftarrow c \text{ in } [\langle a, y\rangle]): T(A \times B)$$

Remark 3.1. To understand why a category with finite products and a monad is not enough to interpret the metalanguage (and where the natural transformation t is needed), one has to look at the interpretation of a let-expression

$$\text{let}\ \frac{\Gamma \vdash_{\text{ml}} e_1: T\tau_1 \quad \Gamma, x: \tau_1 \vdash_{\text{ml}} e_2: T\tau_2}{\Gamma \vdash_{\text{ml}} (\text{let}_T\, x \Leftarrow e_1 \text{ in } e_2): T\tau_2}$$

where Γ is a typing context. Let $g_1: c \to Tc_1$ and $g_2: c \times c_1 \to Tc_2$ be the interpretations of $\Gamma \vdash_{\text{ml}} e_1: T\tau_1$ and $\Gamma, x: \tau_1 \vdash_{\text{ml}} e_2: T\tau_2$, respectively, where c is the interpretation of the typing context Γ and c_i is the interpretation of the type τ_i; then the interpretation of $\Gamma \vdash_{\text{ml}} (\text{let}_T\, x \Leftarrow e_1 \text{ in } e_2): T\tau_2$ ought to be a morphism $g: c \to Tc_2$. If (T, η, μ) is the *identity monad*, i.e., T is the identity functor over $\mathscr{C}$ and η and μ are the identity natural transformation over T, then computations get identified with values. In this case $(\text{let}_T\, x \Leftarrow e_1 \text{ in } e_2)$ can be replaced by $[e_1/x]e_2$, so g is simply $\langle \text{id}_c, g_1\rangle$; $g_2: c \to c_2$. In the general case Table 3 suggests that $-;-$ above is indeed composition in the Kleisli category, therefore $\langle \text{id}_c, g_1\rangle$; g_2 should be replaced by $\langle \text{id}_c, g_1\rangle$; g_2^*. But in $\langle \text{id}_c, g_1\rangle$: g_2^* there is a type mismatch, since the codomain of $\langle \text{id}_c, g_1\rangle$ is $c \times Tc_1$, while the domain of Tg_1 is $T(c \times c_1)$. The natural transformation $\mathrm{t}_{A,B}: A \times TB \to T(A \times B)$ mediates between these two objects, so that g can be defined as $\langle \text{id}_c, g_1\rangle; t_{c,c_1}; g_2^*$.

DEFINITION 3.2. A *strong monad* over a category $\mathscr{C}$ with (explicitly given) finite products is a monad (T, η, μ) together with a natural transformation $\mathrm{t}_{A,B}$ from $A \times TB$ to $T(A \times B)$ s.t.

$$\begin{tikzcd} TA & \\ 1 \times TA \arrow[u, "r_{TA}"] \arrow[r, "\mathrm{t}_{1,A}"] & T(1 \times A) \arrow[ul, "Tr_A"'] \end{tikzcd}$$

$$\begin{tikzcd} (A \times B) \times TC \arrow[r, "\mathrm{t}_{A \times B, C}"] \arrow[d, "\alpha_{A,B,TC}"'] & T((A \times B) \times C) \arrow[dr, "T\alpha_{A,B,C}"] & \\ A \times (B \times TC) \arrow[r, "\mathrm{id}_A \times \mathrm{t}_{B,C}"] & A \times T(B \times C) \arrow[r, "\mathrm{t}_{A, B \times C}"] & T(A \times (B \times C)) \end{tikzcd}$$

$$\begin{tikzcd} A \times B \arrow[d, "\mathrm{id}_A \times \eta_B"'] \arrow[dr, "\eta_{A \times B}"] & & \\ A \times TB \arrow[r, "\mathrm{t}_{A,B}"] & T(A \times B) & \\ A \times T^2B \arrow[u, "\mathrm{id}_A \times \mu_B"] \arrow[r, "\mathrm{t}_{A,TB}"] & T(A \times TB) \arrow[r, "T\mathrm{t}_{A,B}"] & T^2(A \times B) \arrow[ul, "\mu_{A \times B}"'] \end{tikzcd}$$

where r and α are the natural isomorphisms

$$r_A\colon (1 \times A) \to A, \qquad \alpha_{A,B,C}\colon (A \times B) \times C \to A \times (B \times C).$$

Remark 3.3. The diagrams above are taken from Kock (1972), where a characterization of strong monads is given in terms of $\mathscr{C}$-enriched categories (see Kelly, 1982). Kock fixes a commutative monoidal closed category $\mathscr{C}$ (in particular a cartesian closed category), and in this setup he establishes a one-one correspondence between *strengths* $\mathrm{st}_{A,B}\colon B^A \to (TB)^{TA}$ and *tensorial strengths* $\mathrm{t}_{A,B}\colon A \otimes TB \to T(A \otimes B)$ for a endofunctor T over $\mathscr{C}$ (see Theorem 1.3 in Kock, 1972). Intuitively a strength $\mathrm{st}_{A,B}$ *internalises* the action of T on morphisms from A to B, and more precisely it makes (T, st) a $\mathscr{C}$-enriched endofunctor on $\mathscr{C}$ enriched over itself (i.e., the hom-object $\mathscr{C}(A, B)$ is B^A). In this setting the diagrams of Definition 3.2 have the following meaning:

- the first two diagrams are (1.7) and (1.8) in Kock (1972), saying that t is a tensorial strength of T. So T can be made into a $\mathscr{C}$-enriched endofunctor.
- the last two diagrams say that $\eta\colon \mathrm{Id}_{\mathscr{C}} \to T$ and $\mu\colon T^2 \to T$ are $\mathscr{C}$-enriched natural transformations, where $\mathrm{Id}_{\mathscr{C}}$, T, and T^2 are enriched in the obvious way (see Remark 1.4 in Kock, 1972).

There is another purely categorical characterisation of strong monads, suggested to us by G. Plotkin, in terms of $\mathscr{C}$-indexed categories (see

Johnstone and Pare, 1978). Both characterisations are instances of a general methodological principle for studying programming languages (or logics) categorically (see Moggi, 1989b):

> when studying a complex language the 2-category **Cat** of small categories, functors, and natural transformations may not be adequate; however, one may replace **Cat** with a different 2-category, whose objects captures better some *fundamental structure* of the language, while *less fundamental structure* can be modelled by *2-categorical concepts*.

Monads are a 2-categorical concept, so we except notions of computations for a complex language to be modelled by monads in a suitable 2-category.

The first characterisation takes a commutative monoidal closed structure on $\mathscr{C}$ (used in Lafont (1988) and Seely (1987) to model a fragment of *linear logic*), so that $\mathscr{C}$ can be enriched over itself. Then a strong monad over a cartesian closed category $\mathscr{C}$ is just a monad over $\mathscr{C}$ in the 2-category of $\mathscr{C}$-enriched categories.

The second characterisation takes a class $\mathscr{D}$ of display maps over $\mathscr{C}$ (used in Myland and Pitts, 1987) to model *dependent types*), and defines a $\mathscr{C}$-indexed category $\mathscr{C}/_{\mathscr{D}}\_$. Then a strong monad over a category $\mathscr{C}$ with finite products amounts to a monad over $\mathscr{C}/_{\mathscr{D}}\_$ in the 2-category of $\mathscr{C}$-indexed categories, where $\mathscr{D}$ is the class of first projections (corresponding to constant type dependency).

In general the natural transformation t has to be given explicitly as part of the additional structure. However, t is uniquely determined (but it may not exist) by T and the cartesian structure on $\mathscr{C}$, when $\mathscr{C}$ has enough points.

PROPOSITION 3.4 (Uniqueness). *If (T, η, μ) is a monad over a category $\mathscr{C}$ with finite products and enough points (i.e., $\forall h\colon 1 \to A.h; f = h; g$ implies $f = g$ for any $f, g\colon A \to B$), then $(T, \eta, \mu, \mathrm{t})$ is a strong monad over $\mathscr{C}$ if and only if $\mathrm{t}_{A,B}$ is the unique family of morphisms s.t. for all points $a\colon 1 \to A$ and $b\colon 1 \to TB$*

$$\langle a, b\rangle; \mathrm{t}_{A,B} = b; T(\langle !_B; a, \mathrm{id}_B\rangle),$$

where $!_B\colon B \to 1$ is the unique morphism from B to the terminal object.

Proof. Note that there is at most one $\mathrm{t}_{A,B}$ s.t. $\langle a, b\rangle$; $\mathrm{t}_{A,B} = b$; $T(\langle !_B; a, \mathrm{id}_B\rangle)$ for all points $a\colon 1 \to A$ and $b\colon 1 \to TB$, because $\mathscr{C}$ has enough points.

First we show that if $(T, \eta, \mu, \mathrm{t})$ is a strong monad, then $\mathrm{t}_{A,B}$ satisfies the equation above. By naturality of t and by the first diagram in Definition 3.2 the following diagram commutes:

$$\begin{tikzcd}
1 \arrow[r, "\langle a, b\rangle"] \arrow[dr, "\langle \mathrm{id}_1, b\rangle"'] & A \times TB \arrow[r, "\mathrm{t}_{A,B}"] & T(A \times B) \\
& 1 \times TB \arrow[u, "a \times \mathrm{id}_{TB}"] \arrow[r, "\mathrm{t}_{1,B}"] \arrow[dr, "r_{TB}"'] & T(1 \times B) \arrow[u, "T(a + \mathrm{id}_B)"'] \arrow[d, "Tr_B"] \\
& & TB
\end{tikzcd}$$

Since r_B is an isomorphism (with inverse $\langle !_B, \mathrm{id}_B\rangle$), the two composite morphisms $\langle a, b\rangle$; $\mathrm{t}_{A,B}$ and $\langle \mathrm{id}_1, b\rangle$; r_{TB}; $T(r_B^{-1})$; $T(a \times \mathrm{id}_B)$ from 1 to $T(A \times B)$ must coincide. But the second composition can be rewritten as b; $T(\langle !_B; a, \mathrm{id}_B\rangle)$.

Second we have to show that if t is the unique family of morphisms satisfying the equation above, then $(T, \eta, \mu, \mathrm{t})$ is a strong monad. This amount to proving that t is a natural transformation and that the three diagrams in Definition 3.2 commute. The proof is a tedious diagram chasing, which relies on $\mathscr{C}$ having enough points. For instance, to prove that $\mathrm{t}_{1,A}$; $Tr_A = r_{TA}$ it is enough to show that $\langle \mathrm{id}_1, a\rangle$; $\mathrm{t}_{1,A}$; $Tr_A = \langle \mathrm{id}_1, a\rangle$; r_{TA} for all points a: $1 \to A$. ∎

EXAMPLE 3.5. We go through the monads given in Example 1.4 and show that they have a tensorial strength.

- **partiality** $TA = A_\perp (= A + \{\perp\})$

$\mathrm{t}_{A,B}(a, \perp) = \perp$ and $\mathrm{t}_{A,B}(a, b) = \langle a, b\rangle$ (when $b \in B$)

- **nondeterminism** $TA = \mathscr{P}_{\mathrm{fin}}(A)$

$\mathrm{t}_{A,B}(a, c) = \{\langle a, b\rangle \mid b \in c\}$

- **side-effects** $TA = (A \times S)^S$

$\mathrm{t}_{A,B}(a, c) = (\lambda s\colon S.(\mathrm{let}\langle b, s'\rangle = c(s)\ \mathrm{in}\ \langle\langle a, b\rangle, s'\rangle))$

- **exceptions** $TA = (A + E)$

$t_{\mathrm{A},B}(a, \mathrm{inr}(e)) = \mathrm{inr}(e)$ (when $e \in E$) and

$t_{A,B}(a, \mathrm{inl}(b)) = \mathrm{inl}(\langle a, b\rangle)$ (when $b \in B$)

- **continuations** $TA = R^{(R^A)}$

$t_{A,B}(a, c) = (\lambda k\colon R^{A \times B}.c(\lambda b\colon B.k(\langle a, b\rangle)))$

- **interactive input** $TA = (\mu\gamma.A + \gamma^U)$

$\mathrm{t}_{A,B}(a, c)$ is the tree obtained by replacing leaves of c labelled by b with the leaf labelled by $\langle a, b\rangle$

- **interactive output** $TA = (\mu\gamma.A + (U \times \gamma))$

$\mathrm{t}_{A,B}(a, \langle s, b\rangle) = \langle s, \langle a, b\rangle\rangle$.

Remark 3.6. The tensorial strength t induces a natural transformation $\psi_{A,B}$ from $TA \times TB$ to $T(A \times B)$, namely

$$\psi_{A,B} = c_{TA,TB};\, \mathrm{t}_{TB,A};\, (c_{TB,A};\, \mathrm{t}_{A,B})^*$$

where c is the natural isomorphism $c_{A,B}\colon A \times B \to B \times A$.

The morphism $\psi_{A,B}$ has the correct domain and codomain to interpret the pairing of a computation of type A with one of type B, obtained by first evaluating the first argument and then the second, namely

$$c_1\colon TA,\, c_2\colon TB \xmapsto{\psi_{A,B}} (\text{let } x \Leftarrow c_1 \text{ in } (\text{let } y \Leftarrow c_2 \text{ in } [\langle x, y\rangle])) \colon T(A \times B)$$

There is also a dual notion of pairing, $\tilde{\psi}_{A,B} = c_{TA,TB};\, \psi_{B,A};\, Tc_{B,A}$ (see Kock, 1972), which amounts to first evaluating the second argument and then the first.

3.1. *Interpretation and formal system*

We are now in a position to give the metalanguage for algebraic terms, its interpretation and inference rules.

TABLE 8

Interpretation of Types in the Metalanguage

Rule	Syntax		Semantics
A			
	$\overline{\vdash_{\mathrm{ml}} A \text{ type}}$	$=$	$[\![A]\!]$
T			
	$\vdash_{\mathrm{ml}} \tau \text{ type}$	$=$	c
	$\overline{\vdash_{\mathrm{ml}} T\tau \text{ type}}$	$=$	Tc
1			
	$\overline{\vdash_{\mathrm{ml}} 1 \text{ type}}$	$=$	1
$\times$			
	$\vdash_{\mathrm{ml}} \tau_1 \text{ type}$	$=$	c_1
	$\vdash_{\mathrm{ml}} \tau_2 \text{ type}$	$=$	c_2
	$\overline{\vdash_{\mathrm{ml}} \tau_1 \times \tau_2 \text{ type}}$	$=$	$c_1 \times c_2$
$\varnothing$			
	$\vdash_{\mathrm{ml}} \tau_i \text{ type } (1 \leqslant i \leqslant n)$	$=$	c_i
	$\overline{x_1 : \tau_1, \ldots, x_n : \tau_n \vdash}$	$=$	$c_1 \times \cdots \times c_n$

Definition 3.7 (*metalanguage*). An interpretation $[\![-]\!]$ of the metalanguage in a category $\mathscr{C}$ with terminal object $!_A\colon A \to 1$, binary products $\pi_i^{A_1,A_2}\colon A_1 \times A_2 \to A_i$, and a strong monad $(T, \eta, \mu, \mathrm{t})$ is parametric in an interpretation of the symbols in the signature and is defined by induction on the derivation of well-formedness for types (see Table 8), terms and equations (see Table 9).

Finite products $\pi_i^{A_1,\ldots,A_n}\colon A_1 \times \cdots \times A_n \to A_i$ used to interpret contexts and variables are defined by inductin on n:

TABLE 9

Interpretation of Terms in the Metalanguage

Rule	Syntax		Semantics
var_i			
	$\vdash_{\mathrm{ml}} \tau_i \text{ type } (1 \leqslant i \leqslant n)$	$=$	c_i
	$x_1\colon \tau_1, \ldots, x_n\colon \tau_n \vdash x_i\colon \tau_i$	$=$	$\pi_i^{c_1,\ldots,c_n}$
$*$			
	$\Gamma \vdash *\colon 1$	$=$	$!_{[\![\Gamma]\!]}$
$\langle\ \rangle$			
	$\Gamma \vdash e_1\colon \tau_1$	$=$	g_1
	$\Gamma \vdash e_2\colon \tau_2$	$=$	g_2
	$\Gamma \vdash \langle e_1, e_2 \rangle\colon \tau_1 \times \tau_2$	$=$	$\langle g_1, g_2 \rangle$
π_i			
	$\Gamma \vdash e\colon \tau_1 \times \tau_2$	$=$	g
	$\Gamma \vdash \pi_i(e)\colon \tau_1$	$=$	$g; \pi_i^{[\![\tau_1]\!],[\![\tau_2]\!]}$
$f\colon \tau_1 \to \tau_2$			
	$\Gamma \vdash_{\mathrm{ml}} e_1\colon \tau_1$	$=$	g
	$\Gamma \vdash_{\mathrm{ml}} f(e_1)\colon \tau_2$	$=$	$g; [\![f]\!]$
$[-]_T$			
	$\Gamma \vdash_{\mathrm{ml}} e\colon \tau$	$=$	g
	$\Gamma \vdash_{\mathrm{ml}} [e]_T\colon T\tau$	$=$	$g; \eta_{[\![\tau]\!]}$
let			
	$\Gamma \vdash_{\mathrm{ml}} e_1\colon T\tau_1$	$=$	g_1
	$\Gamma, x\colon \tau_1 \vdash_{\mathrm{ml}} e_2\colon T\tau_2$	$=$	g_2
	$\Gamma \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x \Leftarrow e_1 \text{ in } e_2)\colon T\tau_2$	$=$	$\langle \mathrm{id}_{[\![\Gamma]\!]}, g_1 \rangle; t_{[\![\Gamma]\!],[\![\tau_1]\!]}; g_2^*$
eq			
	$\Gamma \vdash_{\mathrm{ml}} e_1\colon \tau$	$=$	g_1
	$\Gamma \vdash_{\mathrm{ml}} e_2\colon \tau$	$=$	g_2
	$\Gamma \vdash_{\mathrm{ml}} e_1 =_\tau e_2$	$\Leftrightarrow$	$g_1 = g_2$

$$
\begin{array}{ll}
0 & A_1 \times \cdots \times A_0 \triangleq 1 \\
n+1 & A_1 \times \cdots \times A_{n+1} \triangleq (A_1 \times \cdots \times A_n) \times A_{n-1} \\
& \quad -\pi_{n+1}^{A_1, \ldots, A_{n+1}} = \pi_2^{(A_1 \times \cdots \times A_n), A_{n+1}} \\
& \quad -\pi_i^{A_1, \ldots, A_{n+1}} = \pi_1^{(A_1 \times \cdots \times A_n), A_{n+1}}; \pi_i^{A_1, \ldots, A_n}
\end{array}
$$

The inference rules for the metalanguage (see Table 10) are divided into three groups:

- general rules for many sorted equational logic
- rules for finite products
- rules for T.

PROPOSITION 3.8. *Every theory $\mathcal{T}$ of the metanlanguage, viewed as a category $\mathcal{F}(\mathcal{T})$, is equipped with finite products and a strong monad whose tensorial strength is*

$$\mathrm{t}_{\tau_1, \tau_2} = [x \colon \tau_1 \times T\tau_2 \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_2 \Leftarrow \pi_2 x \text{ in } [\langle \pi_1 x, x_2 \rangle]_T) \colon T(\tau_1 \times \tau_2)]_{\mathcal{T}}.$$

Proof. Similar to that of Proposition 2.5. ∎

Once we have a metalanguage for algebraic terms it is straightforward to add data types characterised by *universal properties* and to extend the categorical semantics accordingly.[5] For instance, if we want to have function spaces, then we simply require the category $\mathscr{C}$ (where the metalanguage is interpreted) to have exponentials B^A and add the inference rules for the simply typed λ-calculus (see Table 11) to those for the metalanguage. From a programming language perspective the situation is more delicate. For instance, the semantics of functional types should reflect the choice of *calling mechanism*:[6]

- in call-by-value a procedure of type $A \to B$ expects a value of type A and computes a result of type B, so the interpretation of $A \to B$ is $(TB)^A$;
- in call-by-name a procedure of type $A \to B$ expects a computation of type A, which is evaluated only when needed, and computes a result of type B, so the interpretation of $A \to B$ is $(TB)^{TA}$.

In both cases the only exponentials needed to interpret the functional types of a programming language are of the form $(TB)^A$. By analogy with partial cartesian closed categories (pccc), where only *p-exponentials* are required

[5] The next difficult step in extending the metalanguage is the combination of dependent types and computations, which is currently under investigation.

[6] Call-by-need does not have a simple categorical semantics, since the environment in which an expression is evaluated may itself undergo evaluation.

TABLE 10

Inference Rules of the Metalanguage

refl $\dfrac{\Gamma \vdash e:\tau}{\Gamma \vdash e =_\tau e}$

symm $\dfrac{\Gamma \vdash e_1 =_\tau e_2}{\Gamma \vdash e_2 =_\tau e_1}$

trans $\dfrac{\Gamma \vdash e_1 =_\tau e_2 \quad \Gamma \vdash e_2 =_\tau e_3}{\Gamma \vdash e_2 =_\tau e_3}$

congr $\dfrac{\Gamma \vdash e_1 =_{\tau_1} e_2}{\Gamma \vdash f(e_1) =_{\tau_2} f(e_2)} f\colon \tau_1 \to \tau_2$

subst $\dfrac{\Gamma \vdash e:\tau \quad \Gamma, x:\tau \vdash \phi}{\Gamma \vdash [e/x]\phi}$

Inference Rules of Many Sorted Equational Logic

$1.\eta$ $\Gamma \vdash * =_1 x$

$\langle\ \rangle.\xi$ $\dfrac{\Gamma \vdash e_1 =_{\tau_1} e_1' \quad \Gamma \vdash e_2 =_{\tau_2} e_2'}{\Gamma \vdash \langle e_1, e_2\rangle =_{\tau_1\times\tau_2} \langle e_1', e_2'\rangle}$

$\times.\beta$ $\dfrac{\Gamma \vdash e_1:\tau_1 \quad \Gamma \vdash e_2:\tau_2}{\Gamma \vdash \pi_i(\langle e_1, e_2\rangle) =_{\tau_i} e_i}$

$\times.\eta$ $\dfrac{\Gamma \vdash e:\tau_1\times\tau_2}{\Gamma \vdash \langle \pi_1(e), \pi_2(e)\rangle =_{\tau_1\times\tau_2} e}$

Rules for Product Types

$[-].\xi$ $\dfrac{\Gamma \vdash_{\mathrm{ml}} e_1 =_\tau e_2}{\Gamma \vdash_{\mathrm{ml}} [e_1]_T =_{T\tau} [e_2]_T}$

$\mathrm{let}.\xi$ $\dfrac{\Gamma \vdash_{\mathrm{ml}} e_1 =_{T\tau_1} e_2 \quad \Gamma, x:\tau_1 \vdash_{\mathrm{ml}} e_1' =_{T\tau_2} e_2'}{\Gamma \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x \Leftarrow e_1 \text{ in } e_1') =_{T\tau_2} (\mathrm{let}_T\, x \Leftarrow e_2 \text{ in } e_2')}$

ass $\dfrac{\Gamma \vdash_{\mathrm{ml}} e_1 : T\tau_1 \quad \Gamma, x_1:\tau_1 \vdash_{\mathrm{ml}} e_2 : T\tau_2 \quad \Gamma, x_2:\tau_2 \vdash_{\mathrm{ml}} e_3 : T\tau_3}{\Gamma \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_2 \Leftarrow (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } e_2) \text{ in } e_3) =_{T\tau_3} (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } (\mathrm{let}_T\, x_2 \Leftarrow e_2 \text{ in } e_3))}$

$T.\beta$ $\dfrac{\Gamma \vdash_{\mathrm{ml}} e_1 : \tau_1 \quad \Gamma, x_1:\tau_1 \vdash_{\mathrm{ml}} e_2 : T\tau_2}{\Gamma \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_1 \Leftarrow [e_1]_T \text{ in } e_2) =_{T\tau_2} [e_1/x_1]e_2}$

$T.\eta$ $\dfrac{\Gamma \vdash_{\mathrm{ml}} e_1 : T\tau_1}{\Gamma \vdash_{\mathrm{ml}} (\mathrm{let}_T\, x_1 \Leftarrow e_1 \text{ in } [x_1]_T) =_{T\tau_1} e_1}$

Rules for Computational Types

TABLE 11

Rules for Function Spaces

app.ξ	$\dfrac{\Gamma \vdash e_1 =_{\tau_1} e_1' \quad \Gamma \vdash e =_{\tau_1 \to \tau_2} e'}{\Gamma \vdash ee_1 =_{\tau_2} e'e_1'}$
$\lambda.\xi$	$\dfrac{\Gamma, x: \tau_1 \vdash e_1 =_{\tau_2} e_2}{\Gamma \vdash (\lambda x: \tau_1.e_1) =_{\tau_1 \to \tau_2} (\lambda x: \tau_1.e_2)}$
$\to.\beta$	$\dfrac{\Gamma \vdash e_1: \tau_1 \quad \Gamma, x: \tau_1 \vdash e_2: \tau_2}{\Gamma \vdash (\lambda x: \tau_1.e_2)e_1 =_{\tau_2} [e_1/x]e_2}$
$\to.\eta$	$\dfrac{\Gamma \vdash e: \tau_1 \to \tau_2}{\Gamma \vdash (\lambda x: \tau_1.ex) =_{\tau_1 \to \tau_2} e}\, x \notin \mathrm{DV}(\Gamma)$

to exist (see Moggi (1986), Rosolini (1986)), we adopt the following definition of λ_c-model:

DEFINITION 3.9. A *λ_c-model is a category $\mathscr{C}$ with finite products, a strong monad $(T, \eta, \mu, \mathrm{t})$ satisfying the mono requirement (i.e., η_A mono for every $A \in \mathscr{C}$), and T-exponential $(TB)^A$ for every A, $B \in \mathscr{C}$.*

Remark 3.10. The definition of λ_c-model generalises that of pccc, in the sense that every pccc can be viewed as a λ_c-model. By analogy with p-exponentials, a T-exponential can be defined by giving an isomorphism $\mathscr{C}_T(C \times A, B) \cong \mathscr{C}(C, (TB)^A)$ natural in $C \in \mathscr{C}$. We refer to Moggi (1989c) for the interpretation of a call-by-value programming language in a λ_c-model and the corresponding formal system, the λ_c-calculus.

4. Strong Monads over a Topos

In this section we show that, as far as monads or strong monads are concerned, we can assume w.l.o.g. that they are over a topos (see Theorem 4.9). The proof of Theorem 4.9 involves non-elementary notions from category theory, and we postpone it until after discussing some applications, with particular emphasis on further extensions of the metalanguage and on conservative extension results.

Let us take as formal system for toposes the type theory described in Lambek and Scott (1986). This is a many sorted intuitionistic higher order logic with equality and with a set of types satisfying the following closure properties:[7]

[7] Lambek and Scott do not require closure under function spaces and subsets $\{x \in A \mid \phi(x)\}$.

- the terminal object 1, the natural number object N, and the subobject classifier Ω are types
- if A is a type, then the power object PA is a type
- if A and B are types, then the binary product $A \times B$ and the function space $A \to B$ are types
- if A is a type and ϕ: $A \to \Omega$ is a predicate, then $\{x \in A \mid \phi(x)\}$ is a type.

Notation 4.1. We introduce some notational conventions for formal systems:

- ML_T is the metalanguage for algebraic terms, whose set of types is closed under terminal object, binary products, and TA;
- $\lambda\mathrm{ML}_T$ is the extension of ML_T with function spaces $A \to B$ (interpreted as exponentials);
- HML_T is the type theory described above (see Lambek and Scott, 1986) extended with objects of computations TA;
- PL is the programming language for algebraic terms (see Moggi, 1989c);
- λ_cPL is teh extension of PL with function spaces $A \to B$ (interpreted as T-exponentials), called λ_c-calculus in Moggi (1989c).

DEFINITION 4.2. We say that a formal system $(L_2, \vdash_2)$, where $\vdash_2 \subseteq \mathscr{P}(L_2) \times L_2$ is a formal consequence relation[8] over L_2, is a *conservative* extension of $(L_1, \vdash_1)$ provided $L_1 \subseteq L_2$ and $\vdash_1$ is the restriction of $\vdash_2$ to $\mathscr{P}(L_1) \times L_1$.

THEOREM 4.3. *HML_T is a conservative extension of ML_T and $\lambda\mathrm{ML}_T$. In particular $\lambda\mathrm{ML}_T$ is a conservative extension of ML_T.*

Proof. The first result follows from Theorem 4.9, which implies that for every model $\mathscr{C}$ of ML_T the Yoneda embedding maps the interpretation of an ML_T-term in $\mathscr{C}$ to its interpretation in $\hat{\mathscr{C}}$, and the faithfulness of the Yoneda embedding, which implies that two ML_T-terms have the same interpretation in $\mathscr{C}$ iff they have the same interpretation in $\hat{\mathscr{C}}$. The second result follows, because the Yoneda embedding preserves function spaces. The third conservative extension result follows immediately from the first two. ∎

The above result means that we can think of computations naively in terms of sets and functions, provided we treat them intuitionistically, and can use the full apparatus of higher-order (intuitionistic) logic instead of the less expressive many sorted equational logic.

[8] For instance, in the case of $ML_{\mathscr{T}}$ the elements of L are well-formed equality judgements $\Gamma \vdash_{\mathrm{ml}} e_1 =_\tau e_2$ and $P \vdash C$ iff there exists a derivation of C, where all assumptions are in P.

Before giving a conservative extension result for the programming language, we have to express the mono requirement, equivalence and existence in HML_T. The idea is to extend the translation from PL-terms to ML_T-terms given in Definition 2.7 and exploit the increased expressiveness of HML_T over ML_T to axiomatise the mono requirement and translate existence and equivalence assertions (see Remark 2.1):

- the **mono requirement** for τ, i.e., that η_τ be mono, is axiomatised by

$$\mathrm{mono}.\tau\ (\forall x, y: \tau.[x]_T =_{T\tau} [y]_T \to x =_\tau y)$$

- the **equalising requirement** for τ, i.e. that η_τ be the equaliser of $T(\eta_\tau)$ and $\eta_{T\tau}$, is axiomatised by $(\mathrm{mono}.\tau)$ and the axiom

$$\mathrm{eqls}.\tau\ (\forall x: T\tau.[x]_T =_{T^2\tau} (\mathrm{let}_T\ y \Leftarrow x \text{ in } [[y]_T]_T) \to (\exists!\, y: \tau.x =_{T\tau} [y]_T))$$

- the **translation** $-^\circ$ is extended to assertions and functional types as follows:

— $(e_1 \equiv_\tau e_2)^\circ \triangleq e_1^\circ =_{T\tau} e_2^\circ$
— $(e_1 \downarrow_\tau)^\circ \triangleq (\exists!\, x: \tau.e_1^\circ =_{T\tau} [x]_T)$
— $(\tau_1 \rightharpoonup \tau_2)^\circ \triangleq \tau_1^\circ \to T\tau_2^\circ$.

THEOREM 4.4. $\mathrm{HML}_T + \{(\mathrm{mono}.\tau) \mid \tau$ *type of* $\mathrm{PL}\}$ (*i.e., is built using only base types*, 1, TA, *and* $A \times B$) *is a conservative extension of* PL (*after translation*). *Similarly*, $\mathrm{HML}_T + \{(\mathrm{mono}.\tau) \mid \tau$ *type of* $\lambda_c\mathrm{PL}\}$ (*i.e.*, τ *is built using only base types*, 1, TA, $A \times B$ *and* $A \to B$) *is a conservative extension of* λ_cPL (*after translation*).

Proof. The proof proceeds as in the previous theorem. The only additional step is to show that for every type τ of PL (or λ_cPL) the axiom $(\mathrm{mono}.\tau)$ holds in $\hat{\mathscr{C}}$, under the assumption that $\mathscr{C}$ satisfies the mono requirement. Let c be the interpretation of τ in $\mathscr{C}$ (therefore Yc is the interpretation of τ in $\hat{\mathscr{C}}$), then the axiom $(\mathrm{mono}.\tau)$ holds in $\hat{\mathscr{C}}$ provided $\hat{\eta}_{Yc}$ is a mono. η_c is mono (by the mono requirement), so $\hat{\eta}_{Yc} = Y(\eta_c)$ is mono (as Y preserves monos). ∎

In the theorem above only types from the programming language have to satisfy the mono requirement. Indeed, $\mathrm{HML}_T + \{(\mathrm{mono}.\tau) \mid \tau$ type of $\mathrm{HML}_T\}$ is not a conservative extension of PL (or λ_cPL).

LEMMA 4.5. *If* (T, η, μ) *is a monad over a topos* $\mathscr{C}$ *satisfying the mono requirement, then it satisfies also the equalising requirement.*

Proof. See Lemma 6 on p. 110 of Barr and Wells (1985). ∎

In other words, for any type τ the axiom (eqls.τ) is derivable in HML_T from the set of axioms $\{(\mathrm{mono}.\tau) \mid \tau$ type of $\mathrm{HML}_T\}$. In general, when $\mathscr{C}$ is not a topos, the mono requirement does not entail the equalising requirement; one can easily define strong monads (over a Heyting algebra) that satisfy the mono but not the equalising requirement (just take $T(A) = A \vee B$, for some element $B \neq \bot$ of the Heyting algebra). In terms of formal consequence relation this means that in HML_T + mono requirement the existence assertion $\Gamma \vdash_{\mathrm{pl}} e \downarrow_\tau$ is derivable from $\Gamma \vdash_{\mathrm{pl}} [e] \equiv_{T\tau} (\mathrm{let}\, x \Leftarrow e \,\mathrm{in}\, [x])$, while such derivation is not possible in $\lambda_c\mathrm{PL}$. We do not known whether HML_T + equalising requirement is a conservative extension of PL + equalising requirement, or whether $\lambda_c\mathrm{PL}$ is a conservative extension of PL.

A language which combines computations and higher order logic, like HML_T, seems to be the ideal framework for program logics that go beyond proving equivalence of programs, like Hoare's logic for partial correctness of imperative languages. In HML_T (as well as ML_T and PL) one can describe a programming language by introducing additional constant and axioms. In $\lambda\mathrm{ML}_T$ or $\lambda_c\mathrm{PL}$ such constants correspond to program-constructors, for instance:

- *lookup*: $L \to TU$, which given a location $l \in L$ produces the value of such location in the current store, and *update*: $L \times U \to T1$, which changes the current store by assigning to $l \in L$ the value $u \in U$;
- *if*: $Bool \times TA \times TA \to TA$ and *while*: $T(Bool) \times T1 \to T1$;
- *new*: $1 \to TL$, which returns a newly created location;
- *read*: $1 \to TU$, which computes a value by reading it from the input, and *write*: $U \to T1$, which writes a value $u \in U$ on the output.

In HML_T one can describe also a program logic, by adding constants $p\colon TA \to \Omega$ corresponding to properties of computations.

EXAMPLE 4.6. Let T be the monad for non-deterministic computations (see Example 1.4); then we can define a predicate *may*: $A \times TA \to \Omega$ such that $may(a, c)$ is true iff the value a is a possible outcome of the computation c (i.e. $a \in c$). However, there is a more uniform way of defining the *may* predicate of any type. Let $\Diamond\colon T\Omega \to \Omega$ be the predicate such that $\Diamond(X) = \top$ iff $\top \in X$, where Ω is the set $\{\bot, \top\}$ (note that $\Diamond(-) = may(\top, -)$). Then $may(a, c)$ can be defined as $\Diamond(\mathrm{let}_T\, x \Leftarrow c \,\mathrm{in}\, [a =_\tau x]_T)$.

The previous example suggests that predicates defined *uniformly* on computations of any type can be better described in terms of *modal operators* $\gamma\colon T\Omega \to \Omega$, relating a computation of truth values to a truth value. This possibility has not been investigated in depth, so we will give only a tentative definition.

DEFINITION 4.7. If (T, η, μ) is a monad over a topos $\mathscr{C}$, then a *T-modal operator* is a T-algebra $\gamma\colon T\Omega \to \Omega$, i.e.,

$$\begin{array}{ccccc} T^2\Omega & \xrightarrow{\mu\Omega} & T\Omega & \xleftarrow{\eta\Omega} & \Omega \\ {\scriptstyle T\gamma}\downarrow & & \downarrow{\scriptstyle \gamma} & \swarrow{\scriptstyle \mathrm{id}_{T\Omega}} & \\ T\Omega & \xrightarrow[\gamma]{} & \Omega & & \end{array}$$

where Ω is the subobject classifier in $\mathscr{C}$.

The commutativity of the two diagrams above can be expressed in the metalanguage:

- $x\colon \Omega \vdash \gamma([x]_T) \leftrightarrow x$
- $c\colon T^2\Omega \vdash \gamma(\text{let } x \Leftarrow c \text{ in } x) \leftrightarrow \gamma(\text{let } x \Leftarrow c \text{ in } [\gamma(x)]_T)$,

We consider some examples and *non-examples* of modal operators.

EXAMPLE 4.8. For the monad T of non-deterministic computations (see Example 1.4) there are only two modal operators $\Box$ and $\Diamond$:

- $\Box(X) = \bot$ iff $\bot \in X$;
- $\Diamond(X) = \top$ iff $\top \in X$.

Given a nondeterministic computation e of type τ and a predicate $A(x)$ over τ, i.e., a term of type Ω, then $\Box(\text{let}_T\, x \Leftarrow e \text{ in } [A(x)]_T)$ is true iff all possible results of e satisfy $A(x)$.

For the monad T of computations with side-effects (see Example 1.4) there is an operator $\Box\colon (\Omega \times S)^S \to \Omega$ that can be used to express Hoare's triples:

- $\Box f = \top$ iff for all $s \in S$ there exists $s' \in S$ s.t. $fs = \langle \top, s' \rangle$

  this operator does not satisfy the second equivalence, as only one direction is valid, namely $c\colon T^2\Omega \vdash \gamma(\text{let } x \Leftarrow c \text{ in } [\gamma(x)]_T) \to \gamma(\text{let } x \Leftarrow c \text{ in } x)$.

Let $P\colon U \to \Omega$ and $Q\colon U \times U \to \Omega$ be predicates over storable values, $e \in T1$ a computation of type 1, and $x, y \in L$ locations. The intended meaning of the triple $\{P(x)\}\, e\{Q(x, y)\}$ is "if in the initial state the content u of x satisfies $P(u)$, then in the final state (i.e., after executing e) the content v of y satisfies $Q(u, v)$". This intended meaning can be expressed formally in terms of the modal operator $\Box$ and the program-constructors *lookup* and *update* as follows:

$$\forall u\colon U.\, P(u) \to \Box(\text{let}_T\, v \Leftarrow (\mathit{update}(x, u); e; \mathit{lookup}(y)) \text{ in } [Q(u, v)]_T)$$

where $-;-: TA \times TB \to TB$ is the derived operation $e_1; e_2 \triangleq (\text{let}_T\, x \Leftarrow e_1$ in $e_2)$ with x not free in e_2.

Finally, we state the main theorem and outline its proof. In doing so we assume that the reader is familiar with non-elementary concepts from category theory.

THEOREM 4.9. *Let $\mathscr{C}$ be a small category, $\hat{\mathscr{C}}$ the topos of presheaves over $\mathscr{C}$, and Y the Yoneda embedding of $\mathscr{C}$ into $\hat{\mathscr{C}}$. Then for every monad (T, η, μ) over $\mathscr{C}$, there exists a monad $(\hat{T}, \hat{\eta}, \hat{\mu})$ over $\hat{\mathscr{C}}$ such that the following diagram commutes*[9]

$$\begin{tikzcd} \mathscr{C} \arrow[r, "T"] \arrow[d, "Y"'] & \mathscr{C} \arrow[d, "Y"] \\ \hat{\mathscr{C}} \arrow[r, "\hat{T}"'] & \hat{\mathscr{C}} \end{tikzcd}$$

and for all $a \in \mathscr{C}$ the following equations hold:

$$\hat{\eta}_{Ya} = Y(\eta_a), \qquad \hat{\mu}_{Ya} = Y(\mu_a).$$

Moreover, for every strong monad $(T, \eta, \mu, \mathrm{t})$ over $\mathscr{C}$, there exists a natural transformation $\hat{\mathrm{t}}$ such that $(\hat{T}, \hat{\eta}, \hat{\mu}, \hat{\mathrm{t}})$ is a strong monad over $\hat{\mathscr{C}}$ and for all $a, b \in \mathscr{C}$ the following equation holds

$$\hat{\mathrm{t}}_{Ya, Yb} = Y(\mathrm{t}_{a,b}),$$

where we have implicitly assumed that the Yoneda embedding preserves finite products on the nose, i.e., the following diagrams commute

$$\begin{tikzcd} 1 \arrow[r, "1"] \arrow[dr, "1"'] & \mathscr{C} \arrow[d, "Y"] & \mathscr{C} \times \mathscr{C} \arrow[l, "\times"'] \arrow[d, "Y \times Y"] \\ & \hat{\mathscr{C}} & \hat{\mathscr{C}} \times \hat{\mathscr{C}} \arrow[l, "\times"] \end{tikzcd}$$

and for all $a, b \in \mathscr{C}$ the following equations hold:

$$!_{Ya} = Y(!_a), \qquad \pi_i^{Ya, Yb} = Y(\pi_i^{a,b}).$$

DEFINITION 4.10 (MacLane, 1971). Let $T: \mathscr{C} \to \mathscr{D}$ be a functor between

[9] This is a simplifying assumption. For our purposes it would be enough to have a natural isomorphism σ; T; $Y \to Y$; $\hat{T}$, but then the remaining equations have to be patched. For instance, the equation relating η and $\hat{\eta}$ would become $\hat{\eta}_{Ya} = Y(\eta_a); \sigma_a$.

two small categories and $\mathscr{A}$ a cocomplete category. Then the *left Kan extension* $L_T^{\mathscr{A}}: \mathscr{A}^{\mathscr{C}} \to \mathscr{A}^{\mathscr{D}}$ is the left adjoint of $\mathscr{A}^T$ and can be defined as

$$L_T^{\mathscr{A}}(F)(d) = \mathrm{Colim}^{\mathscr{A}}_{T\downarrow d}(\pi; F),$$

where $F: \mathscr{C} \to \mathscr{A}$, $d \in \mathscr{D}$, $T\downarrow d$ is the comma category whose objects are pairs $\langle c \in \mathscr{C}, f: Tc \to d\rangle$, $\pi: T\downarrow d \to \mathscr{C}$ is the projection functor (mapping a pair $\langle c, f\rangle$ to c), and $\mathrm{Colim}^{\mathscr{A}}_I: \mathscr{A}^I \to \mathscr{A}$ (with I a small category) is a functor mapping an I-diagram in $\mathscr{A}$ to its colimit.

The following proposition is a 2-categorical reformulation of Theorem 1.3.10 of Makkai and Reyes (1977). For the sake of simplicity, we use the strict notions of 2-functor and 2-natural transformation, although we should have used pseudo-functors and pseudo-natural transformations.

PROPOSITION 4.11. *Let* **Cat** *be the 2-category of small categories,* **CAT** *the 2-category of locally small categories, and* $-$: **Cat** $\to$ **CAT** *the inclusion 2-functor. Then the following* $\hat{-}$: **Cat** $\to$ **CAT** *is a 2-functor:*

- *if $\mathscr{C}$ is a small category, then $\hat{\mathscr{C}}$ is the topos of presheaves* $\mathbf{Set}^{\mathscr{C}^{op}}$
- *if $T: \mathscr{C} \to \mathscr{D}$ is a functor, then $\hat{T}$ is the left Kan extension* $L^{\mathbf{Set}}_{T^{op}}$
- *if $\sigma: S \to T: \mathscr{C} \to \mathscr{D}$ is a natural transformation and $F \in \hat{\mathscr{C}}$, then $\hat{\sigma}_F$ is the natural transformation corresponding to* $\mathrm{id}_{\hat{T}F}$ *via the sequence of steps*

$$\begin{tikzcd}
\hat{\mathscr{C}}(F, T^{op}; \hat{T}F) \arrow[d, "\hat{\mathscr{C}}(F, \sigma^{op}; \hat{T}FD)"] & \hat{\mathscr{D}}(\hat{T}F, \hat{T}F) \arrow[l, mapsto, "\sim"'] \\
\hat{\mathscr{C}}(F, S^{op}; \hat{T}F) \arrow[r, mapsto, "\sim"] & \hat{\mathscr{D}}(\hat{S}F, \hat{T}F)
\end{tikzcd}$$

Moreover, $Y: - \to \hat{-}$ *is a 2-natural transformation.*

Since monads are a 2-categorical concept (see Street, 1972), the 2-functor 2 maps monads in **Cat** to monads in **CAT**. Then, the statement of Theorem 4.9 about lifting of monads follows immediately from Proposition 4.11. It remains to define the lifting $\hat{\mathrm{t}}$ of a tensorial strength t for a monad (T, η, μ) over a small category $\mathscr{C}$.

PROPOSITION 4.12. *If $\mathscr{C}$ is a small category with finite products and T is an endofunctor over $\mathscr{C}$, then for every natural transformation $\mathrm{t}_{a,b}: a \times Tb \to T(a \times b)$ there exists a unique natural transformation $\hat{\mathrm{t}}_{F,G}: F \times \hat{T}G \to \hat{T}(F \times G)$ s.t. $\hat{\mathrm{t}}_{Ya,Yb} = Y(\mathrm{t}_{a,b})$ for all $a, b \in \mathscr{C}$.*

Proof. Every $F \in \hat{\mathscr{C}}$ is isomorphic to the colimit $\mathrm{Colim}^{\hat{\mathscr{C}}}_{Y\downarrow F}(\pi; Y)$ (shortly $\mathrm{Colim}_i\, Yi$), where Y is the Yoneda embedding of $\mathscr{C}$ into $\hat{\mathscr{C}}$. Similarly G is isomorphic to $\mathrm{Colim}_j\, Yj$. Both functors $(- \times \hat{T}-)$ and

$\hat{T} - \times -)$ from $\hat{\mathscr{C}} \times \hat{\mathscr{C}}$ to $\hat{\mathscr{C}}$ preserves colimits (as $\hat{T}$ and $- \times F$ are left adjoints) and commute with the Yoneda embedding (as $Y(a \times b) = Ya \times Yb$ and $\hat{T}(Ya) = Y(Ta)$). Therefore, $F \times \hat{T}G$ and $\hat{T}(F \times G)$ are isomorphic to the colimits $\mathrm{Colim}_{i,j}\, Yi \times \hat{T}(Yj)$ and $\mathrm{Colim}_{i,j}\, \hat{T}(Yi \times Yj)$, respectively. Let $\hat{\mathrm{t}}$ be the natural transformation we are looking for; then

$$\begin{tikzcd} Yi \times \hat{T}(Yj) \arrow[r, "Y(\mathrm{t}_{i,j})"] \arrow[d, "f \times \hat{T}g"'] & \hat{T}(Yi \times Yj) \arrow[d, "\hat{T}(f \times g)"] \\ F \times \hat{T}(G) \arrow[r, "\hat{\mathrm{t}}_{F,G}"'] & \hat{T}(F \times G) \end{tikzcd}$$

for all $f\colon Yi \to F$ and $g\colon Yj \to g$ (by naturality of $\hat{\mathrm{t}}$ and $\hat{\mathrm{t}}_{Yi,Yj} = Y(\mathrm{t}_{i,j})$). But there exists exactly one morphism $\hat{\mathrm{t}}_{F,G}$ making the diagram above commute, as $\langle \mathrm{t}_{i,j} \mid i, j \rangle$ is a morphism between diagrams in $\mathscr{C}$ of the same shape, and these diagrams have colimit cones $\langle f \times \hat{T}g \mid f, g \rangle$ and $\langle \hat{T}(f \times g) \mid f, g \rangle$, respectively. ∎

Remark 4.13. If T is a monad of partial computations, i.e., it is induced by a domination $\mathscr{M}$ on $\mathscr{C}$ s.t. $P(\mathscr{C}, \mathscr{M})(a, b) \cong \mathscr{C}(a, Tb)$, then the lifting $\hat{T}$ is the monad of partial computations induced by the dominion $\hat{\mathscr{M}}$ on $\hat{\mathscr{C}}$, obtained by lifting $\mathscr{M}$ to the topos of presheaves, as described in Rosolini (1988). For other monads, however, the lifting is not the *expected* one. For instance, if T is the monad of side-effects $(- \times S)^S$, then $\hat{T}$ is not (in general) the endofunctor $(- \times YS)^{YS}$ on the topos of presheaves.

Conclusions and Further Research

The main contribution of this paper is the category-theoretic semantics of computations and the general principle for extending it to more complex languages (see Remark 3.3 and Section 4), while the formal systems presented are a straightforward fallout, easy to understand and relate to other calculi.

Our work is just an example of what can be achieved in the study of programming languages by using a category-theoretic methodology, which avoids irrelevant syntactic detail and focus instead on the important structures underlying programming languages. We believe that there is a great potential to be exploited here. Indeed, in Moggi (1989b) we give a categorical account of phase distinction and program modules, that could lead to the introduction of higher order modules in programming languages like ADA or ML (see Harper, Mitchell, and Moggi, 1990), while in Moggi (1989a) we propose a "modular approach" to Denotational Semantics based on the idea of monad-constructor (i.e., an endofunctor on the category of monads over a category $\mathscr{C}$).

The metalanguage open also the possibility to develop a new Logic of Computable Functions (see Scott, 1969), based on an abstract semantics of computations rather than domain theory, for studying axiomatically different notions of computation and their relations. Some recent work by Crole and Pitts (1990) has considered an extension of the metalanguage equipped with a logic for inductive predicates, which goes beyond equational reasoning. A more ambitious goal would be to try exploiting the capabilities offered by higher-order logic in order to give a uniform account of various program logics, based on the idea of "T-modal operator" (see Definition 4.7).

The semantics of computations corroborates the view that (constructive) proofs and programs are rather unrelated, although both of them can be understood in terms of functions. Indeed, monads (and comonads) used to model logical modalities, e.g., possibility and necessity in modal logic or why not and of course of linear logic, usually do not have a tensorial strength. In general, one should expect types suggested by logic to provide a more fine-grained type system without changing the *nature* of computations.

We have identified monads as important to modeling notions of computations, but *computational monads* seem to have additional properties; e.g., they have a tensorial strength and may satisfy the mono requirement. It is likely that there are other properties of computational monads still to be identified, and there is no reason to believe that such properties have to be found in the literature on monads.

Acknowledgments

I have many people to thank for advice, suggestions and criticisms, in particular R. Amadio, R. Burstall, M. Felleisen, R. Harper, F. Honsell, M. Hyland, B. Jay, A. Kock, Y. Lafont, G. Longo, R. Milner, A. Pitts, G. Plotkin, J. Power, and C. Talcott.

Received February 21, 1990; final manuscript received October 23, 1990

References

Barr, M., and Wells, C. (1985), "Toposes, Triples and Theories," Springer-Verlag, Berlin/New York.

Crole, R. L., and Pitts, A. M. (1990), New foundations for fixpoint computations, *in* "4th LICS Conference," IEEE, New York.

Constable, R. L., and Smith, S. F. (1987), Partial objects in constructive type theory, *in* "2nd LICS Conference," IEEE, New York.

Constable, R. L., and Smith, S. F. (1987), Computational foundations of basic recursive function theory, *in* "3rd LICS Conference," IEEE, New York.

FELLEISEN, M., AND FRIEDMAN, D. P. (1989), A syntactic theory of sequential state, *Theoret. Comput. Sci.* **69**, No. 3.

FELLEISEN, M., FRIEDMAN, D. P., KOHLBECKER, E., AND DUBA, B. (1986), Reasoning with continuations, *in* "1st LICS, Conference," IEEE, New York.

FOURMAN, M. P. (1977), The logic of topoi, *in* "Handbook of Mathematical Logic" (J. Barwise, Ed.), Studies in Logic, Vol. 90, North-Holland, Amsterdam.

GORDON, M. J. C., MILNER, R., AND WADSWORTH, C. P. (1979), "Edinburgh LCF: A Mechanized Logic of Computation," Lecture Note in Computer Science, Vol. 78, Springer-Verlag, Berlin/New York.

GUNTER, C., AND SCOTT, S. (1989), "Semantic Domains," Technical Report MS-CIS-89-16, Dept. of Comp. and Inf. Science, Univ. of Pennsylvania. [To appear in North-Holland Handbook of Theoretical Computer Science].

HARPER, R., MITCHELL, J., AND MOGGI, E. (1990), Higher-order modules and the phase distinction, *in* "17th POPL.," *Assoc. Comput. Mach.*, New York.

HYLAND, J. M. E., AND PITTS, A. M. (1987), The theory of constructions: Categorical semantics and topos-theoretic models, *in* "Proc. AMS Conf. on Categories in Comp. Sci. and Logic (Boulder, 1987)."

JOHNSTONE, P. T. AND PARE, R. (Eds.) (1978), "Indexed Categories and their Applications," Lecture Notes in Mathematics, Vol. 661, Springer-Verlag, Berlin/New York.

KELLY, G. M. (1982), "Basic Concepts of Enriched Category Theory," Cambridge Univ. Press, London/New York.

KOCK, A. (1972), Strong functors and monoidal monads, *Arch. Math.* **23**.

KOCK, A., AND REYES, G. E. (1977), Doctrines in categorical logic, *in* "Handbook of Mathematical Logic" (J. Barwise, Ed.), Studies in Logic, Vol. 90, North-Holland, Amsterdam.

LAFONT, Y. (1988), The linear abstract machine, *Theoret. Comput. Sci.* **59**.

LAMBEK, J., AND SCOTT, P. J. (1986), "Introduction to Higher-Order Categorical Logic," Cambridge Studies in Advanced Mathematics, Vol. 7, Cambridge Univ. Press, London/New York.

MACLANE, S. (1971), "Categories for the Working Mathematician," Springer-Verlag, Berlin/New York.

MANES, E. (1976), "Algebraic Theories," Graduate Texts in Mathematics, Vol. 26, Springer-Verlag, Berlin/New York.

MASON, I. A. (1988), Verification of programs that destructively manipulate data, *Sci. Comput. Programming* **10**.

MOGGI, E. (1986), Categories of partial morphisms and the partial lambda-calculus, *in* "Proceedings Workshop on Category Theory and Computer Programming, Guildford 1985," Lecture Notes in Computer Science, Vol. 240, Springer-Verlag, Berlin/New York.

MOGGI, E. (1988), "The Partial Lambda-Calculus," Ph.D. thesis, University of Edinburgh.

MOGGI, E. (1989a), "An Abstract View of Programming Languages," Technical Report ECS-LFCS-90-113, Edinburgh Univ., Dept. of Comp. Sci., Lecture Notes for course CS 359, Stanford Univ.

MOGGI, E. (1989b), A category-theoretic account of program modules, *in* "Proceedings of the Conference on Category Theory and Computer Science, Manchester, UK, Sept. 1989," Lecture Notes in Computer Science, Vol. 389, Springer-Verlag, Berlin/New York.

MOGGI, E. (1989c), Computational lambda-calculus and monads, *in* "4th LICS Conference," IEEE, New York.

MOSSES, P. (1989), "Denotational Semantics," Technical Report MS-CIS-89-16, Dept. of Comp. and Inf. Science, University of Pennsylvania. [To appear in Handbook of Theoretical Computer Science, North-Holland, Amsterdam].

MAKKAI, M., AND REYES, G. (1977), "First Order Categorial Logic," Springer-Verlag, Berlin/New York.

MASON, I., AND TALCOTT, C. (1989a), Programming, transforming, and proving with function abstractions and memories, *in* "16th Colloquium on Automata, Languages and Programming," EATCS, Monographson Theoretical Computer Science, Springer-Verlag, Berlin/New York.

MASON, I., AND TALCOTT, C. (1989b), A sound and complete axiomatization of operational equivalence of programs with memory, *in* "POPL 89," Assoc. Comput. Mach., New York.

PLOTKIN, G. D. (1975), Call-by-name, call-by-value and the λ-calculus, *Theoret. Comput. Sci.* **1**.

PLOTKIN, G. D. (1985), "Denotational Semantics with Partial Functions," Lecture Notes at C.S.L.I. Summer School.

ROSOLINI, G. (1986), "Continuity and Effectiveness in Topoi," Ph.D. thesis, University of Oxford.

SCHMIDT, D. A. (1986), "Denotational Semantics: A Methodology for Language Development," Allyn & Bacon, Rockleigh, NO.

SCOTT, D. S. (1969), A type-theoretic alternative to CUCH, ISWIM, OWHY, Oxford Notes.

SCOTT, D. S. (1979), Identity and existence in intuitionistic logic, *in* "Applications of Sheaves" (M. P. Fourman, C. J. Mulvey, and D. S. Scott, Eds.), Lecture Notes in Mathematics, Vol. 753, Springer-Verlag, Berlin/New York.

SCOTT, D. S. (1980), Recating theories of the λ-calculus, *in* "To H. B. Curry: Essays in Combinatory Logic, lambda Calculus and Formalisms" (R. Hindley and J. Seldin, Eds.), Academic Press.

SEELY, R. A. G. (1987), Linear Logic, $*$-autonomous categories and cofree coalgebras, *in* "Proceedings, AMS Conference on Categories in Computer Science and Logic (Boulder 1987)."

SHARMA, K. (1984), Syntactic aspects of the non-deterministic lambda calculus, Master's thesis, Washington State University. [Available as Internal Report CS-84-127 of the Computer Science Department]

SMITH, M., AND PLOTKIN, G. (1982), The category-theoretic solution of recursive domain equations, *SIAM J. Comput.* **11**.

STREET, R. (1972), The formal theory of monads, *J. Pure Appl. Algebra* **2**.