Lecture: Monad and Adjunctions

November 7, 2018

Moral 1 Monads are "shadows" of adjunctions!

Moral 2 If (T, μ, η) is a monad on $\mathcal{C}$, then Ta is a generalized space associated with a; μ_a is a degeneralization morphism; η_a is a canonical embedding of a into the generalized space Ta.

Moral 3 Morphisms in the Kleisli category $\mathcal{C}_T$ are generalizations of morphisms in the base category $\mathcal{C}$.

More precisely, every adjunction gives rise to a monad, and every monad gives rise to (multiple) adjunctions.

Definition 1. Let $\mathcal{C}$ be a category. A **monad** on $\mathcal{C}$ is a triple (T, μ, η) where $T : \mathcal{C} \to \mathcal{C}$ is an endofunctor, $\mu : T^2 \Rightarrow T$ and $\eta : id_{\mathcal{C}} \Rightarrow T$ are natural transformations such that the following two diagrams commute:

$$\begin{tikzcd}
T^3 \arrow[r, "T\mu"] \arrow[d, "\mu_T"'] & T^2 \arrow[d, "\mu"] \\
T^2 \arrow[r, "\mu"] & T
\end{tikzcd}
\qquad
\begin{tikzcd}
T \arrow[r, "\eta_T"] \arrow[dr, equal] & T^2 \arrow[d, "\mu"] & T \arrow[l, "T\eta"'] \arrow[dl, equal] \\
& T &
\end{tikzcd}$$

Theorem 1 (Adjunction $\Rightarrow$ Monad)**.** *Given*

$$\begin{tikzcd}
\mathcal{C} \arrow[r, bend left, "F"] \arrow[r, phantom, "\perp"] & \mathcal{D} \arrow[l, bend left, "G"]
\end{tikzcd}$$

$(GF, G\epsilon F, \eta)$ is a monad on $\mathcal{C}$.

Proof. The associative diagram for $G\epsilon F$ is a special case of horizontal composition of natural transformations:

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "F"] & \mathcal{D} \arrow[r, "FG"] \arrow[r, bend right, "id_{\mathcal{D}}"'] & \mathcal{D} \arrow[r, "FG"] \arrow[r, bend right, "id_{\mathcal{D}}"'] & \mathcal{D} \arrow[r, "G"] & \mathcal{C}
\end{tikzcd}$$

The unit diagram for η follows from triangle equalities:

$$\begin{tikzcd}
GFc \arrow[r, "GF\eta_c"] \arrow[dr, equal] & GFGFc \arrow[d, "G\epsilon_{Fc}"] & GFc \arrow[l, "\eta_{GFc}"'] \arrow[dl, equal] \\
& GFc &
\end{tikzcd}$$

where

$$G\epsilon_{Fc} \cdot \eta_{GFc} = (G\epsilon \cdot \eta_G)_{Fc} = (id_G)_{Fc} = id_{GFc}$$

and

$$G\epsilon_{Fc} \cdot GF\eta_c = G(\epsilon_{Fc} \cdot F\eta_c) = G(id_F)_c = id_{GFc}$$

$\square$

Definition 2. Let (T, μ, η) be a monad acting on $\mathcal{C}$. The **Kleisli category** $\mathcal{C}_T$ is the category with $ob\mathcal{C}_T = ob\mathcal{C}$ and a morphism $f : a \rightsquigarrow b$ in $\mathcal{C}$ is a morphism $f : a \to Tb$ in $\mathcal{C}$. Compositions are given by

$$\begin{tikzcd} a \arrow[r, squiggly, "f"] & b \arrow[r, squiggly, "g"] & c \end{tikzcd}$$

$$\begin{tikzcd} a \arrow[r, "f"] & Tb \arrow[r, "Tg"] & T^2c \arrow[r, "\mu_c"] & Tc \end{tikzcd}$$

The **Eilenberg-Moore category** C^T (also known as the category of free T-algebras) is the category with objects (a, h), where $h : Ta \to a$ is a structure morphism that makes the following two diagrams commute,

$$\begin{tikzcd} a \arrow[r, "\eta_a"] \arrow[dr, equal] & Ta \arrow[d, "h"] \\ & a \end{tikzcd} \qquad \begin{tikzcd} T^a \arrow[r, "Th"] \arrow[d, "\mu_a"] & Ta \arrow[d, "h"] \\ Ta \arrow[r, "h"] & a \end{tikzcd}$$

and morphisms $f : (a, h) \to (b, k)$ that makes the following diagram commutes in $\mathcal{C}$:

$$\begin{tikzcd} Ta \arrow[r, "Tf"] \arrow[d, "h"] & Tb \arrow[d, "k"] \\ a \arrow[r, "f"] & b \end{tikzcd}$$

Theorem 2 (Monad $\Rightarrow$ Adjunction)**.** *Let (T, μ, η) be a monad acting on $\mathcal{C}$. Then there are adjunctions F^T, G^T, F_T, G_T such that the following diagram commutes*

$$\begin{tikzcd} \mathcal{C}_T \arrow[dr, bend right, "G_T"'] & & \mathcal{C}^T \arrow[dl, bend right, "G^T"'] \\ & \mathcal{C} \arrow[ul, bend right, "F_T"'] \arrow[ur, bend right, "F^T"'] & \end{tikzcd}$$

Proof. We define F_T, G_T as follows:

$$\begin{tikzcd} \mathcal{C}_T & a \arrow[r, squiggly, "f"] \arrow[d, "G_T"] & b \arrow[d, "G_T"] & a \arrow[r, squiggly, "\eta_b \cdot g"] & b \\ \mathcal{C} & Ta \arrow[r, "\mu_b \cdot Tf"] & Tb & a \arrow[u, "F_T"] \arrow[r, "g"] & b \arrow[u, "F_T"'] \end{tikzcd}$$

Check:

$$(G_T(F_T f)) = G_T(\eta_b \cdot f) = \mu_b \cdot T(\eta_b \cdot f) = (\mu_b \cdot T\eta_b) \cdot Tf = Tf$$

Moreover, there is a natural bijection between $\mathcal{C}_T(F_T a, b) = \mathcal{C}_T(a, b)$ and $\mathcal{C}(a, G_T b) = \mathcal{C}(a, Tb)$. In fact, in this case, the bijection is an "equality" given how the morphisms in $\mathcal{C}_T$ are defined.

In particular, we note that the counit ϵ_T at b, which is a morphism from $F_T G_T b = Tb$ to b in $\mathcal{C}_T$, just is 1_{Tb}.

Now as for the Eilenberg-Moore category, we define F^T and G^T as follows,

$$\begin{tikzcd} \mathcal{C} & a \arrow[r, "f"] \arrow[d, "F^T"] & b \arrow[d, "F^T"] & a \arrow[r, "g"] & b \\ \mathcal{C}^T & (Ta, \mu_a) \arrow[r, "f"] & (Tb, \mu_b) & (a, h) \arrow[u, "G^T"] \arrow[r, "g"] & (b, k) \arrow[u, "G^T"'] \end{tikzcd}$$

Let $\epsilon_{(a,h)} = h$. We leave it as an exercise to check that ϵ so defined is indeed a natural transformation from $F^T G^T$ to $id_{\mathcal{C}^T}$. We check that $F^T \dashv G^T$ by showing that η and ϵ satisfy the triangle equalities.

$$\begin{tikzcd}
F^T a = (Ta, \mu_a) \arrow[r, "F^T \eta"] \arrow[dr, equal] & F^T G^T F^T a = (T^2 a, \mu_{Ta}) \arrow[d, "\epsilon F^T = \mu_{Ta}"] \\
& F^T a = (Ta, \mu_a)
\end{tikzcd}$$

which commutes by the associative square of μ.

Similarly,

$$\begin{tikzcd}
G^T(a,h) = a \arrow[r, "\eta G^T = \eta"] \arrow[dr, equal] & G^T F^T G^T(a,h) = Ta \arrow[d, "G^T \epsilon_{(a,h)} = h"] \\
& G^T(a,h) = a
\end{tikzcd}$$

which commutes by the fact that h is the left inverse of η_a. □

Lastly, if $\mathcal{C}$ is a category, let $\mathrm{Adj}_T(\mathcal{C})$ be the category of adjunctions whose objects are fully specified adjunctions and morphisms are adjunctions transformations. We show that $\mathcal{C}_T$ and $\mathcal{C}^T$ are respectively the initial and terminal objects in this category, in the following sense:

Theorem 3. *Let (T, μ, η) be a monad acting on $\mathcal{C}$. If $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ is an adjunction that gives rise to this monad, then there is a unique functor $K : \mathcal{C}_T \to \mathcal{D}$ and $L : \mathcal{D} \to \mathcal{C}^T$ such that the following diagram commutes:*

$$\begin{tikzcd}
\mathcal{C}_T \arrow[r, dotted, "K"] \arrow[dr, shift left, "F_T"'] & \mathcal{D} \arrow[r, dotted, "L"] \arrow[d, shift right, "F"'] & \mathcal{C}^T \arrow[dl, shift left, "F^T"] \\
& \mathcal{C} \arrow[ul, shift left, "G_T"] \arrow[u, shift right, "G"'] \arrow[ur, shift left, "G^T"] &
\end{tikzcd}$$

Proof. The left triangle commutes just in case the following squares commute:

$$\begin{tikzcd}
\mathcal{C}_T \arrow[r, "G_T"] \arrow[d, "K"] & \mathcal{C} \arrow[r, "F_T"] \arrow[d, equal] & \mathcal{C}_T \arrow[d, "K"] \\
\mathcal{D} \arrow[r, "G"] & \mathcal{C} \arrow[r, "F"] & \mathcal{D}
\end{tikzcd}$$

Now the commutativity of the right square forces

$$KF_T a = Ka = Fa, \quad \forall a : ob\mathcal{C}.$$

Moreover, note that

$$\mathcal{C}_T(a,b) = \mathcal{C}(a, Tb) = \mathcal{C}(a, G_T F_T b) = \mathcal{C}(a, GKF_T b) \cong \mathcal{D}(Fa, KF_T b) = \mathcal{D}(KF_T a, KF_T b).$$

Recall that if $F \dashv G$, then

$$\mathcal{C} \qquad a \xrightarrow{f} GKF_T b$$

$$\mathcal{D} \qquad Fa \xrightarrow{Ff} FGKF_T b \xrightarrow{\epsilon_{KF_T b}} KF_T b$$

But $\epsilon_{KF_T b} = \epsilon_{Fb}$. So $Kf = \epsilon_{Fb} \cdot Ff$. □

Hence K exists and is unique.

Now the right triangle commutes just in case the following squares commute:

$$\begin{tikzcd}
\mathcal{C}^T \arrow[r, "G^T"] & \mathcal{C} \arrow[r, "F^T"] & \mathcal{C}^T \\
\mathcal{D} \arrow[r, "G"] \arrow[u, "L"] & \mathcal{C} \arrow[r, "F"] \arrow[u, equal] & \mathcal{D} \arrow[u, "L"]
\end{tikzcd}$$

Since G_T is the projection functor, the commutativity of the left square forces

$$Ld = (Gd, \gamma_d), \quad \forall d : ob\mathcal{D}$$

We verify that $\gamma_d = G\epsilon_d$. By previous observation, we know $\gamma_d = \epsilon^T_{(Gd,\gamma d)} = \epsilon^T_{Ld}$. By the triangle inequality $G\epsilon \cdot \eta_G$, we have

$$\begin{tikzcd}
Gd = G^T Ld \arrow[r, "\eta_{Gd}"] \arrow[dr, equal] & G^T F^T G^T Ld = GFGd \arrow[d, "G^T \epsilon^T_{Ld} = \gamma_d"] \arrow[d, bend left, "G\epsilon_d"] \\
& G^T Ld = Gd
\end{tikzcd}$$

Therefore, $\gamma_d = G\epsilon_d$. On morphisms, since $LG^T = G$, we have $Lf = Gf$, which one can verify is indeed a a structure morphism.

Example 1 (Powerset Monad)**.** Consider the powerset monad $(\mathcal{P}, \mu, \eta)$ on **Set** defined by

$$\begin{aligned}
\mu_X &: \mathcal{P}\mathcal{P}X \to \mathcal{P}X \quad S \mapsto \bigcup S \\
\eta_X &: X \to \mathcal{P}X \quad x \mapsto \{x\}
\end{aligned}$$

This gives content to our Moral 2: $\mathcal{P}X$ is a "generalized set" associated with X; μ_X de-generalizes $\mathcal{P}\mathcal{P}X$ to $\mathcal{P}X$ by collapsing the set-brackets; η_X embeds X into $\mathcal{P}X$.

The crucial observation is that $\mathbf{Set}_{\mathcal{P}}$, the Kleisli category of the powerset monad, is the category of sets with relations!

Here is a short argument why:

$$\begin{aligned}
\mathbf{Set}^{\mathcal{P}}(X, Y) &= \mathbf{Set}(X, \mathcal{P}Y) \\
&= \mathbf{Set}(X, \mathbf{Set}(Y, 2)) \\
&= \mathbf{Set}(X \times Y, 2) \\
&= \mathcal{P}(X \times Y)
\end{aligned}$$

So morphisms in $\mathbf{Set}^{\mathcal{P}}$ are precisely set-theoretic relations!

Example 2 (Giry Monad)**.** Recall that the Giry monad is the triple $(\mathcal{M}, \mu, \eta)$ on **Meas** where

- $\mathcal{M}$ sends each measurable space $(X, \mathcal{F})$ to the measurable space of its probability measures, i.e. $\mathcal{M}(X) = \{\nu : \nu \text{ is a probability measure on } X\}$.

> (Crash course on measure theory for those who are keen) A measurable space is a set X endowed with a σ-algebra $\mathcal{F}$, where $\mathcal{F}$ is a field of subsets of X (subsets that are "measurable") such that $\mathcal{F}$ contains $\emptyset$ and is closed under complementation as well as countable union.
>
> Given two measurable space $(X, \mathcal{F})$ and $(Y, \mathcal{G})$, we say a function $f : X \to Y$ is measurable if $f^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{G}$. That is, the preimage of a measurable set under f is also measurable (but f does not need to send a measurable set to a measurable set; in particular, $f(X)$ does not need to be measurable).
>
> Given a measurable space $(X, \mathcal{F})$, the canonical σ-algebra associated with $\mathcal{M}(X)$ is the coarsest σ-algebra such that the functions $ev_A : \mathcal{M}(X) \to [0, 1]$ given by $ev_A(\nu) = \nu(A)$ is measurable (with respect to the Borel algebra of $[0, 1]$ - the σ-algebra generated by open intervals in $[0, 1]$) for all $A \in \mathcal{F}$.

- $\mu_X : \mathcal{M}(\mathcal{M}(X)) \to \mathcal{M}(X)$ is given by $\pi \mapsto \mathcal{E}(\pi)$, where $\mathcal{E}(\pi)$ is a probability measure on X that assigns each measurable set A the probability $\int_{\nu \in \mathcal{M}(X)} \nu(A)\pi(d\nu)$ - the average probability of A weighted by π.

- $\nu_X : X \to \mathcal{M}(X)$ sends each element $x \in X$ to the Dirac measure δ_x.

Again our Moral 2 and Moral 3 apply in this example: we can think of $\mathcal{M}(X)$ as a (randomized) generalized space associated with X; μ_X as de-randomization, and ν_X as a canonical embedding of X into its stochastic generalization (where elements in X correspond to deterministic measures).

The Kelisli category of this monad is rather beautiful. Note that morphisms in $\mathbf{Meas}_{\mathcal{M}}$ are measurable functions $Q : X \to \mathcal{M}(Y)$. Alternatively (by currying), we can think of Q as as a two-place function that takes (x, A) as input (with $x \in X$, and A a measurable set in Y), and returns a probability. This is precisely the *conditional probability* function (just rewrite $Q(x, A)$ as $Q(A|x)$). So morphisms in $\mathbf{Meas}_{\mathcal{M}}$ are also called Markov kernels or transition probabilities, and $\mathbf{Meas}_{\mathcal{M}}$ is often referred to as the stochastic category **Stoch**.

In particular, it is not hard to see that the terminal object in $\mathbf{Meas}_{\mathcal{M}}$ is $(1, \{\emptyset, 1\})$. Morphisms $P : 1 \rightsquigarrow X$ in $\mathbf{Meas}_{\mathcal{M}}$ are, according to the proposed interpretation, conditional probabilities $P(.|1 = 1)$. But since there is no stochasticity associated with $1 = 1$, these are indeed the *unconditional probability measures defined on* X! So this gives us an alternative perspective on probability theory: probability measures on $(X, \mathcal{F})$ are morphisms from 1 to X in $\mathbf{Meas}_{\mathcal{M}}$. Moreover, just as we can think of set-theoretic elements as injective functions $f : 1 \to X$ in **Set**, we can also think of probability functions $P : 1 \rightsquigarrow X$ as stochastic elements in X!