Monads Need Not Be Endofunctors

Thorsten Altenkirch, University of Nottingham
James Chapman, Institute of Cybernetics, Tallinn
Tarmo Uustalu, Institute of Cybernetics, Tallinn

ScotCats, Edinburgh, 21 May 2010

Motivation

- Monads are the most successful pattern in functional programming and Type Theory.
- Useful for modelling effects (e.g. error, state, etc), but also other programming idioms (e.g. generalized syntactic structures).
- Monads, and constructions on monads (such as monad transformers) are key to reusable strutures.
- Frequently, we find structures that fail to be monads as if for the only reason that the underlying functor is not an endofunctor.
- E.g., untyped/typed lambda calculus syntax (over finite contexts), finite-dimensional vector spaces etc.
- Can/should one develop a theory of such structures?

Example: Vector spaces

- Let $\mathbb{F}$ be the skeletal category of finite sets ($|\mathbb{F}| = \mathbb{N}$).
- $J_{\mathrm{f}} \in \mathbb{F} \to \mathbf{Set}$ is the obvious embedding.
- Let $(R, +, 0, \times, 1)$ be a semiring.
- We define

$$
\begin{aligned}
&\mathsf{Vec} \in |\mathbb{F}| \to |\mathbf{Set}| \\
&\mathsf{Vec}\, m =_{\mathrm{df}} J_{\mathrm{f}}\, m \to R \\
&\eta_m \in J_{\mathrm{f}}\, m \to \mathsf{Vec}\, m \\
&\eta_m\, i =_{\mathrm{df}} \lambda j.\ \mathrm{if}\ i = j\ \mathrm{then}\ 1\ \mathrm{else}\ 0 \\
&(-)^* \in (J_{\mathrm{f}}\, m \to \mathsf{Vec}\, n) \to (\mathsf{Vec}\, m \to \mathsf{Vec}\, n) \\
&A^*\, \vec{a} =_{\mathrm{df}} \lambda j. \sum_{i \in \underline{m}} \vec{a}\, i \times A\, i\, j
\end{aligned}
$$

- Check that:

$$
\begin{aligned}
k^* \circ \eta_X &= k \\
\eta_X^* &= \mathrm{id}_{\mathsf{Vec}\, X} \\
(l^* \circ k)^* &= l^* \circ k^*
\end{aligned}
$$

Relative monads

- Given a category $\mathbb{C}$ and another category $\mathbb{J}$ with a functor $J \in [\mathbb{J}, \mathbb{C}]$.
- A *relative monad* is given by
  - an object function $T \in |\mathbb{J}| \to |\mathbb{C}|$,
  - for any object $X \in |\mathbb{J}|$, a map $\eta_X \in \mathbb{C}(J\,X, T\,X)$ (unit),
  - for any objects $X, Y \in |\mathbb{J}|$ and map $k \in \mathbb{C}(J\,X, T\,Y)$, a map $k^* \in \mathbb{C}(T\,X, T\,Y)$ (Kleisli extension)

  satisfying

  - for any $X, Y \in |\mathbb{J}|$, $k \in \mathbb{C}(J\,X, T\,Y)$, $k^* \circ \eta_X = k$,
  - for any $X \in |\mathbb{J}|$, $\eta_X^* = \mathrm{id}_{T\,X} \in \mathbb{C}(T\,X, T\,X)$,
  - for any $X, Y, Z \in |\mathbb{J}|$, $k \in \mathbb{C}(J\,X, T\,Y)$, $\ell \in \mathbb{C}(J\,Y, T\,Z)$, $(\ell^* \circ k)^* = \ell^* \circ k^* \in \mathbb{C}(T\,X, T\,Z)$.
- T is functorial with $T\,f = (\eta \circ J\,f)^*$; η and $(-)^*$ are natural.

Relative monads (ctd)

- Clearly $T = \mathrm{Vec}$ with $\mathbb{J} = \mathbb{F}$ and $J = J_{\mathrm{f}}$ is an instance.
- Ordinary monads arise as as the special case where $\mathbb{J} =_{\mathrm{df}} \mathbb{C}$, $J =_{\mathrm{df}} \mathsf{Id}_{\mathbb{C}}$.
- Any monad $(T, \eta, (-)^*)$ on $\mathbb{C}$ restricts to a relative monad $(T^\flat, \eta^\flat, (-)^{*\flat})$ on J defined by $T^\flat X =_{\mathrm{df}} T\,(J\,X)$, $\eta^\flat_X =_{\mathrm{df}} \eta_{J\,X}$, $k^{*\flat} =_{\mathrm{df}} k^*$.

Example: Untyped lambda calculus syntax

- Define T as the initial algebra of $F \in [\mathbb{F}, \mathbf{Set}] \to [\mathbb{F}, \mathbf{Set}]$ defined by $F\,G\,X =_{\mathrm{df}} J\,X + (G\,X \times G\,X + G\,(1 + X))$ (the terms of untyped lambda calculus).
- T is a relative monad, with η the inclusion of variables to terms and $(-)^*$ substitution.

Example: Typed lambda calculus syntax

- Let Ty be the set of types of typed lambda calculus (over some base types).
- Let $\mathbb{F} \downarrow \mathrm{Ty}$ be the category whose objects are pairs (Γ, ρ) where $\Gamma \in |\mathbb{F}|$ and $\rho \in \Gamma \to \mathrm{Ty}$ and maps from (Γ, ρ) to (Γ', ρ') are maps $f \in \mathbb{F}(\Gamma, \Gamma')$ such that $\rho = \rho' \circ f$ (the contexts and context maps).
- Let $J \in \mathbb{F} \downarrow \mathrm{Ty} \to [\mathrm{Ty}, \mathbf{Set}]$ be the natural embedding.
- T (the terms) can be defined as an initial algebra of a suitable endofunctor on $[\mathbb{F} \downarrow \mathrm{Ty}, [\mathrm{Ty}, \mathbf{Set}]]$.
- T is a relative monad.

Example: Indexed Functors

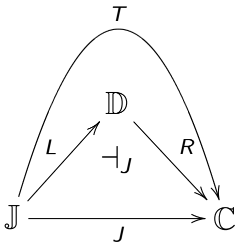
- Let $\mathbf{U}$ be the category of small sets.
- The functor $J_{\mathbf{U}} \in [\mathbf{U}, \mathbf{Cat}]$ views a small set as a category.
- $\mathrm{IF} \in [\mathbf{U}, \mathbf{Cat}]$ defined by $\mathrm{IF}\, A =_{\mathrm{df}} [[J_{\mathbf{U}}\, A, \mathbf{U}], \mathbf{U}]$ gives rise to a relative monad.
- The definitions of η and $(-)^*$ correspond to the continuation monad (apart from the size issue).
- This showed up in our work on *indexed containers* (LICS 09), which also form a relative monad.

Relative adjunctions

- Given two categories $\mathbb{C}$, $\mathbb{D}$ together with a third category J and a functor $J \in \mathbb{J} \to \mathbb{C}$.
- Given $L \in [\mathbb{C}, \mathbb{D}]$, $R \in [\mathbb{D}, \mathbb{C}]$: $L \dashv_J R$ (L is a relative left adjoint to R), if

$$\mathbb{C}(J\,X, R\,Y) \simeq \mathbb{D}(L\,X, Y)$$

- A relative adjunction gives rise to a relative monad $T = R \cdot L$.



Kleisli and Eilenberg-Moore constructions

- Given a relative monad we can define its initial ($\mathbf{Kl}(T)$) and terminal ($\mathbf{EM}(T)$) splitting as a relative adjunction.
- $|\mathbf{Kl}(T)| = |\mathbb{J}|$ and $\mathbf{Kl}(T)(X, Y) =_{\mathrm{df}} \mathbb{C}(J\,X, T\,Y)$.
- Kleisli categories for the examples:

  Vector spaces Finite dimensional vector spaces

  λ calculus (untyped/typed) contexts and substitutions.

  Indexed Functors Functors between different slices.
- To define $\mathbf{EM}(T)$ we define the notion of an EM-algebra without referring to μ.
- An EM-algebra is given by family of maps

$$a_X \in \mathbb{C}(J\,X, A) \to \mathbb{C}(T\,X, A)$$

such that $a\,\rho \circ \eta = \rho$ and $a(a\,\rho \circ k) = a\,\rho \circ k^*$

Relative Monads as monoids?

- Can we have a monoid form of relative monads?
- Here is a calculation in the end-coend calculus:

$$
\begin{aligned}
&\int_{X,Y\in|\mathbb{J}|} \mathbb{C}(JX, TY) \to \mathbb{C}(TX, TY) \\
\cong\ & \int_{Y\in|\mathbb{J}|} \mathbb{C}\Big(\int^{X\in|\mathbb{J}|} \mathbb{C}(JX, TY) \bullet TX, TY\Big) \\
\cong\ & \int_{Y\in|\mathbb{J}|} \mathbb{C}(\mathrm{Lan}_J\, T\, (TY), TY) \\
\cong\ & [\mathbb{J}, \mathbb{C}](\mathrm{Lan}_J\, T \cdot T, T)
\end{aligned}
$$

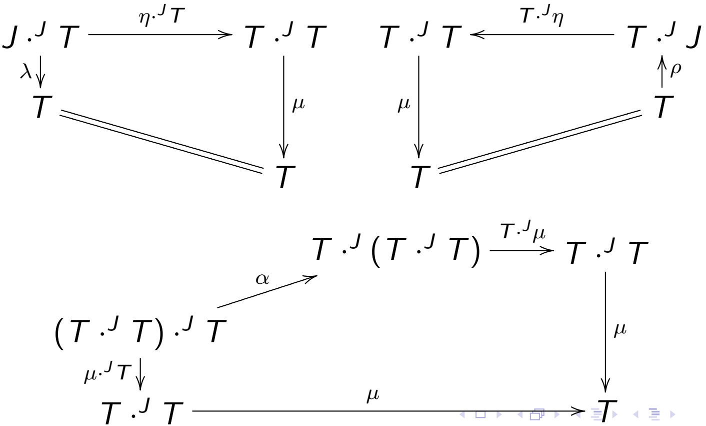
- Assume henceforth that $\mathrm{Lan}_J \in [\mathbb{J}, \mathbb{C}] \to [\mathbb{C}, \mathbb{C}]$ exists.

$[\mathbb{J}, \mathbb{C}]$ is lax monoidal

- We can define
  - for any objects $F, G \in |[\mathbb{J}, \mathbb{C}]|$, an object $G \cdot^J F \in |[\mathbb{J}, \mathbb{C}]|$ by $G \cdot^J F =_{\mathrm{df}} \mathrm{Lan}_J\, G \cdot F$.
- We can also define
  - for any object $F \in |[\mathbb{J}, \mathbb{C}]|$, a map $\lambda_F \in [\mathbb{J}, \mathbb{C}](\mathrm{Lan}_J\, J \cdot F, F)$,
  - for any object $F \in |[\mathbb{J}, \mathbb{C}]|$, a map $\rho_F \in [\mathbb{J}, \mathbb{C}](F, \mathrm{Lan}_J\, F \cdot J)$,
  - for any objects $F, G, H \in |[\mathbb{J}, \mathbb{C}]|$, a map $\alpha_{H,G,F} \in [\mathbb{J}, \mathbb{C}](\mathrm{Lan}_J\, (\mathrm{Lan}_J\, H \cdot G) \cdot F, \mathrm{Lan}_J\, H \cdot \mathrm{Lan}_J\, G \cdot F)$.
- $([\mathbb{J}, \mathbb{C}], J, \cdot^J, \lambda, \rho, \alpha)$ is a lax monoidal category, i.e., $\cdot^J$ is functorial, λ, ρ, α are natural (however not generally isomorphisms) and satisfy certain coherence conditions.

Relative monads = lax monoids

- Relative monads on J are the same as lax monoids in the lax monoidal structure on $[\mathbb{J}, \mathbb{C}]$,
  i.e., triples (T, η, μ) with $T \in |[\mathbb{J}, \mathbb{C}]|$, $\eta \in [\mathbb{J}, \mathbb{C}](J, T)$ and $\mu \in [\mathbb{J}, \mathbb{C}](T \cdot^J T, T)$ such that



Assume further conditions on J...

- Assume that, in addition to the existence of Lan_J, J further satisfies these conditions:
  - J is fully faithful, i.e., for any $X, Y \in |\mathbb{J}|$, there is an inverse to the canonical map $J_{X,Y} \in \mathbb{J}(X, Y) \to \mathbb{C}(J\,X, J\,Y)$ given by $J_{X,Y} f =_{\mathrm{df}} J\,f$,
  - J is dense, i.e., for any $X, Y \in |\mathbb{C}|$, there is an inverse to the canonical map $K_{X,Y} \in \mathbb{C}(X, Y) \to [\mathbb{J}^{\mathrm{op}}, \mathbf{Set}](\mathbb{C}(J\,-, X), \mathbb{C}(J\,-, Y))$ given by $K_{X,Y}\,g\,f =_{\mathrm{df}} g \circ f$,
  - For any $F \in \mathbb{J} \to \mathbb{C}$, $X \in |\mathbb{J}|$, $Y \in |\mathbb{C}|$, there is an inverse to the canonical map $L^F_{X,Y} \in \mathrm{Lan}_J(\mathbb{C}(J\,X, F\,-))\,Y \to \mathbb{C}(J\,X, \mathrm{Lan}_J\,F\,Y)$.
- The functors $J \in \mathbb{F} \to \mathbf{Set}$ and $J \in \mathbb{F} \downarrow \mathrm{Ty} \to [\mathrm{Ty}, \mathbf{Set}]$ enjoy these properties.

$[\mathbb{J}, \mathbb{C}]$ is monoidal, relative monads = monoids

- Then ρ, λ, α have inverses definable in terms of J^{-1}, K^{-1}, L^{-1}.
- Hence $[\mathbb{J}, \mathbb{C}]$ is (properly) monoidal.
- A relative monad T on J is a (proper) monoid in $[\mathbb{J}, \mathbb{C}]$.

Relative monads extend to monads

- We also get that T extends to a monad on $\mathbb{C}$ (a monoid in the strict monoidal category $([\mathbb{C},\mathbb{C}],\mathsf{Id},\cdot)$).
- Define
  - $T^\sharp =_{\mathrm{df}} \mathrm{Lan}_J\, T$,
  - $\eta^\sharp =_{\mathrm{df}}\ \mathsf{Id} \xrightarrow{\lambda^{-1}_{\mathsf{Id}}} \mathrm{Lan}_J\, J \xrightarrow{\mathrm{Lan}_J\,\eta} \mathrm{Lan}_J\, T$,
  - $\mu^\sharp =_{\mathrm{df}}$

    $$\mathrm{Lan}_J\, T \cdot \mathrm{Lan}_J\, T \xrightarrow{\alpha^{-1}_{T,T,\mathsf{Id}}} \mathrm{Lan}_J\,(Lan_J\, T \cdot T) \xrightarrow{\mathrm{Lan}_J\,\mu} \mathrm{Lan}_J\, T$$
- $(T^\sharp,\eta^\sharp,\mu^\sharp)$ is a monad on $\mathbb{C}$.
- E.g., untyped lambda calculus syntax extends to a monad on **Set**, typed lambda calculus syntax to a monad on $[\mathrm{Ty},\mathbf{Set}]$.

Relative monads extend to monads (ctd)

- Furthermore, the defining adjunction of Lan_J,

$$[\mathbb{C}, \mathbb{C}] \underset{\mathrm{Lan}_J}{\overset{- \cdot J}{\rightleftarrows}} [\mathbb{J}, \mathbb{C}] \qquad (\top)$$

  lifts to an adjunction

$$\mathbf{Mnd}(\mathbb{C}) \underset{(-)^{\sharp}}{\overset{(-)^{\flat}}{\rightleftarrows}} \mathbf{RMnd}(J) \qquad (\top)$$

Summary

- No conditions on J
  - Monads restrict to relative monads
  - Huber's theorem, Kleisli, E-M constructions
- Lan_J exists
  - $[\mathbb{J}, \mathbb{C}]$ lax monoidal, relative monads $=$ lax monoids
- Further conditions on J
  - $[\mathbb{J}, \mathbb{C}]$ monoidal, relative monads $=$ monoids
  - Relative monads extend to monads, coreflection

Arrows

- Given a category $\mathbb{J}$, a (weak) *arrow* on $\mathbb{J}$ is given by
  - an object function $R \in |\mathbb{J}| \times |\mathbb{J}| \to \mathbf{Set}$,
  - for any objects $X, Y \in |\mathbb{J}|$, a function $\mathsf{pure} \in \mathbb{J}(X, Y) \to R(X, Y)$,
  - for any $X, Y, Z \in |\mathbb{J}|$, a function $(\lll) \in R(Y, Z) \times R(X, Y) \to R(X, Z)$

  satisfying
  - $\mathsf{pure}\,(g \circ f) = \mathsf{pure}\, g \lll \mathsf{pure}\, f$,
  - $r \lll \mathsf{pure}\,\mathrm{id} = r$,
  - $\mathsf{pure}\,\mathrm{id} \lll r = r$,
  - $t \lll (s \lll r) = (t \lll s) \lll r$.
- R extends to a functor $\mathbb{J}^{\mathrm{op}} \times \mathbb{J} \to \mathbf{Set}$ (an endoprofunctor on $\mathbb{J}$); pure and $\lll$ are natural.

Arrows = relative monads on Yoneda

- Assume $\mathbb{J}$ is small. Let $\mathbb{C} =_{\mathrm{df}} [\mathbb{J}^{\mathrm{op}}, \mathbf{Set}]$, $J\,Y\,X = \mathbb{J}(X, Y)$ (the Yoneda embedding).
- Lan_J exists, J is well-behaved.
- An arrow on $\mathbb{J}$ (a functor $R \in \mathbb{J}^{\mathrm{op}} \times \mathbb{J} \to \mathbf{Set}$ with structure) is the same as a relative monad on J (a functor $T \in \mathbb{J} \to [\mathbb{J}^{\mathrm{op}}, \mathbf{Set}]$ with structure).
- Cf. Jacobs et al. (2006): Arrows on $\mathbb{J}$ are the same as monoids in the monoidal structure on $[\mathbb{J}^{\mathrm{op}} \times \mathbb{J}, \mathbf{Set}]$ (the category of endoprofunctors on $\mathbb{J}$).

Conclusions

- Relative monads are a natural generalization of monads.
- They are smoothly formulated in Manes's style, the monoid form needs left Kan extensions.
- A large part of monad theory carries over with minimal adjustments. There is a clear relationship to ordinary monads.
- They cover important examples for programming, in particular, examples with size issues.
- They also subsume arrows.