Chalmers | Göteborgs Universitet
Alejandro Russo, Computer Science and Engineering

Advanced Functional Programming TDA342/DIT260

Tuesday, March 15, 2016, Hörsalsvägen (yellow brick building), 8:30-12:30.

Alejandro Russo, tel. 031 772 6156

- The maximum amount of points you can score on the exam: 60 points. The grade for the exam is as follows:

  Chalmers: **3**: 24 - 35 points, **4**: 36 - 47 points, **5**: 48 - 60 points.
  GU: Godkänd 24-47 points, Väl godkänd 48-60 points
  PhD student: 36 points to pass.

- Results: within 21 days.

- **Permitted materials (Hjälpmedel):** Dictionary (Ordlista/ordbok).

  You may bring up to two pages (on one A4 sheet of paper) of pre-written notes – a "summary sheet". These notes may be typed or handwritten. They may be from any source. If this summary sheet is brought to the exam it must also be handed in with the exam (so make a copy if you want to keep it).

- **Notes:**

  – Read through the paper first and plan your time.
  – Answers preferably in English, some assistants might not read Swedish.
  – If a question does not give you all the details you need, you may make reasonable assumptions. Your assumptions must be clearly stated. If your solution only works under certain conditions, state them.
  – Start each of the questions on a new page.
  – The exact syntax of Haskell is not so important as long as the graders can understand the intended meaning. If you are unsure just put in an explanation of your notation.
  – Hand in the summary sheet (if you brought one) with the exam solutions.
  – As a recommendation, consider spending around 1h 20 minutes per exercise. However, this is only a recommendation.
  – To see your exam: *by appointment (send email to Alejandro Russo)*

FUNCTOR TYPE-CLASS **class** *Functor* c **where** $fmap :: (a \to b) \to c\ a \to c\ b$	IDENTITY $fmap\ id \equiv id$ **where** $id = \lambda x \to x$

MAP FUSION
$fmap\ (f \circ g) \equiv fmap\ f \circ fmap\ g$

Figure 1: Functors

Problem 1: (**Functors**) As its name implies, a binary tree is a tree with a two-way branching structure, i.e., a left and a right sub tree. In Haskell, such trees can be defined as follows.

$$
\begin{aligned}
&\textbf{data}\ Tree\ a\ \textbf{where}\\
&\quad Leaf\ :: a \to Tree\ a\\
&\quad Node :: Tree\ a \to Tree\ a \to Tree\ a
\end{aligned}
$$

a) Show that *Tree a* is a functor. For that, you should *provide* an instance for the *Functor* type-class and *prove* that *fmap* for *finite trees*, i.e., $fmap :: (a \to b) \to Tree\ a \to Tree\ b$, fulfills the laws for functors – see Figure 1.

Important: Assume that f and g are total, i.e., they do not raise any errors or loop indefinitely when applied to an argument. If your proof is by induction, you should indicate *induction on what* (e.g., in the length of the list). Justify every step in your proof.

(8p)

b) As with lists, it is also useful to "fold" over trees. Given a tree t with elements e_1, e_2, ..., e_n and an operator $\oplus$, folding over the tree t with operator $\oplus$ intuitively means to *intercalate* the operator among the elements of the tree, i.e., $e_1 \oplus e_2 \oplus e_3 \oplus \ldots \oplus e_n$. For simplicity, we assume that the operator $\oplus$ is always associative. We call the function implementing folding over trees *foldT*.

$$foldT :: (a \to a \to a) \to Tree\ a \to a$$

By using *foldT*, we can now express a bunch of useful functions on trees.

P_1 $height\_tree = foldT\ (\lambda l\ r \to max\ l\ r + 1) \circ fmap\ (const\ 0)$	P_2 $sum\_tree = foldT\ (+)$

P_3
$leaves = foldT\ (+\!\!+) \circ fmap\ (\lambda x \to [x])$

Program P_1 computes the height of a tree. Program P_2 sums all the numbers in a tree. Program P_3 extracts all the elements of a tree.

Your task is to implement *foldT*. *(4p)*

c) There is a relation between mapping functions over trees' leaves and lists. More specifically, we have the following equation for finite and well-defined trees.

$$map\ f \circ leaves \equiv leaves \circ fmap\ f$$

It is the same to first extract the leaves and then map the function (left-hand side), as it is to map the function first and then extracting the leaves (right-hand side).

Your task is to prove that the equation holds.

You can assume the following properties and definition for this exercise and the rest of the exam!

(.)	Assoc. (.)	(id left)	(id right)	(eta)
$(f \circ g)\ x = f\ (g\ x)$	$(f \circ g) \circ z = f \circ (g \circ z)$	$id \circ f = f$	$f \circ id = f$	$\lambda x \to f\ x \equiv f$

(cons.0)	((++).0)	((++).1)
$x : [\,] = [x]$	$[\,] +\!\!+\ ys = ys$	$(x : xs) +\!\!+\ ys = x : (xs +\!\!+\ ys)$

(Assoc. (++))	(*map*.0)	(*map*.1)
$xs +\!\!+\ (ys +\!\!+\ zs) \equiv (xs +\!\!+\ ys) +\!\!+\ zs$	$map\ f\ [\,] = [\,]$	$map\ f\ (x : xs) = f\ x : map\ f\ xs$

You cannot assume any property that relates (++), *map*, and *fmap* – if you need such properties, you should prove them too! *(8p)*

$$
\begin{array}{lll}
\textbf{class}\ \mathit{Monad}\ m\ \textbf{where} & \textsc{Left Identity} & \textsc{Right Identity} \\
\quad \mathit{return} :: a \to m\ a & \mathit{return}\ x \gg\!= f \equiv f\ x & m \gg\!= \mathit{return} \equiv m \\
\quad (\gg\!=)\ :: m\ a \to (a \to m\ b) \to m\ b & &
\end{array}
$$

$$
\begin{array}{l}
\textsc{Associativity}\ (x\ \textsc{does not appear in}\ k_1\ \textsc{and}\ k_2) \\
(m \gg\!= k_1) \gg\!= k_2 \equiv m \gg\!= (\lambda x \to k_1\ x \gg\!= k_2)
\end{array}
$$

Figure 2: Monads

Problem 2. (**Monads**) During the lectures we said that a data type m is a monad if we can define the primitives *return* and $(\gg\!=)$, and that m fulfills the monadic laws – see Figure 2. There is, however, an alternative interface for monads described as follows.

$$
\begin{array}{lll}
\textbf{class}\ \mathit{MonadAlternative}\ m\ \textbf{where} & \textsc{Identity} & \textsc{Map fusion} \\
\quad \mathit{return}' :: a \to m\ a & \mathit{fmap}'\ \mathit{id}\ m \equiv m & \mathit{fmap}'\ (f \circ g) \equiv \mathit{fmap}'\ f \circ \mathit{fmap}'\ g \\
\quad \mathit{join}\ \ :: m\ (m\ a) \to m\ a & & \\
\quad \mathit{fmap}'\ :: (a \to b) \to m\ a \to m\ b & &
\end{array}
$$

$$
\begin{array}{lll}
A_1 & A_2 & A_3 \\
\mathit{fmap}'\ f \circ \mathit{return}' \equiv \mathit{return}' \circ f & \mathit{join} \circ \mathit{fmap}'\ \mathit{return}' \equiv \mathit{id} & \mathit{join} \circ \mathit{return}' \equiv \mathit{id}
\end{array}
$$

$$
\begin{array}{ll}
A_4 & A_5 \\
\mathit{join} \circ \mathit{fmap}'\ \mathit{join} \equiv \mathit{join} \circ \mathit{join} & \mathit{join} \circ \mathit{fmap}'\ (\mathit{fmap}'\ f) \equiv \mathit{fmap}'\ f \circ \mathit{join}
\end{array}
$$

This interface requires m to be a functor and introduces an operation called *join*. Furthermore, *return*′, *join*, and *fmap*′ are required to obey various different laws.

a) Your task consists of showing that the alternative interface is enough to implement *return* and $(\gg\!=)$. In other words, if you define *return*′, *fmap*′, and *join* for certain data type m, then you can show that m is an instance of the type-class *Monad* in Haskell. You should provide the following type-class instance:

$$
\begin{array}{l}
\textbf{instance}\ \mathit{MonadAlternative}\ m \Rightarrow \mathit{Monad}\ m\ \textbf{where} \\
\quad \mathit{return} = \ldots \\
\quad (\gg\!=)\ \ = \ldots
\end{array}
$$

(6p)

b) Assuming the laws for the alternative monadic interface, you should show that the implementation that you gave in the previous question is indeed a monad in the traditional sense, i.e. it fulfills the laws from Figure 2. *(14p)*

Problem 3: (**EDSL**) *Information-flow control* (IFC) is a promising technology to guarantee confidentiality of data when manipulated by untrusted code, i.e. code written by someone else. In IFC, data gets classified either as *public* (low) or *secret* (high), where public information can flow into secret entities but not vice versa. We encode the sensitivity of data as abstract data types, and the allowed flows of information in the type-class *CanFlowTo* – see Figure 3.

```
  -- Security level for public data
data L
  -- Security level for secret data
data H
  -- allowed flows of information
class l `CanFlowTo` l' where
  -- Public data can flow into public entities
instance L `CanFlowTo` L where
  -- Public data can flow into secret entities
instance L `CanFlowTo` H where
  -- Secret data can flow into secret entities
instance H `CanFlowTo` H where
```

Figure 3: Allowed flows of information

To build secure programs which do not leak secrets, we build a small EDSL in Haskell with two core concepts: *labeled values* and *secure computations*. Labeled values are simply data tagged with a security level indicating its sensitivity. For example, a weather report is a public piece of data, so we can model it as a public labeled string *weather_report* :: *Labeled L String*. Similarly, a credit card number is sensitive, so we model it as a secret integer *cc_number* :: *Labeled H Integer*.

A secure computation is an entity of type *MAC l a*, which denotes a computation that handles data at sensitivity level *l* and produces a result (of type *a*) of this level. In order to remain secure, secure computations can only observe data that "can flow to" the computation (see primitive *unlabel* below), and can only create labeled values provided that information from the computation "can flow to" the newly created labeled value (see primitive *label* below). We describe the API for the EDSL in Figure 4, and provide a *deep-embedded* implementation for the API in Figure 5.

a) Your task is to take the implementation in Figure 5 and obtain an "intermediate embedding" by removing *Bind* from the *MAC l a* data type. As a result, *runMAC* will no longer run *Bind*; instead, the defintion of (>>=) will change. After your modifications, it is important to show that you can faithfully implement the *whole* EDSL API.

Important: If you alter the definition of *MAC l a*, or any other function in the deep-embedded implementation, you need to show that your modifications are correct by deriving them.

Help: You can assume that $runMAC\ (m \gg\!= f) \equiv runMAC\ m \gg\!= runMAC \circ f$ *(12p)*

```
-- Types
newtype Labeled l a
data    MAC l a
-- Labeled values
label    :: (l `CanFlowTo` l') => a -> MAC l (Labeled l' a)
unlabel  :: (l' `CanFlowTo` l) => Labeled l' a -> MAC l a
-- MAC monad
return   :: a -> MAC l a
(>>=)    :: MAC l a -> (a -> MAC l b) -> MAC l b
joinMAC  :: (l `CanFlowTo` l') => MAC l' a -> MAC l (Labeled l' a)
-- Run function
runMAC   :: MAC l a -> IO a
```

Figure 4: EDSL API

```
-- Types
newtype Labeled l a = MkLabeled a
data MAC l a where
  Label   :: (l `CanFlowTo` l') => Labeled l' a -> MAC l (Labeled l' a)
  Unlabel :: (l' `CanFlowTo` l) => Labeled l' a -> MAC l a
  Join    :: (l `CanFlowTo` l') => MAC l' a   -> MAC l (Labeled l' a)
  Return  :: a -> MAC l a
  Bind    :: MAC l a -> (a -> MAC l b) -> MAC l b
-- Labeled values
label    = Label . MkLabeled
unlabel  = Unlabel
-- MAC operations
joinMAC = Join
instance Monad (MAC l) where
  return = Return
  (>>=)  = Bind
-- Run function
runMAC (Label lv)               = return lv
runMAC (Unlabel (MkLabeled v))  = return v
runMAC (Join mac_a)             = runMAC mac_a >>= return . MkLabeled
runMAC (Return a)               = return a
runMAC (Bind mac f)             = runMAC mac >>= runMAC . f
```

Figure 5: Deep-embedded implemention

b) We would like to add the function *output* to the EDSL in order to print out messages. Ideally, we will have two output channels, one for public data and one for secret values. However, for simplicity, we assume that we have only one output channel: the screen. To mimic having two output channels, however, we will pre-append some text to indicate on which channel data is being sent. See the functions *add_location* and *print_cc* below.

```
-- outputting in a public channel
add_location :: Labeled L String → MAC L ()
add_location lstr = do
  str  ← unlabel lstr
  msg ← label (str ++ "Gothenburg")
           :: MAC L (Labeled L String)
  output msg
```

```
-- outputting in a secret channel
print_cc :: Labeled H Int → MAC H ()
print_cc lcc = do
  number ← unlabel lcc
  msg      ← label ("CC number "
                       ++ show number)
                   :: MAC H (Labeled H String)
  output msg
```

If we call *add_location* with a weather report, then it prints out a message in the public channel.

```
> let weather = MkLabeled "Sunny, 31 degrees, " :: Labeled L String
  in runMAC (add_location weather)
public channel : Sunny, 31 degrees, Gothenburg
```

By contrast, if we call *print_cc* with a credit card number, then it sends the credit card digits to the secret channel.

```
> let cc_number = MkLabeled 1234 :: Labeled H Int
  in runMAC (print_cc cc_number)
private channel : CC number 1234
```

Observe that the implementation of *output* depends on the type of the labeled value taken as argument, i.e. *output* is overloaded. Your task is to extend the definitions of *MAC l a*, (>>=), and *runMAC* to include the primitive *output* in the EDSL. *(8p)*