Chapter 7: Computations lifted to a functor context II
Part 2: Laws and structure of semimonads

Sergei Winitzki

Academy by the Bay

2018-03-11

Semimonad laws I: The intuitions

What properties of functor block programs do we expect to have?

- In `x ← c`, the value of `x` will *go over items* held in container `c`
- Manipulating items in container is followed by a generator:

```
x ← cont1                      y ← cont1
y = f(x)                         .map(x ⇒ f(x))
z ← cont2(y)                   z ← cont2(y)
```

```
cont1.flatMap(x ⇒ cont2(f(x))) = cont1.map(f).flatMap(y ⇒ cont2(y))
```

- Manipulating items in container is preceded by a generator:

```
x ← cont1                      x ← cont1
y ← cont2(x)                   z ← cont2(x)
z = f(y)                         .map(f)
```

```
cont1.flatMap(cont2).map(f) = cont1.flatMap(x ⇒ cont2(x).map(f))
```

- After `x ← c`, further computations will use *all those* `x`

```
x ← cont                       y ← for { x ← cont
y ← p(x)                                 yy ← p(x) } yield yy
z ← cont2(y)                   z ← cont2(y)
```

```
cont.flatMap(x ⇒ p(x).flatMap(cont2)) = cont.flatMap(p).flatMap(cont2)
```

Semimonad laws II: The laws for `flatMap`

To use the concise notation, denote `flatMap` by flm
A **semimonad** S^A has $\text{flm}^{[S,A,B]} : (A \Rightarrow S^B) \Rightarrow S^A \Rightarrow S^B$ with 3 laws:

❶ $\text{flm}\,(f^{A \Rightarrow B} \circ g^{B \Rightarrow S^C}) = \text{fmap}\,f \circ \text{flm}\,g$ (naturality in A)

$$S^A \xrightarrow[\text{flm}\,(f^{A \Rightarrow B} \circ g^{B \Rightarrow S^C})]{\text{fmap}\,f^{A \Rightarrow B}} \; S^B \xrightarrow{\text{flm}\,g^{B \Rightarrow S^C}} S^C$$

❷ $\text{flm}\left(f^{A \Rightarrow S^B} \circ \text{fmap}\,g^{B \Rightarrow C}\right) = \text{flm}\,f \circ \text{fmap}\,g$ (naturality in B)

$$S^A \xrightarrow[\text{flm}\,(f^{A \Rightarrow S^B} \circ \text{fmap}\,g^{B \Rightarrow C})]{\text{flm}\,f^{A \Rightarrow S^B}} \; S^B \xrightarrow{\text{fmap}\,g^{B \Rightarrow C}} S^C$$

❸ $\text{flm}\left(f^{A \Rightarrow S^B} \circ \text{flm}\,g^{B \Rightarrow S^C}\right) = \text{flm}\,f \circ \text{flm}\,g$ (composition)

$$S^A \xrightarrow[\text{flm}\left(f^{A \Rightarrow S^B} \circ \text{flm}\,g^{B \Rightarrow S^C}\right)]{\text{flm}\,f^{A \Rightarrow S^B}} \; S^B \xrightarrow{\text{flm}\,g^{B \Rightarrow S^C}} S^C$$
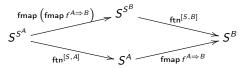
Is there a shorter formulation of the laws?

The methods `flatten` (denoted by ftn) and `flatMap` are equivalent:

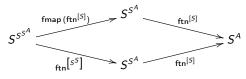$$\text{ftn}^{[S,A]} : S^{S^A} \Rightarrow S^A = \text{flm}^{[S,S^A,A]}(m^{S^A} \Rightarrow m)$$

$$\text{flm}\left(f^{A \Rightarrow S^B}\right) = \text{fmap } f \circ \text{ftn}$$

It turns out that `flatten` has only 2 laws:

1 $\text{fmap}\left(\text{fmap } f^{A \Rightarrow B}\right) \circ \text{ftn}^{[S,B]} = \text{ftn}^{[S,A]} \circ \text{fmap } f$ (naturality)

2 $\text{fmap}\left(\text{ftn}^{[S]}\right) \circ \text{ftn}^{[S]} = \text{ftn}^{[S^S]} \circ \text{ftn}^{[S]}$ (associativity)

Denote for brevity $q_\uparrow \equiv \mathrm{fmap}^{[S]}\, q$ for any function q

Express $\mathrm{flm}\, f = f_\uparrow \circ \mathrm{ftn}$ and substitute that into flm's 3 laws:

1. $\mathrm{flm}\,(f \circ g) = f_\uparrow \circ \mathrm{flm}\, g$ gives $(f \circ g)_\uparrow \circ \mathrm{ftn} = f_\uparrow \circ g_\uparrow \circ \mathrm{ftn}$
   – this law holds automatically due to functor composition law

2. $\mathrm{flm}\,(f \circ g_\uparrow) = \mathrm{flm}\, f \circ g_\uparrow$ gives $(f \circ h)_\uparrow \circ \mathrm{ftn} = f_\uparrow \circ \mathrm{ftn} \circ h$;
   using the functor composition law, we reduce this to $h_\uparrow \circ \mathrm{ftn} = \mathrm{ftn} \circ h$
   – this is the naturality law for `flatten`

3. $\mathrm{flm}\,(f \circ \mathrm{flm}\, g) = \mathrm{flm}\, f \circ \mathrm{flm}\, g$ with functor composition law gives
   $f_\uparrow \circ g_{\uparrow\uparrow} \circ \mathrm{ftn}_\uparrow \circ \mathrm{ftn} = f_\uparrow \circ \mathrm{ftn} \circ g_\uparrow \circ \mathrm{ftn}$; using ftn's naturality and omitting
   the common factor $f_\uparrow \circ g_{\uparrow\uparrow}$, we get ftn's associativity: $\mathrm{ftn}_\uparrow \circ \mathrm{ftn} = \mathrm{ftn} \circ \mathrm{ftn}$

- `flatten` has the simplest type signature and the fewest laws
- It is usually easy to check naturality!
  - **Parametricity theorem**: Any *fully parametric* code for a function of
    type $F^A \Rightarrow G^A$ implements a natural transformation $F \rightsquigarrow G$
- Checking `flatten`'s associativity needs more work
  - The "accumulated context" is completely different for these monads

The `cats` library has a `FlatMap` type class, defining `flatten` via `flatMap`

Semimonad laws IV: Checking the laws of `flatten`

- Implement `flatten` for these functors and check the laws (see code):
  - `Option` monad: $F^A \equiv 1 + A$; ftn : $1 + (1 + A) \Rightarrow 1 + A$
  - `Either` monad: $F^A \equiv Z + A$; ftn : $Z + (Z + A) \Rightarrow Z + A$
  - `List` monad: $F^A \equiv \text{List}^A$; ftn : $\text{List}^{\text{List}^A} \Rightarrow \text{List}^A$
  - Writer monad: $F^A \equiv A \times W$; ftn : $(A \times W) \times W \Rightarrow A \times W$
  - Reader monad: $F^A \equiv R \Rightarrow A$; ftn : $(R \Rightarrow (R \Rightarrow A)) \Rightarrow R \Rightarrow A$
  - State: $F^A \equiv S \Rightarrow A \times S$; ftn : $(S \Rightarrow (S \Rightarrow A \times S) \times S) \Rightarrow S \Rightarrow A \times S$
  - Continuation monad: $F^A \equiv (A \Rightarrow R) \Rightarrow R$;
    ftn : $((((A \Rightarrow R) \Rightarrow R) \Rightarrow R) \Rightarrow R) \Rightarrow (A \Rightarrow R) \Rightarrow R$
- Code implementing these `flatten` functions is *fully parametric* in A
  - Naturality of these functions follows from parametricity theorem
- Associativity needs to be checked for each monad!
- Example of a useful semimonad that is *not* a full monad:
  - $F^A \equiv A \times V \times W$; ftn $((a \times v_1 \times w_1) \times v_2 \times w_2) = a \times v_1 \times w_1$
- Examples of *non-associative* (i.e. wrong) implementations of `flatten`:
  - $F^A \equiv A \times W \times W$; ftn $((a \times v_1 \times v_2) \times w_1 \times w_2) = a \times w_2 \times w_1$
  - $F^A \equiv \text{List}^A$, but `flatten` concatenates the nested lists in reverse order

Exercises II

1. Implement

Structure of semimonads

How to recognize a semimonad by its type?

Intuition from `flatten`: reshuffle data in F^{F^A} to fit into F^A

Some constructions of exponential-polynomial semimonads

1. $F^A = Z$ (constant functor) for a fixed type Z
2. $F^A \equiv G^A \times H^A$ for any semimonads G^A and H^A
3. $F^A \equiv G^A + G^{F^A}$ (recursive) for any functor G^A

* Worked examples II: Constructions of filterable functors I

(2) The `fmapOpt` laws hold for $F^A \times G^A$ if they hold for F^A and G^A

- For $f^{A \Rightarrow 1+B}$, get $\text{fmapOpt}_F(f) : F^A \Rightarrow F^B$ and $\text{fmapOpt}_G(f) : G^A \Rightarrow G^B$
- Define $\text{fmapOpt}_{F \times G} f \equiv p^{F^A} \times q^{G^A} \Rightarrow \text{fmapOpt}_F(f)(p) \times \text{fmapOpt}_G(f)(q)$
- Identity law: $f = \text{id}_{\diamond_{\text{Opt}}}$, so $\text{fmapOpt}_F f = \text{id}$ and $\text{fmapOpt}_G f = \text{id}$
  - Hence we get $\text{fmapOpt}_{F+G}(f)(p \times q) = \text{id}(p) \times \text{id}(q) = p \times q$
- Composition law:

$$(\text{fmapOpt}_{F \times G} f_1 \circ \text{fmapOpt}_{F+G} f_2)(p \times q)$$
$$= \text{fmapOpt}_{F \times G}(f_2)\left(\text{fmapOpt}_F(f_1)(p) \times \text{fmapOpt}_G(f_1)(q)\right)$$
$$= (\text{fmapOpt}_F f_1 \circ \text{fmapOpt}_F f_2)(p) \times (\text{fmapOpt}_G f_1 \circ \text{fmapOpt}_G f_2)(q)$$
$$= \text{fmapOpt}_F(f_1 \diamond_{\text{Opt}} f_2)(p) \times \text{fmapOpt}_G(f_1 \diamond f_2)(q)$$
$$= \text{fmapOpt}_{F \times G}(f_1 \diamond_{\text{Opt}} f_2)(p \times q)$$

- Exactly the same proof as that for functor property for $F^A \times G^A$
  - this is because `fmapOpt` corresponds to a generalized functor
- New proofs are necessary only when using non-filterable functors
  - these are used in constructions $4 - 6$

* Worked examples II: Constructions of filterable functors II

(5) The `fmapOpt` laws hold for $F^A \equiv 1 + A \times G^A$ if they hold for G^A

- For $f^{A \Rightarrow 1+B}$, get $\mathsf{fmapOpt}_G(f) : G^A \Rightarrow G^B$
- Define $\mathsf{fmapOpt}_F(f)(1 + a^A \times q^{G^A})$ by returning $0 + b \times \mathsf{fmapOpt}_G(f)(q)$ if the argument is $0 + a \times q$ and $f(a) = 0 + b$, and returning $1 + 0$ otherwise
- Identity law: $f = \mathsf{id}_{\diamond_{\mathbf{Opt}}}$, so $f(a) = 0 + a$ and $\mathsf{fmapOpt}_G f = \mathsf{id}$
  - Hence we get $\mathsf{fmapOpt}_F(\mathsf{id}_{\diamond_{\mathbf{Opt}}})(1 + a \times q) = 1 + a \times q$
- Composition law: need only to check for arguments $0 + a \times q$, and only when $f_1(a) = 0 + b$ and $f_2(b) = 0 + c$, in which case $(f_1 \diamond_{\mathbf{Opt}} f_2)(a) = 0 + c$; then

$$(\mathsf{fmapOpt}_F f_1 \circ \mathsf{fmapOpt}_F f_2)(0 + a \times q)$$
$$= \mathsf{fmapOpt}_F(f_2) \left(\mathsf{fmapOpt}_F(f_1)(0 + a \times q) \right)$$
$$= \mathsf{fmapOpt}_F(f_2) \left(0 + b \times \mathsf{fmapOpt}_G(f_1)(q) \right)$$
$$= 0 + c \times (\mathsf{fmapOpt}_G f_1 \circ \mathsf{fmapOpt}_G f_2)(q)$$
$$= 0 + c \times \mathsf{fmapOpt}_G(f_1 \diamond_{\mathbf{Opt}} f_2)(q)$$
$$= \mathsf{fmapOpt}_F(f_1 \diamond_{\mathbf{Opt}} f_2)(0 + a \times q)$$

This is a "greedy filter": if $f(a)$ is empty, will delete all data in G^A

* Worked examples II: Constructions of filterable functors III

(6) The `fmapOpt` laws hold for $F^A \equiv G^A + A \times F^A$ if they hold for G^A

- For $f^{A \Rightarrow 1+B}$, we have $\text{fmapOpt}_G(f) : G^A \Rightarrow G^B$ and $\text{fmapOpt}'_F(f) : F^A \Rightarrow F^B$ (for use in recursive arguments as the inductive assumption)

- Define $\text{fmapOpt}_F(f)(q^{G^A} + a^A \times p^{F^A})$ by returning $0 + \text{fmapOpt}'_F(f)(p)$ if $f(a) = 1 + 0$, and $\text{fmapOpt}_G(f)(q) + b \times \text{fmapOpt}'_F(f)(p)$ otherwise

- Identity law: $\text{id}_{\diamond_{\text{Opt}}}(x) \neq 1 + 0$, so $\text{fmapOpt}_F(\text{id}_{\diamond_{\text{Opt}}})(q + a \times p) = q + a \times p$

- Composition law:
  $(\text{fmapOpt}_F(f_1) \circ \text{fmapOpt}_F(f_2))(q + a \times p) = \text{fmapOpt}_F(f_1 \diamond_{\text{Opt}} f_2)(q + a \times p)$

- For arguments $q + 0$, the laws for G^A hold; so assume arguments $0 + a \times p$. When $f_1(a) = 0 + b$ and $f_2(b) = 0 + c$, the proof of the previous example will go through. So we need to consider the two cases $f_1(a) = 1 + 0$ and $f_1(a) = 0 + b$, $f_2(b) = 1 + 0$

- If $f_1(a) = 1 + 0$ then $(f_1 \diamond_{\text{Opt}} f_2)(a) = 1 + 0$; to show $\text{fmapOpt}'_F(f_2)(\text{fmapOpt}'_F(f_1)(p)) = \text{fmapOpt}'_F(f_1 \diamond_{\text{Opt}} f_2)(p)$, use the inductive assumption about $\text{fmapOpt}'_F$ on p

- If $f_1(a) = 0 + b$ and $f_2(b) = 1 + 0$ then $(f_1 \diamond_{\text{Opt}} f_2)(a) = 1 + 0$; to show $\text{fmapOpt}_F(f_2)(0 + b \times \text{fmapOpt}'_F(f_1)(p)) = \text{fmapOpt}'_F(f_1 \diamond_{\text{Opt}} f_2)(p)$, rewrite $\text{fmapOpt}_F(f_2)(0 + b \times \text{fmapOpt}'_F(f_1)(p)) = \text{fmapOpt}'_F(f_2)(\text{fmapOpt}'_F(f_1)(p))$ and again use the inductive assumption about $\text{fmapOpt}'_F$ on p

This is a "list-like filter": if $f(a)$ is empty, will recurse into nested F^A data

Worked examples II: Constructions of filterable functors IV

Use known filterable constructions to show that
$F^A \equiv (\text{Int} \times \text{String}) \Rightarrow (1 + \text{Int} \times A + A \times (1 + A) + (\text{Int} \Rightarrow 1 + A + A \times A \times \text{String}))$
is a filterable functor

- Instead of implementing `Filterable` and verifying laws by hand, we analyze the structure of this data type and use known constructions
- Define some auxiliary functors that are parts of the structure of F^A,
  - $R_1^A = (\text{Int} \times \text{String}) \Rightarrow A$ and $R_2^A = \text{Int} \Rightarrow A$
  - $G^A = 1 + \text{Int} \times A + A \times (1 + A)$ and $H^A = 1 + A + A \times A \times \text{String}$
- Now we can rewrite $F^A = R_1 \left[G^A + R_2 \left[H^A \right] \right]$
  - G^A is filterable by construction 5 because it is of the form $G^A = 1 + A \times K^A$ with filterable functor $K^A = 1 + \text{Int} + A$
  - K^A is of the form $1 + A + X$ with constant type X, so it is filterable by constructions 1 and 3 with the `Option` functor $1 + A$
  - H^A is filterable by construction 5 with $H^A = 1 + A \times (1 + A \times \text{String})$, while $1 + A \times \text{String}$ is filterable by constructions 5 and 1
- Constructions 3 and 4 show that $R_1 \left[G^A + R_2 \left[H^A \right] \right]$ is filterable

Note that there are more than one way of implementing `Filterable` here

* Exercises II

1. Implement a `Filterable` instance for `type F[T] = G[H[T]]` assuming that the functor `H[T]` already has a `Filterable` instance (construction 4). Verify the laws rigorously (i.e. by calculations, not tests).

2. For `type F[T] = Option[Int ⇒ Option[(T, T)]]`, implement a `Filterable` instance. Show that the filterable laws hold by using known filterable constructions (avoiding explicit proofs or tests).

3. Implement a `Filterable` instance for $F^A \equiv G^A + \text{Int} \times A \times A \times F^A$ (recursive) for a filterable functor G^A. Verify the laws rigorously.

4. Show that $F^A = 1 + A \times G^A$ is in general *not* filterable if G^A is an arbitrary (non-filterable) functor; it is enough to give an example.

5. Show that $F^A = 1 + G^A + H^A$ is filterable if $1 + G^A$ and $1 + H^A$ are filterable (even when G^A and H^A are by themselves not filterable).

6. Show that the functor $F^A = A + (\text{Int} \Rightarrow A)$ is not filterable.

7. Show that one can define `deflate`: $C^{1+A} \Rightarrow C^A$ for any contrafunctor C^A (not necessarily filterable), similarly to how one can define `inflate`: $F^A \Rightarrow F^{1+A}$ for any functor F^A (not necessarily filterable).