Monads

Monads occur whenever there are adjoint functors. Thus, they are an important construction in category theory.

Concepts

Recall the definition of adjoint functors.

- There are categories $\mathcal{C}$ and $\mathcal{D}$
- There are functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$. The functor F is the left adjoint of the functor G, which is the the right adjoint of F.
- There is a natural transformation $u : 1_{\mathcal{C}} \to GF$, called the *unit* of the adjunction.
- There is a natural transformation $v : FG \to 1_{\mathcal{D}}$, called the *counit* of the adjunction
- We have $1_F = v_F \circ F(u)$ and $1_G = G(v) \circ u_G$.
- The maps $g \mapsto g^u = G(g) \circ u_A$ and its inverse $f \mapsto f_v = v_B \circ F(f)$ provide a natural identification between $\mathcal{D}(A, G(B))$ and $\mathcal{C}(G(A), B)$.

In many of the examples of adjoint functors, we saw that GF is a functor from **Set** to itself which was of a special kind, so we and look at its properties in the light of what is given above.

- $\mathcal{C}$ is a category
- $GF : \mathcal{C} \to \mathcal{C}$ is a functor.
- $u : 1_{\mathcal{C}} \to GF$ is a natural transofrmation
- We have a natural transformation $G(v_F) : GFGF \to GF$ such that the following diagram commutes

$$\begin{array}{ccc} GF(A) & \xrightarrow{GF(u_A)} & GFGF(A) \\ & {\scriptstyle 1_{GF(A)}} \searrow\!\!\searrow & \downarrow {\scriptstyle G(v_{F(A)})} \\ & & GF(A) \end{array}$$

- We also have other commutative diagrams. For example, the following diagram follows from the identity $1_{G(B)} = G(v_B) \circ u_{G(B)}$, by substituting $B = F(A)$.

$$\begin{array}{ccc} GF(A) & \xrightarrow{u_{GF(A)}} & GFGF(A) \\ & {\scriptstyle 1_{GF(A)}} \searrow\!\!\searrow & \downarrow {\scriptstyle G(v_{F(A)})} \\ & & GF(A) \end{array}$$

This suggests the following definition of a monad M.

- We have a category $\mathcal{C}$.
- We have a functor M from $\mathcal{C}$ to itself.

- We have a natural transformation $u : 1_{\mathcal{C}} \to M$. This is called the "identity" or "unit" of the monad.
- We have a natural transformation $m : MM \to M$. This is called the "multiplication" or "operation" of the monad.
- We have the commutative diagram (which is like saying that u is the identity for the multiplication m)

$$\begin{array}{ccc} M(A) & \xrightarrow{u_{M(A)}} & MM(A) \\ {\scriptstyle M(u_A)}\big\downarrow & \overset{1_{M(A)}}{\diagdown\!\!\diagdown} & \big\downarrow{\scriptstyle m} \\ MM(A) & \xrightarrow{\ m\ } & M(A) \end{array}$$

- We have the commutative diagram (which is like saying that m is associative)

$$\begin{array}{ccc} MMM(A) & \xrightarrow{M(m_A)} & MM(A) \\ \big\downarrow{\scriptstyle m_{M(A)}} & & \big\downarrow{\scriptstyle m_A} \\ MM(A) & \xrightarrow{\ m\ } & M(A) \end{array}$$

Monads from adjoint functors of forgetful functors to Set

We look at the monads that we have already constructed in the the situation where the category is **Set**.

In the first case, by factoring through monads, we associate to a set S, the underlying set associated with the free monad $M(S) = S^*$. We have already constructed the morphisms $u : S \to S^*$ given by $s \mapsto (s)$ and $m : M(S^*) \to S^*$ which sends a tuple of tuples to the concatenation of these tuples.

$$((a_1, \ldots, a_p), (b_1, \ldots, b_q), \ldots, (c_1, \ldots, c_r)) \\ \mapsto (a_1, \ldots, a_p, b_1, \ldots, b_q, \ldots, c_1, \ldots, c_r)$$

In the second case, we factor through Abelian groups. We associate to a set S, the underlying set of the free Abelian group $\langle S \rangle$ whose basis (over $\mathbb{Z}$) is given by $\{e_s\}_{s \in S}$. We have already seen that the map $S \to \langle S \rangle$ is given by $s \mapsto e_s$. The map $m : \langle\langle S \rangle\rangle \to \langle S \rangle$ is given by sending, for each $a \in \langle S \rangle$, the element e_a in $\langle\langle S \rangle\rangle$ to the element a. Since $\langle\langle S \rangle\rangle$ is generated freely by e_a, this determines m.

A similar construction can be carried out when we factor through commutative rings. We will examine this case a little differently below.

Abelianisation

We also studied the example where G is the forgetful functor from **Ab** to **Grp** and F is the functor from **Grp** which associates to a group K, the Abelian group $K^{\mathrm{ab}} = K/[K, K]$.

In this case, the functor $M = GF$ from **Grp** to itself, which associates to a group K, the group K^{ab}. The natural transformation u is the homomorphism $K \to K/[K,K] = K^{\text{ab}} = MK$. For an Abelian group A, we have $A = A^{\text{ab}}$ so $MMK = MK = K^{\text{ab}}$ so the natural tranformation $m : MM \to M$ is the identity homomorphism.

One easily checks the commutativity of the required diagrams.

Power Set functor

We have a functor P from **Set** to itself that associates to a set S, the power set $P(S)$ that parametrises subsets of S. Note that if $f : S \to T$ is a set map, the associated set map $P(f) : P(S) \to P(T)$ is the map that takes a subset R of S to its image $f(R)$ in T. One easily checks that this is a functor.

There is a natural transformation u given by $u_S : S \to P(S)$ that takes an element s of S to the singleton set $\{s\}$ which is an element of $P(S)$. A subset T of $P(S)$ gives an indexed collection $\{R_t\}_{t\in T}$ of subsets R_t of S for each t. We define $m(T) = \cup_{t\in T} R_t$ which is the union of the R_t as subsets of S. This defines a set map $m_S : P(P(S)) \to P(S)$ which we check is a natural transformation $m : PP \to P$

One then checks the commutativity of the required diagrams.

Algebraic structures and Monads

In each of the cases of monads associated with forgetful functors above, there is a natural category through which M "factors". To what extent is this true in general?

Given a monad M on a category $\mathcal{C}$, we define an algebra of type M to be an object A of $\mathcal{C}$ with a morphism $e : M(A) \to A$ such that the following diagrams commute:

- The identity of M leads to identity on A.

$$\begin{array}{ccc} A & \xrightarrow{u_A} & M(A) \\ & \overset{1_A}{\searrow\!\!\searrow} & \downarrow e \\ & & A \end{array}$$

- The operation of A is associative in the sense of M

$$\begin{array}{ccc} MM(A) & \xrightarrow{M(e)} & M(A) \\ \downarrow m_A & & \downarrow e \\ M(A) & \xrightarrow{e} & A \end{array}$$

Polynomial monad and commutative rings

Let us see what this means for the monad C which associates to a set S the underlying set of the ring $\mathbb{Z}[S]$ of plynomials in variables x_s for s in S.

The map $u_S : S \to \mathbb{Z}[S]$ sends an element s to the variable x_s. This plays the role of "identity" for the monad C as it gives a morphism $u_S : S \to C(S)$.

When R is a commutative ring, and $f : T \to R$ is a set map, we have a homomorphism $\mathbb{Z}[T] \to R$ which sends which sends the variable x_t to $f(t)$ for every t in T. If $p \in \mathbb{Z}[T]$ is a polynomial in the variables $(x_t)_{t\in T}$, then its image in R can be seen as the result of evaluating the polynomial p by substituting x_t by $f(t)$ for each element $t \in T$.

Applying this to $R = \mathbb{Z}[S]$ and the identity map $1_R : R \to R$, we get a map $m_S : \mathbb{Z}\left[\mathbb{Z}[S]\right] \to \mathbb{Z}[S]$. The ring $\mathbb{Z}\left[\mathbb{Z}[S]\right]$ is the collection of polynomials q in the variables x_p where p varies over polynomials in the variables $(x_s)_{s\in S}$. Under the map m_S, we send q to the polynomial obtained when we substitute x_p by p.

This gives the multiplication map $m_S : CC(S) \to C(S)$. With the natural transformations u_S and m_S, one checks that C becomes a monad.

A C-algebra (or an algebra of type C) comprises have a set A with a set map $e : C(A) \to A$ which satisfiy some identities (or commutative diagrams).

Note that $C(A) = \mathbb{Z}[A]$ is the polynomial ring in the variables x_a for elements a in A. So for every polynomial p in the variables $(x_a)_{a\in A}$, we have an element $e(p)$ in A. In particular, we can define some natural elements and operations for A as follows.

- The element $0 \in A$ as the image $e(0)$ of the 0 polynomial.
- The element $1 \in A$ as the image $e(1)$ of the 0 polynomial.
- Given $a, b \in A$ we define $a + b$ as the image $e(x_a + x_b)$ of the polynomial $x_a + x_b$.
- Given $a, b \in A$ we define $a \cdot b$ as the image $e(x_a x_b)$ of the polynomial $x_a x_b$.
- Given $a \in A$ we define $-a$ (the additive inverse of a) as the image $e(-x_a)$ of the polynomial $-x_a$.

First of all, we can apply e to the polynomial which is just the single variable x_a for $a \in A$. Note that $x_a = u_A(a)$ as seen above. The "identity" property for the monad says that $e(u_A(a)) = a$ which means that $e(x_a) = a$ as expected! In particular, we note that $e(x_0) = 0$, $e(x_1) = 1$ and so on.

The map $m_A : CC(A) \to C(A)$ takes a polynomial q in the variables $(x_p)_{p\in C(A)}$ to the polynomial $m_A(q)$ in the variables $(x_a)_{a\in A}$ obtained from q by substituting x_p by p.

On the other hand, we see that the map $C(e) : CC(A) \to C(A)$ takes a polynomial q in the variables $(x_p)_{p\in C(A)}$ to $C(e)(q)$ which is obtained from q by substituting x_p by $x_{e(p)}$.

The second diagram for the C-algbra A says that $e(m_A(q)) = e(C(e)(q))$ for every q. This will lead to various required identities.

- Since $x_a + x_b = x_b + x_a$ and $x_a \cdot x_b = x_b \cdot x_a$ in $C(A)$ we see that $a+b = b+a$ and $a \cdot b = b \cdot a$ as these are defined to be the images under e of the above polynomials.
- Consider the polynomial $q = x_{x_a} + x_0$ where x_a and 0 are considered as elements of $C(A)$. Under m_A the image of q is $x_a + 0 = x_a$, so $e(m_A((q)) = a$. On the other hand $C(e)(q) = x_a + x_0$ since $e(x_a) = a$ and $e(x_0) = 0$, so $e(C(e)(q)) = e(x_a + x_0) = a + 0$. Thus, the above identity says that $a = a + 0$.
- Similarly, consider $q = x_{x_a} x_1$, where x_a and 1 are considered as elements of $C(A)$. Its image under m_A is $x_a \cdot 1 = x_a$ while its image under $C(e)$ is $x_a \cdot x_1$ in $C(A)$. The images of these two in A must be equal which shows that $a \cdot 1 = a$.
- We now consider $q = x_{x_a} + x_{-x_a}$. Its image under m_A is $x_a - x_a = 0$ and its image under $C(e)$ is $x_a + x_{-a}$. Hence, we obtain the identity $a + (-a) = 0$.
- Consider $q = x_{x_a} + x_{x_b + x_c}$, where x_a and $x_b + x_c$ are considered as elements of $C(A)$. Its image under m_A is $x_a + x_b + x_c$ while its image under $C(e)$ is $x_a + x_{b+c}$. It follows that the images of these under e are equal, giving us the identity $e(x_a + x_b + x_c) = a + (b + c)$. Since this is true for all a, b, c we see that $a + (b + c) = c + (a + b) = (a + b) + c$, where the last identity is due to commutativity of addition which was proved above.
- Similarly, consider $q = x_{x_a} x_{x_b x_c}$, where x_a and $x_b x_c$ are considered as elements of $C(A)$. Its image under m_A is $x_a x_b x_c$ while its image under $C(e)$ is $x_a x_{b \cdot c}$. As above, taking images under e we obtain the identity $e(x_a x_b x_c) = a \cdot (b \cdot c)$ for a, b and c. We then get the identity $a \cdot (b \cdot c) = c \cdot (a \cdot b) = (a \cdot b) \cdot c$ by also using commutativity of multiplication which was proved above.
- Finally, consider $q = x_{x_a x_b} + x_{x_a x_c}$. Its image under m_A is $x_a x_b + x_a x_c = x_a(x_b + x_c)$, while its image under $C(e)$ is $x_{a \cdot b} + x_{a \cdot c}$. Taking images under e, we obtain the identity $e(x_a(x_b + x_c)) = a \cdot b + a \cdot c$. At the same time, we see that if $q = x_{x_a} x_{x_b + x_c}$, then $m_A(q) = x_a(x_b + x_c)$ and $C(e)(q) = x_a x_{b+c}$; applying e gives the identity $e(x_a(x_b + x_c)) = a \cdot (b + c)$. Combining these gives $a \cdot b + a \cdot c = a \cdot (b + c)$.

In conclusion, we see that if A is an algebra of type C, then A is a commutative ring in a natural way and $e : C(A) \to A$ is the map that takes a polynomial p in the variables $(x_a)_{a \in A}$ to the result of substituting x_a by a in p.